Hierarchical Accountability in Government: Theory and Evidence*

Razvan Vlaicu†
University of Maryland

Alexander Whalley‡
University of California - Merced
& NBER

February 2012

Abstract

How are the incentives of an indirectly-elected policymaker different from those of a directly-elected policymaker? Using a principal-monitor-agent model with adverse selection and moral hazard we show that having a monitor allows an uninformed principal to credibly insulate the agent from popular pressure, weakening his pandering incentives. However, this happens only when the agent's preference bias is low and/or the principal's policy perceptions are weak. Otherwise the principal preempts agent discretion by prescribing popular policies through the monitor. We test the model's predictions on a panel of U.S. cities, using a new instrument for manager government. We find that: (i) indirectly-elected city managers choose popular police employment policies less often than directly-elected mayors, and (ii) this police employment differential varies according to the model's political and informational mechanisms.

JEL Classification: D72, D73, H70.

Keywords: hierarchical accountability, incentives in government, pandering, U.S. city manager, police employment.

*We would like to thank Ryan Booth, Richard Boylan, Aimee Chin, Philip Curry, Jim Davies, Allan Drazen, Price Fishback, Lisa George, John Hatfield, Shawn Kantor, Ethan Kaplan, Trevor Kollmann, Katrina Kosec, Sumon Majumdar, Jean McGloin, Steeve Mongrain, James Svara, Gergely Ujhelyi, Emanuel Vespa, Joel Waldfogel, John Wallis and Jihong Zhao. Erin Moody, Krystal Tapia, Michael Wall and Marko Zivanovic provided excellent research assistance.

†Department of Economics and Department of Government and Politics, 3105 Tydings Hall, College Park, MD 20742, United States. E-mail: vlaicu@econ.umd.edu.

‡School of Social Sciences, Humanities & Arts, 5200 North Lake Road, Merced, CA 95343, United States. E-mail: awhalley@ucmerced.edu.

Electronic copy available at: http://ssrn.com/abstract=1925005

1 Introduction

A recent body of theoretical and empirical research in political economy has advanced our understanding of how elected policymakers - politicians - differ from non-elected policymakers - bureaucrats - in terms of either extrinsic motivation (incentives) or intrinsic motivation (type).[1] Less attention has been given to *indirectly-elected policymakers*, despite the pervasiveness of this accountability form. For instance, about half of U.S. cities operate under a manager charter, where the chief executive is hired/fired by a directly-elected city council and not by city voters. Unlike a politician, these policymakers are protected from removal by voters; unlike a typical bureaucrat, they are not protected by civil service tenure.[2]

Does an indirectly-elected policymaker face incentives more similar to those of a politician or a bureaucrat? (Stillman 1977) A simple intuition might say that indirectly-elected policymakers are more insulated from popular pressure. Persson, Roland and Tabellini (1997) argue that, because of possible collusion with parliament, an indirectly-elected prime minister has more discretion to pursue rent-seeking than a directly-elected president. Other arguments, however, do not support this view. Deno and Mehay (1987) argue that indirectly-elected city managers should track the median voter's preferences, just as the directly-elected city council to which they are accountable.

This paper studies the incentives of indirectly-elected policymakers. We propose a principal-monitor-agent model that uncovers political and informational conditions under which an indirectly-elected policymaker is insulated from popular pressure.[3] We then estimate policy differences between indirectly-elected and directly-elected U.S. city executives during 1960-2000, addressing the potential endogeneity of accountability form.

From the point of view of incentives, a key feature of indirectly-elected policymakers is that their continuation in office is decided by a directly-elected monitor. If the monitor's own continuation in office depends primarily on policymaker performance we say that the policymaker is under *hierarchical accountability*.[4]

[1] Examples include: Besley and Coate (2003), Maskin and Tirole (2004), Alesina and Tabellini (2007), Iaryczower, Lewis and Shum (2010).

[2] "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) Manager government is also used by U.S. counties, as well as cities and counties in Canada, Ireland, and other developed countries. Similarly, U.S. public school superintendents are accountable to a popularly-elected school board in about 80 percent of independent public school districts.

[3] The agency approach to electoral accountability was pioneered by Barro (1973) and Ferejohn (1986).

[4] Related forms of *indirect accountability*, where the monitor has additional responsibilities, include: U.S. senators elected by state legislatures (before the 17th Amendment of 1913, see the empirical analysis of Gailmard and Jenkins 2009); top bureaucrats, such as a central bank governor, selected by a popularly-elected president; federal agency heads monitored by Congress (Weingast and Moran 1983); village heads appointed by district heads in new democracies such as Indonesia (Martinez-Bravo 2011).

Electronic copy available at: http://ssrn.com/abstract=1925005

How is policymaker accountability to the voter mediated through the monitor? The agency literature suggests the answer should depend on the informational asymmetries among the players. If the voter is perfectly informed about policymaker type and policy optimality, he can transmit the same incentives though an reelection-seeking monitor that he would to the policymaker directly, by setting a policy-based performance standard. Thus, unless the monitor colludes with the policymaker, the monitor is irrelevant when the voter is perfectly informed (Persson, Roland and Tabellini 1997).

When the voter cannot perfectly observe policymaker type and policy optimality, however, his assessment of the policymaker has to be based on his *policy perceptions* i.e. his beliefs about optimal policy. This induces directly-elected policymakers to pander to voters, i.e. choose popular policies at the expense of optimal policies (Canes-Wrone, Herron and Shotts 2001). Even if the voter would like to commit to a perceptions-free performance standard this *commitment* is not credible to a directly-elected policymaker because once policy has been chosen the voter will rely on his perceptions to decide whether to keep the policymaker for another term.[5] In an environment plagued by policymaker pandering distortions non-elected officials such as life-tenured judges can outperform reelection-seeking politicians (Maskin and Tirole 2004); term limits can serve as a voter commitment device for reducing politician pandering (Smart and Sturm 2011); and media sometimes limits the informational distortion (Ashworth and Shotts 2010). Can a hierarchical agency relationship similarly alleviate the voter's inability to prevent pandering?

We study an infinitely-repeated principal-monitor-agent model where each period the policymaker chooses a policy, the monitor keeps or fires the policymaker, and the voter retains or replaces the monitor. We introduce commitment and pandering distortions by assuming (i) policymaker and monitor preferences are private information and can be either aligned with or biased away from the voter's preferences (adverse selection), and (ii) the voter is uncertain about policy optimality (moral hazard).[6]

The model's main insight is that, unlike direct accountability, where only perceptions-based voter strategies are credible, hierarchical accountability offers the voter a choice be-

[5] As explicitly noted by Besley and Smart (2007) and Smart and Sturm (2011) in the context of fiscal rules and term limits, respectively.

[6] Hierarchical agency (principal-supervisor-agent) models in contract theory and corporate finance also feature asymmetric information but use incentive structures richer than a binary retention decision. Strausz (1997) studies a model where either the principal or an intermediary can monitor the agent's hidden action. He finds that by delegating to an intermediary the principal can improve both incentives and his ability to commit to a broader range of wage structures. Park (2000) studies optimal firm debt structure under moral hazard and argues that the optimal contract should delegate monitoring to a single senior lender who is allowed to appropriate the full return from his monitoring activities.

tween setting a perceptions-based and a perceptions-free performance standard through the monitor. The first leads to policymaker pandering, while the second allows policymaker discretion. The perceptions-free strategy is now also credible because once policy has been chosen, the voter cannot fire the policymaker even if - according to his perceptions - he is unhappy with the policymaker. The voter can at most replace the monitor. Yet, because the monitor has incentives to transmit the voter's preferred standard the voter cannot benefit from removing the monitor after the fact.

The voter prefers to commit to the performance standard that maximizes his expected welfare.[7] We show that this choice depends on the policymaker's preference bias and the voter's policy perceptions. When the preference bias is low or voter perceptions are weak the voter prefers policymaker truthfulness, and thus will allow the policymaker discretion by prescribing the monitor to keep or fire the policymaker regardless of policy. When the preference bias is high or voter perceptions are strong, the voter prefers policymaker pandering, and thus the voter will preempt policymaker discretion by transmitting a perceptions-based standard through the monitor.[8]

We empirically test our baseline theoretical result of less pandering under hierarchical accountability, as well as these comparative statics, using city-level data from the U.S. There are several reasons why this setting is a suitable testing ground for our theory. First, a large number of U.S. cities are run by city managers, indirectly-elected chief executives with the same major policy responsibilities as directly-elected mayors, i.e. writing the city budget and hiring personnel. Second, because crime has consistently ranked among the top two concerns of city residents since Gallup started to survey local attitudes in 1959 (see Gallup 2000), police employment is susceptible to pandering incentives. We can thus test whether managers employ fewer police officers; for comparison, we can also test whether a similar differential exists for policies that do not evoke a broad popular preference, such as employment of police civilians, i.e. administrators, attorneys, dispatchers, etc.[9] Lastly, while the empirical local

[7] Ashworth, Bueno de Mesquita and Friedenberg (2010) use a similar criterion to select among the multiple equilibria of a two-period model with competence types.

[8] Our commitment argument is distinct from an information-extraction argument where by delegating monitoring the voter gleans more information about policy optimality, by observing the monitor's responses, and thus hold the policymaker to a more accurate standard of performance (Laffont 1999). It is also different from an argument where institutions help *policymakers* commit to welfare-enhancing policies, such as upholding property rights (North and Weingast 1989); in our agency setup the institution's role is rather to help imperfectly-informed *voters* commit to retention strategies that better incentivize policymakers.

[9] The empirical literature has placed disproportionate emphasis on differences in public spending between the two forms of city executive; results have been mixed. See Coate and Knight (2011) for references and for evidence that managers outspend mayors. Their paper is one of the few to provide a theoretical model of differences in fiscal behavior. Their citizen-candidate model attributes differences in spending not to incentives, but to voters electing different policymaker types. Two recent papers have found significant

government literature has treated city government form as exogenous, we address potential endogeneity in government form by exploiting precipitation shocks that influenced early 20th century (Progressive Era) city charter reforms for reasons that are obsolete today.[10]

We find that managers employ fewer police officers per capita but a comparable number of police civilians per capita. These two patterns are robust across a number of specifications. First, the results hold up when we control for an extensive array of geographic, demographic, and institutional factors, including related Progressive reforms. Second, the results are robust when we instrument for government form with precipitation shocks to address potential reverse causality. Third, the pattern survives extensive sensitivity analysis.

We also find supportive evidence for the model's political and informational mechanisms. The officer employment differential increases in the value of political office, measured by city size, and decreases in interest group power, measured by high police unionization. The officer employment differential decreases in the strength of voter crime perceptions, measured by local newspaper sales, and is more pronounced in election years. These patterns are consistent with differences in incentives driven by local agency institutions. Our extensions to the theory model indicate that similar patterns would not be observed under alternative explanations for manager-mayor policy differences, such as patronage motivations, professional peer accountability and policymaker type selection.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes its equilibrium. Section 4 provides a historical background and introduces the data. Section 5 presents the empirical strategy and results. Section 6 concludes.

2 Model

We propose an infinite-horizon agency model with preference types. Our environment features two informational problems: incomplete information about agent policy preferences, i.e. adverse selection, and imperfect information about policy optimality, i.e. moral hazard.

In every period t, where $t = 1, 2, ...$, a community needs to make a policy choice x_t. The set of policy alternatives is $\{0, 1\}$. To illustrate, suppose that the community is a city and the policy issue at stake is law enforcement. Then $x_t = 1$ can be thought of as a policy of "high police," and $x_t = 0$ represents "low police." The model admits alternative interpretations,

differences between managers and mayors in areas other than public spending. Levin and Tadelis (2010) find that managers are more likely to privatize city services compared to mayors. Enikopolov (2010) finds that managers reduce full-time city employment, but have no effect on part-time employment.

[10] An exception is the cross-sectional analysis in Baqir (2002) who uses past city government form as an instrument for current government form. The focus in that paper, however, is on city council size.

depending on the application.

In every period t the community can be in one of two states of the world $s_t \in \{0, 1\}$. In our illustration state $s_t = 1$ can represent "high crime" and state $s_t = 0$ can represent "low crime." The state is identically and independently distributed across periods. Let $p = \mathbb{P}\{s_t = 1\}$ be the probability of the "high crime" state. Assume $\frac{1}{2} < p < 1$. The parameter p captures the strength of the community's perceptions of the high state.[11]

Players and Preferences. There are three players: a voter, a monitor, and a policymaker. The voter gets a unit of payoff when policy matches the state, zero otherwise: $v_t(x_t, s_t) = \mathbf{1}\{x_t = s_t\}$. Since when uninformed the voter perceives the high state $s_t = 1$ as more likely, an uninformed voter prefers policy $x_t = 1$ in expectation. By being the voter's preferred policy in the most likely state of the world, "high police" is the *popular* policy.[12]

A monitor and a policymaker can be either aligned with the voter or not. That depends on their type: "good" or "bad." The proportions of good monitor and policymaker types are μ_t and π_t respectively, where $0 < \mu_t, \pi_t < 1$. Ex ante, before any choices have been made, monitor and policymaker types are independent of each other.

A good type G shares the voter's preferences: $m_t^G(x_t, s_t) = p_t^G(x_t, s_t) = v_t(x_t, s_t)$. He wants to match the policy to the state. A bad type B wants to instead match the policy to an appeal: $m_t^G(x_t, l_t) = p_t^G(x_t, l_t) = \mathbf{1}\{x_t = l_t\}$. The appeal l_t always matches the low state but sometimes mismatches the high state: $\mathbb{P}\{l_t = 0 | s_t = 0\} = 1$ but $\mathbb{P}\{l_t = 1 | s_t = 1\} = 1 - b$, where $0 < b < 1$. We refer to a state-appeal pair (s_t, l_t) as a "contingency."

The parameter b can be thought of as the "bias" a bad type holds against the popular policy. The source of this bias can be the presence of an interest group, such as a public sector union, that is able to influence policymaker and/or monitor preferences through political

[11] 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 (Gallup 2010). The FBI's *Uniform Crime Reports*, in contrast, show a declining trend in crime starting in the early 1990s. To keep the model simple we assume perceptions are based on the actual crime distribution. For our interpretations to go through we only need that voter crime perceptions increase in the mean of the actual crime distribution.

[12] In U.S. cities this popular preference is driven by crime perceptions (see previous footnote) together with a belief that police reduces crime (see the literature linking "fear of crime" to "police strength" reviewed in Stucky 2005). Additional support for this assumption comes from the American National Elections Study (ANES). The six editions from 1966 through 1976 contain "thermometer" questions measuring public sentiment toward a dozen or so social groups, among them policemen. On a scale from 0 to 100, with 50 measuring a "neutral" feeling, policemen's reading is between 75.22-80.46, with an average over the six surveys of 78.17. By comparison, lawyers score 65.38, city/county officials 63.26, and politicians 52.76. From a different angle, an ABC News/*Washington Post* national poll from March 2011, on possible solutions to the current state and local budget crises, finds that popular opposition to cutbacks was "broadest and deepest" for firefighters, teachers, and police officers, with 89 (75), 86 (74), 86 (70) percent, respectively, opposing (strongly opposing) cuts.

means.[13] One could also assume bad types are biased in favor of the popular policy; however, an interest group appeal against the popular policy matches our empirical context where police unions in U.S. cities have been shown to favor pay and benefits over high employment.[14]

Timing. The game starts in period $t = 1$ with an incumbent monitor and an incumbent policymaker in office. Their type is good with probability $\mu_1 := \mu$ and $\pi_1 := \pi$ respectively. The game consists of the infinite repetition of the following stage game.

1) The policymaker observes the state s_t, the appeal l_t, and the monitor's type. The policymaker chooses a policy x_t.

2) The monitor observes the state s_t, the appeal l_t, and the policy choice x_t. The monitor does not observe the policymaker's type. The monitor makes a firing decision y_t: keep or fire the policymaker.

3) The voter observes the policy choice x_t and the firing decision y_t. The voter does not observe the state, the appeal, the policymaker type or the monitor type. The voter makes a replacement decision: retain or replace the monitor.

Every period an exogenous event may occur that forces an incumbent to retire by the end of the period. This becomes known to all players at the beginning of each period. In those cases the monitor and the policymaker make decisions as lame-ducks. In every period the probability that the incumbent seeks to remain in office is σ^M for the monitor and σ^P for the policymaker, where $0 < \sigma^M, \sigma^P < 1$. They receive payoffs only when in office and, once out, cannot run for office again. The probabilities σ^M and σ^P thus function as implicit discounting factors. The voter discounts future payoffs by a factor β, where $\beta < 1$.[15]

A policymaker thus steps down either because he is fired or because an exogenous event forces him to retire. In either case he is replaced with a challenger who is a good type with probability $\bar{\pi}_{t+1} = \pi_t$. In words, the challenger's reputation is equal to the incumbent's prior reputation. An analogous assumption is made for the monitor.[16]

Notice how in both principal-agent relationships, voter-monitor and monitor-policymaker,

[13] For theories of interest group influence through information and campaign contributions see e.g. Baron (1994) and Prat (2002).

[14] See e.g. Rynecki and Morse (1981) and Carter and Sapp (1992). Trejo (1991) and Valletta (1993) find positive wage effects and negative employment effects of police unionization in U.S. cities.

[15] An infinite-horizon model allows office value to be endogenous to the accountability form. In a two-period model like Maskin and Tirole (2004) and Vlaicu (2008) office value (ego rent) is an exogenous parameter.

[16] This assumption captures the intuition that the challanger pool adjusts to the prior reputation of the current incumbent, e.g. an initially strong incumbent discourages weak challengers from running. Technically, the assumption helps make the environment stationary. See Besley (2006) for a discussion of stationarity in infinite-horizon political agency models.

the agent has more information than the principal. Laffont (1999) and Gailmard and Jenkins (2009) provide arguments in support of the view that accountable representatives have more information than the public. While it is less clear what the monitor-policymaker informational asymmetry is, in the next section we argue that alternative assumptions about this informational asymmetry do not change the equilibrium of the game.

The model naturally allows a comparison of hierarchical to direct accountability by dropping the monitor from the model and allowing the voter to retain or remove the policymaker directly. The informational asymmetry between voter and policymaker remains the same: the policymaker knows the state but the voter does not and will never learn it. To reiterate, in the context of our application the assumption says that in every electoral term the voter observes the level of police, high or low, but the voter cannot tell if a different level would have been a better policy.

Equilibrium Concept. The game defined above is a dynamic game of incomplete information, because every period the structure of the game depends on policymaker and monitor reputations (π_t, μ_t). This pair of variables is the natural game-state. We adopt the widely-used Markov Perfect Equilibrium (MPE) concept. In this equilibrium strategies depend only on the current period's game-state and actions.

The equilibrium consists of Markovian strategies for the voter, monitor and policymaker, as well as voter and monitor beliefs that satisfy the following conditions (we restrict attention to pure strategies):

(i) The voter's replacement strategy maximizes his lifetime expected payoff and is sequentially rational, given other players' strategies.

(ii) The monitor's firing strategy maximizes his lifetime expected payoff and is sequentially rational, given other players' strategies.

(iii) The policymaker's strategy is a best response to monitor and voter strategies.

(iv) Voter and monitor beliefs are consistent with equilibrium strategies on the equilibrium path. We impose no restrictions on off-equilibrium-path beliefs.[17]

In the next section we show that the *direct accountability* MPE is stationary, i.e. strategies depend on current period observables and not on the game-state π_t (one-dimensional in this case, since there is no monitor). This is because the voter's strategy depends on

[17]The voter can be thought of as a principal who offers an implicit contract to his agent. Sequential rationality then requires that the contract be optimal not only ex ante, but also ex post. This interpretation is common in the literature, see e.g. Ferejohn (1986) and Ashworth, Bueno de Mesquita and Friedenberg (2010). It implies that if there are multiple subgame-perfect retention strategies the voter will play the one that maximizes his ex ante welfare. We apply the same interpretation to the monitor-as-principal.

beginning-of-period policymaker reputation π_t only through the current period action x_t. Even if the voter remembers a longer history those actions become irrelevant for the current retention decision, since they are already incorporated in the challenger pool composition. To facilitate comparison, we adopt a stationary MPE concept for the *hierarchical accountability* game, i.e. strategies depend on current period observables and not on the game-state (π_t, μ_t); as we will see this game also admits non-stationary equilibria.

3 Equilibrium

This section characterizes the incentives of an indirectly-elected policymaker and compares them with those of a directly-elected policymaker. It then studies several extensions that help to empirically distinguish the incentives model from alternative explanations for differential policymaking behavior: neutral policy issues (e.g. civilian police employment), salary effects, policymaker type selection, and electoral effects.

Hierarchical vs. Direct Incentives. When the voter can directly replace the policymaker, he has four possible pure firing strategies: two perceptions-based ("keep iff popular policy," "keep iff unpopular policy"), and two perceptions-free ("keep no matter what," "fire no matter what"). Which strategy is both optimal for the voter and credible to the policymaker?

The perceptions-free strategies allow the policymaker discretion, by either insulating or term-limiting him, and thus create incentives for policymaker truthfulness. However they are not credible, because when policymakers are truthful a popular policy signals (i.e. is correlated with) a good type, and so ex post the voter will not be indifferent between a popular incumbent and the challenger.[18] The perceptions-based strategies, on the other hand, are credible because they create in both policymaker types incentives to respond to the voter's perceptions-based performance standard. Of the two policies the voter prefers to prescribe the popular policy since he perceives it more likely to be optimal (by the assumption that $p > \frac{1}{2}$). Thus in equilibrium both policymaker types pander, i.e. choose the popular policy even when know it is not the optimal policy.[19]

A key feature of direct accountability in this informational environment is that it can leave the voter worse off than lack of accountability (Maskin and Tirole 2004). Direct

[18] If policymakers are truthful, a popular policy improves the policymaker's reputation $\tilde{\pi}_t(x_t = 1) = \frac{\pi_t p}{\pi_t p + (1-\pi_t)p(1-b)} > \pi_t$, and an unpopular policy hurts the policymaker's reputation $\tilde{\pi}_t(x_t = 0) < \pi_t$.

[19] While closely related to the concept of populism (Acemoglu, Egorov and Sonin 2011), pandering differs in at least two respects: first, it is not necessarily associated with economic inequality and redistribution, and second, it affects politicians of all ideological stripes, not only those who are naturally inclined to serve the cause of the "common man," e.g. working-class politicians.

accountability creates pandering incentives, which generate p for the voter. Policymaker discretion generates $\pi + (1-\pi)(1-bp)$. Pandering is worse for the voter when:

$$p < \pi + (1-\pi)(1-bp) \tag{1}$$

which holds when the policymaker bias b is low or voter perceptions p are weak. To prevent pandering the voter would have to either always keep or always fire the policymaker. However these strategies are not sequentially rational for the voter, and therefore not credible to the policymaker.

Does the presence of a monitor alleviate or exacerbate the pandering distortion? A monitor faces the same problem as the voter above: choose a firing rule to give the policymaker appropriate incentives. Because the monitor is informed he could induce the policymaker to choose the monitor's preferred policy, by the logic of direct incentives.

However, the voter may want to influence the monitor's strategy by prescribing a monitor response to each policymaker choice. The monitor is more compliant with a voter prescription than a policymaker would be (Lemma 1(ii) in the Theory Appendix) because unlike a policymaker the monitor cannot affect the policy choice. For example, if the voter prescribes the monitor to "keep after popular policy and fire after unpopular policy," once the policymaker has chosen the popular policy the monitor's only profitable response is to comply and keep the policymaker.

Since the monitor is so compliant, the voter can always "dictate" to the monitor any of the four possible policymaker firing rules. What is interesting is that even though only the two perceptions-based rules were credible under direct accountability, here the voter can also credibly transmit the two perceptions-free rules. The reason is that both monitor types comply with the voter's prescription and so the voter cannot benefit from replacing the monitor after the fact. The voter thus has a choice. If he prefers pandering to truthfulness he will transmit pandering incentives; otherwise, the voter can create truthfulness incentives.[20]

Can the voter do better than simply dictating a policymaker firing rule to the monitor? It seems that one way the voter could improve on policymaker pandering and truthfulness is a partially-constraining monitor replacement rule: induce the monitor to reward the popular policy, but leave the monitor unconstrained as to how to respond to the unpopular policy. Presumably this will still prevent a bad policymaker from mismatching the high state, while

[20] If $1-bp < p < \pi + (1-\pi)(1-bp)$ both insulating and term-limiting are optimal but the insulating strategy is no longer stationary: the policymaker's reputation $\tilde{\pi}_t$ evolves dynamically with his policy choices and if it declines to a point where $\tilde{\pi}_t + (1-\tilde{\pi}_t)(1-bp) < p$ the voter prefers to switch to a pandering equilibrium thus increasing expected voter welfare.

not forcing either policymaker into mismatching the low state. However, this voter strategy is not credible because a bad monitor is more willing to keep an unpopular policymaker prompting the voter to ex post punish such a monitor response.

Proposition 1 *A directly-elected policymaker has the incentive to pander to the voter: he chooses the popular policy in every contingency.*

An indirectly-elected policymaker has: (i) pandering incentives, if policymaker bias is high or voter perceptions are strong; or (ii) truthfulness incentives, if policymaker bias is low or voter perceptions are weak.

The key implication of this result is that hierarchical accountability allows for truthfulness incentives not possible under direct accountability. The voter can promote truthfulness by prescribing the monitor to either insulate ("keep no matter what") or term-limit ("fire no matter what") the policymaker.[21] The voter prefers these strategies when equation (1) holds, i.e. either the policymaker bias b against the popular policy is low or the voter's perceptions p of the high state are weak. These two factors thus determine the extent to which indirectly-elected policymakers choose popular policies less frequently than directly-elected policymakers. A third factor is the retention-seeking motivation σ^P: all else equal, a stronger retention motivation amplifies the policy differential.

We emphasize that the voter's benefit from having a monitor is not more information extraction, as when there are checks and balances (Persson, Roland and Tabellini 1997, Laffont 1999), or when media commentary is informative (Ashworth and Shotts 2010). The benefit of hierarchical accountability here is to allow credible voter commitment to a perceptions-free performance standard. Under appropriate political and informational conditions the voter uses his enhanced commitment ability to allow an indirectly-elected policymaker to behave like an "independent bureaucrat" in the sense of Alesina and Tabellini (2007), who acts not in response to popular pressure, but follows his own preferences and information.

Just as the policymaker cannot act on his state information in the direct accountability equilibrium, so in the hierarchical accountability equilibrium the monitor cannot act on his state information. In fact, the hierarchical equilibrium in Proposition 1 does not depend on monitor reputation μ_t. This implies that this equilibrium is robust to alternative assumptions about monitor information, e.g. the monitor knows less about the state than

[21] While we do not provide direct evidence of these types of voter and monitor strategies, due to lack of comprehensive data on manager tenure, an early survey of 48 U.S. cities with manager charters found that "twenty-two city managers in the forty-eight cities studied served a single city for ten or more years" and "some cities had a high rate of turnover in managers." (Stone, Price and Stone 1940).

the policymaker. The hierarchical accountability equilibrium is also robust to assuming the monitor knows the policymaker's type. To see this notice that in the insulating truthful equilibrium the monitor literally learns the policymaker's type in the disagreement contingency (high state, low appeal), since policymakers separate in that contingency. Yet, despite knowing the policymaker's type, the monitor has no incentive to fire an opposite-type policymaker, because the monitor is constrained by the voter's tight replacement rule prescribing the monitor to keep the policymaker no matter what

In the Theory Appendix we show that Proposition 1 is also robust to the possibility that the voter learns the state before the election. When the voter is informed he no longer prescribes the popular policy, but the optimal policy. In a pandering equilibrium the policymaker's incentive to adopt popular policies is then counteracted in the low state and reinforced in the high state. The policymaker is still pandering if the informed incentive is dominated by the uninformed incentive. In a truthful equilibrium the policymaker is unconstrained by an uninformed voter, whereas an informed voter pushes him to adopt optimal policies. This informed incentive may affect only a bad type in the disagreement contingency (high state, low appeal), as long as it is strong enough.[22]

Neutral Policy Issues. 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 the voter. While patronage could potentially affect popular policy issues, such as crime fighting or low taxes, it also typically influences policies such as in-kind transfers or city hall employment whose provision levels do not evoke a broad popular preference. For example, a large city bureaucracy may indicate a thriving city or may convey waste. What does our model imply for policymaker incentives on neutral issues?

Suppose that in the baseline model the bad types' bias affects both states with equal frequency $b = \mathbb{P}\{l_t = 1 | s_t = 0\} = \mathbb{P}\{l_t = 0 | s_t = 1\}$ and $p = \frac{1}{2}$. In this setup neither policy signals a bad type, since both good and bad types prefer any given policy with the same frequency, although at different times. Thus the voter can directly induce both pandering and truthful equilibria. Which one he plays depends on his welfare. In a pandering equilibrium the voter gets the optimal policy half of the time; in a truthful equilibrium he gets it with frequency $\pi + (1 - \pi)(1 - b)$. The voter's choice then depends on policymaker bias b.

Proposition 2 *On neutral policy issues both indirectly-elected and directly-elected policy-*

[22] We note that the truthful equilibrium is more fragile to the arrival of state information than the pandering equilibrium. That is because the truthful equilibrium survives only when the informed incentive is weak, whereas the pandering equilibrium survives when the informed incentive is weaker than the uninformed incentive.

makers have: (i) pandering incentives, if policymaker bias is high, and (ii) truthful incentives, if policymaker bias is low.

The result implies that, all else equal, incentives on neutral issues should not differ by accountability form. A patronage motivation, in contrast, would imply that indirectly-elected policymakers deliver less patronage on neutral issues as well.

Policymaker Salary. In many governments the salaries paid to executive officeholders may differ with their accountability form. What are the policy consequences of policymaker salary differences? We address this question using a model of exogenous salary determining candidate entry similar to Besley (2004).

Let Π be the fraction of good citizens in the population.[23] Let w denote the public office salary. Suppose the private-sector compensations available to good types $w^G \sim U[0, W^G]$ and bad types $w^B \sim U[0, W^B]$ are distributed uniform. A "candidate" is a citizen who runs for public office, i.e. his private-sector compensation is below public office value. A "policymaker" is a candidate who wins. Assume that all candidates have an equal chance of winning. This model generates the following prediction.

Proposition 3 *If policymakers act on incentives, a higher salary weakly increases the frequency of the popular policy.*

In this model citizens have two motivations for seeking public office: a salary motivation and a policy motivation. The first is constant across types. The second differs across types: pandering makes good types more attracted to public office because they value the popular policy more than bad types. As the salary rises, the salary motivation starts to dominate and increases the bad types' initially lower draw to office comparatively more. Thus the fraction of bad candidates, and therefore policymakers, increases. A larger fraction of bad types does not affect the uniqueness of the direct accountability pandering equilibrium, but it can make the pandering equilibrium more appealing to the voter, by equation (1), and thus increase its frequency under hierarchical accountability.[24]

Hierarchical vs. Direct Selection. Indirectly-elected policymakers may choose different policies than directly-elected ones not because they have different incentives, but because

[23] Assume $\Pi > \frac{1}{2}$ so that a good policymaker has the same preferences as the median voter.

[24] In this type of model the sign of the candidate selection effect critically depends on the comparison between good and bad types' office motivations. In Besley (2004), in contrast to our model, bad types' ability to extract rents makes their office motivation relatively stronger. Hence his result that a higher salary appeals to good candidates more.

they are different preference types. One reason for type differences is that the accountability form itself may influence how effectively bad policymaker types are screened out. To explore this possibility, consider a pure selection model where policymakers and monitors act on their preferences, rather than on incentives. In this scenario choices signal types and the voter is able to gradually screen out bad types.

Proposition 4 *Hierarchical selection is less effective than direct selection at screening out bad policymaker types. The selection-driven policy differential increases in: (i) policymaker retention motivation, (ii) policymaker bias, and (iii) voter perceptions.*

Under direct selection the quality of the policymaker pool increases monotonically over time; in the limit all bad types are screened out. Under hierarchical selection the quality of the monitor pool also increases monotonically over time; in the limit all bad monitors are screened out. The quality of the policymaker pool may initially decrease (as long as monitor reputation $\mu_t < \frac{1}{2}$), but as μ_t increases above a half it starts to increase monotonically, although at a slower rate than under direct accountability. The reason is that a bad monitor performs "reverse screening" keeping bad policymakers and firing good ones.

Therefore, if policy were driven purely by preferences then indirectly-elected policymakers would choose the popular policy less often than directly-elected policymakers, by virtue of being more often bad types, i.e. biased against the popular policy. While this pure selection model yields the same qualitative prediction as Proposition 1, namely lower frequency of the popular policy under hierarchical accountability, comparative statics (ii) and (iii) differ. The reason is that here the policy differential is generated by a higher prevalence of bad policymaker types, and thus is increasing in their bias and the opportunities to express it.[25]

Electoral Effects. The notion that pandering incentives distort policymaking leads one to suspect that this behavior is more common around election time. Indeed, there is abundant evidence that various government policies display electoral cycles.[26] This may be due to the fact that voter information varies within an electoral term. To explore the effect of election proximity consider the following simple extension of the model. Suppose that each period t has two sub-periods, called early term (t') and late term (t''). The state is drawn anew every sub-period. The policymaker chooses policy once in early term and again in late term, and the voter observes both before the election.

[25] In the Theory Appendix we show that the selection-driven policy differential is $(\pi_H - \pi_D)\, bp$, which, by Proposition 4, is negative and widens in both b and p.

[26] The evidence is mostly at the national and state level, and is generally weaker for developed countries. For references see Drazen (2000) and Dahlberg and Mork (2011).

We follow Rogoff (1990) and Shi and Svensson (2006) in assuming that late term policy choices contain more information about policymaker type. They model this dynamic informational asymmetry by letting policymaker type evolve according to an AR(1) process. In such a model information signaled through the policy choice in early term t' is irrelevant for the voter in late term t'' because $\mathbb{E}\left[\theta_{(t+1)'}|\theta_{t'}\right] = \mathbb{E}\left[\theta_{(t+1)'}\right]$, i.e. next period policymaker type is mean independent of type early last period t'. We use a related assumption, namely that the incumbent's type "resets" in the middle of each term, implying that policymaker type early next period $\theta_{(t+1)'}$ is statistically independent of type early last period $\theta_{t'}$.[27] The incentives characterized in Proposition 1 imply the following.

Proposition 5 *In election years indirectly-elected policymakers pander less often than directly-elected policymakers. In off-election years both policymakers have truthfulness incentives.*

The intuition is that while late term behavior is driven by voter-induced incentives, early term policy choices do not affect future retention, and so both indirectly- and directly-elected policymakers behave truthfully. Notice also that the voter is not able to screen out bad types between terms, since both types either responded to incentives, were insulated, or term-limited. Thus early next term the ratio of good to bad types should be unchanged.

If, on the other hand, behavior is driven by preferences rather than incentives, then before the election no selection has yet occurred. Thus, on average there should be no differences in policymaker types between accountability forms. After the election, however, each accountability form performs selection according to Proposition 4, namely hierarchical selection is less effective at screening out bad policymaker types.

Proposition 6 *All else equal, in election years indirectly-elected and directly-elected policymakers have the same expected type. In off-election years indirectly-elected policymakers's type is worse than the directly-elected's type.*

Propositions 5 and 6 have opposing implications for how policy is correlated with accountability form along an electoral term. If incentives are at play, then we should observe a negative policy differential in election years, and no differential in off-election years. If selection is at play, then we should observe no policy differential in election years, but a negative policy differential in off-election years.

[27] One way to think about this assumption is that in the second half of a term the incumbent starts to seek support for the upcoming election campaign. He may end up getting the support of groups different than the ones that endorsed him the last time around. This will change his preferences for the second half of his term, and if retained, for the first half of next term. For a different approach to modeling electoral cycles see Martinez (2009).

4 Empirical Application

This section starts by stating the testable hypotheses. It then summarizes the historical facts that motivate our empirical identification strategies. Finally, it discusses our choices of empirical measures for the key theoretical variables.

Empirical Hypotheses. Our theory model implies that hierarchical accountability reduces pandering incentives relative to direct accountability. It also uncovers factors that determine this difference in incentives. We study these predictions empirically in the context of 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). The testable hypotheses in this context are:

(H1) On average managers employ fewer officers than mayors. There is no difference in civilian police employment.

(H2) The officer employment differential: (i) increases in policymaker retention motivation, (ii) decreases in policymaker bias, (iii) decreases when voter crime perceptions are stronger.

(H3) The officer employment differential is more pronounced in election years than in off-election years.

Policy differences between managers and mayors could be generated by factors other than agency-driven incentives. In the next section we interpret our findings against the following alternative explanations: differences in patronage motivations, concurrent political institutions, differences in salary and personal characteristics, professional peer accountability, and selection of preference types.

Historical Background. Thirty years after the ratification of the U.S. Constitution almost all U.S. city chief executives were appointed by state governors or city councils. After Boston and St. Louis started to popularly elect their mayors in 1822, the practice spread quickly to other cities. By 1840 the popular election of city mayors had become nearly universal.[28]

City politics toward the end of the 19th century became highly contentious due to deepening social and ethnic cleavages created by rapid industrialization and massive immigration. The city halls of most large cities became dominated by "machine" politicians drawing their support from workers and immigrants. Patronage, corruption and electoral fraud became

[28] Historians note that the trend toward direct elections coincided with the explosive growth of cities - at rates more than twice as high as the overall U.S. population in the 1820s-1840s - and the emergence of a mass electorate (Judd and Swanstrom 2010).

common strategies to preserve public office. This crisis in city hall accountability sparked an urban reform movement led by local business leaders, social and political elites. The movement coincided with broader social reforms taking place during the Progressive Era (1890-1930). At first municipal reformers targeted not government form, but cities' electoral systems: from district-based partisan to at-large non-partisan elections. The goal was to dilute the power of minorities and parties on which the machines thrived.

In the middle of the Progressive Era another wave of reform called the "efficiency movement" introduced organizational changes in business corporations, in particular a hierarchical administrative structure. The efficiency agenda was a response to the growing complexity of managing larger and larger organizations. The Progressives, who wanted not only cleaner government but also more efficient government, were inspired by the success of the new corporate form and sought to apply the model to city governance. Manager government, first experimented with in the small city of Staunton, Virginia in 1908, soon became part of the municipal reformers' program. The manager plan attained broad recognition when the National Municipal League made it their recommended government form in the 1915 edition of the *Model City Charter*. During the 1910s and 1920s most major cities, including New York City, were debating switching to the manager plan.

Despite strong opposition from incumbent mayors and party bosses the manager plan advanced steadily.[29] Knoke (1982) attributes successful switches to the manager plan to strong business interests, weak unions, high population mobility, small immigrant population, and small city size. As these factors are likely to persist and independently affect policy today, empirically identifying the effect of the manager plan seems challenging.

The manager plan, however, also filled a growing need for technical expertise at the top of city government, a need felt more acutely in times of crisis. The technological boom of the early 20th century had created a demand for technologically-intensive public services, such as paved roads, streetlights and streetcars, levees and sewer systems. The new infrastructure often amplified a crisis. Dayton, Ohio, among the thirty largest U.S. cities in 1900, 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 plan so that an engineer could be appointed to lead the reconstruction effort. After Dayton's first choice for the position – George Washington Goethals, the engineer overseeing the Panama Canal – declined, the

[29] A number of 87 cities adopted a manager charter between 1913-1918, another 153 between 1918-1923, and 84 more between 1923-1928 (Judd and Swanstrom 2010).

engineer Henry M. Waite became the city's first manager.[30]

After the Ohio Flood subsequent crises such as the Great Mississippi Flood of 1927 and the Northeast Flood of 1936 resulted in substantial losses across multiple local jurisdictions and helped swing the balance toward federal takeover of flood control from local authorities. In 1936 Congress passed the Flood Control Act (FCA) that assigned responsibility for flood prevention and management to the Army Corps of Engineers. The hundreds of miles of levees and 375 major reservoirs constructed by the Army Corps of Engineers after 1936 significantly weakened the link between heavy precipitation and the incidence of floods (Arnold 1988).

The connections between natural hazards, infrastructure crises and public demand for technical expertise, combined with the federal takeover of flood control in 1936, suggest an instrumental variable identification strategy: using precipitation shocks during the local flood control era (1900-1936), when they were more likely to have triggered an infrastructure crisis, to isolate exogenous variation in manager government.[31]

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.[32] The sample period for our panel is 1960-2000.

City government form is collected by the International City/County Management Association (ICMA) through surveys sent every five years to municipal officials. The survey results are reported in their *Municipal Year Book*.[33] We point out two features of government form

[30]The voters passed the new charter in August 1913. A contemprary report of the events read: "Even before the disastrous flood of March the feeling was strong for a change of administration, and after the catastrophe, when the city was in dire distress, there came an impelling force of sentiment which paved the way for the change that is now a reality." ("City Manager Takes Charge of Dayton," *New York Times*, January 2, 1914, p. 5)

[31]We subject our instrument to formal validity and exogeneity tests in the next section. Here we note that trends in city managers' educational backgrounds provide additional support for our identification strategy. By 1918, ten years after the first manager government, 95 percent of city managers were engineers. Twenty years later when the new government form had become commonplace, 75 percent of city managers had an engineering degree (Stone, Price, and Stone 1944, p. 265). The shift in flood control from the local to the federal government coincided with a dramatic change in city manager expertise. By 1971 only 33 percent, and typically the older city managers, had degrees in engineering; 39 percent had degrees in liberal arts and 12 percent in public administration. Stillman (1974) notes: "The turbulence of the cities has increasingly caused managers to turn to sociologists and political scientists for answers to complex urban issues." (p. 89)

[32]Two intervening annexations and one merger slightly alter the 1900 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.

[33]If government form changes between survey years, we date it using newspaper sources or city charters.

variation in our sample. First, the maps in Figures 1 and 2 display little geographic clustering in manager government. Second, as Figure 3 shows, most of the time-series variation in manager plan adoption occurs 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*. Police per capita displays wide variation both across and within cities.

The adoption of the manager plan during the Progressive Era often followed a city's electoral system reform. To separate the effect of manager government from at-large and non-partisan electoral reforms, we include these institutions in the analysis as controls. The data comes from the *Municipal Year Book*.[34]

The incentive effect we have characterized theoretically works through two political mechanisms (retention motivation and policymaker bias) and two informational mechanisms (voter perceptions and election proximity). We measure the strength of the retention motivation with city size, available in the U.S. *Census of Population*, based on the idea that the top office in larger cities is more valuable in terms of prestige and compensation. We measure policymaker bias toward an interest group using the fraction of police that are members of a nationally-affiliated union or employee association, available in the 1968 edition of the *Municipal Year Book*. An alternative measure of police unionization is the presence of state-level collective-bargaining laws pertaining to local governments, as coded by Booth and Vespa (2011). For informational mechanisms we use sales of local newspapers, from George and Waldfogel (2006), as a measure of voter crime perceptions and election years as a measure of periods when voter perceptions more likely influence voter behavior and therefore policymaker incentives.[35] City elections data are not available on a systematic basis in any one source and discrepancies among existing sources are prevalent. To reduce errors we corroborated multiple sources, as described in the Online Appendix.

We construct our instrumental variable 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 rainfall, snowfall, and temperature extremes collected from

[34] We also include a measure of early civil service reform, i.e. as of 1937. With the introduction of Social Security in 1939 essentially all local governments set up a civil service agency as it was required as a condition for disbursing Social Security payments to city residents.

[35] Heath and Gilbert (1996) review research on the link between newspaper consumption and fear of crime.

weather stations throughout the U.S. We match a city to the closest weather station based on geographic coordinates.[36] Our main precipitation measure is the yearly sum of rainfall and density-adjusted snowfall at the station level.[37] 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 characteristics controls. The Online Appendix provides the complete list of variables, with details about their sources and measurement.

Table 1 reports summary statistics for our major variables, overall and 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 E suggests that before the advent of the manager plan there were no major differences in either type of police employment. The other statistically significant differences are as follows: manager cities have on average longer distance to nearest river, higher education, higher property crime, higher incidence of Progressive electoral institutions, less state-protected collective bargaining, higher policymaker salary.[38]

Figure 4 provides a first look at how manager governments have affected police employment historically. Manager governments employ fewer officers from 1960 onward, but not before; the same cannot be said for civilians. In the bottom panel we divide cities based on whether they experienced LFC precipitation shocks or not. Cities hit by manager-inducing shocks have lower officer employment after 1960 though the difference attenuates over time.[39] Again, civilian employment does not follow this same pattern.

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

[37] 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.

[38] Our model implies that, if police reduces crime, crime would be higher under manager government. While Table 1 supports this conclusion for property crime rates, the numbers for violent crime rates appear to contradict it. In Table A6 of the Online Appendix we estimate the relationship between manager government and crime rates controlling for year effects, geographic and demographic factors. Point estimates are positive for both property and violent crime.

[39] During our sample period a number of national trends reduced local discretion over police employment. These include court-ordered affirmative action hiring quotas for police departments beginning in 1969 and federal grants to fund police hiring through the Community Oriented Police Services (COPS) program established by Congress in 1994.

5 Empirical Strategy and Results

We start with a baseline model, study its robustness and sensitivity, and then test for the political and informational mechanisms of the theory model.

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

$$\log(Police_{i,t}) = \beta_1 Manager_{i,t} + \beta_2 X_{i,t} + \tau_t + \epsilon_{i,t} \tag{2}$$

where $Police_{i,t}$ is either officer or civilian employment 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, τ_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. Hypothesis (H1) says that $\beta_1 < 0$ 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, 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.

Table 2 presents OLS estimates of β_1. Accounting only for year effects column (1) shows that manager governments employ significantly fewer officers. In columns (2)-(4) we use within-Census-division variation together with an increasing set of controls. The officer differential varies between -6% and -16% depending on how we account for potential confounds. For civilians we find little evidence of a similar statistically significant relationship with government form. The point estimates in columns (5) to (8) are positive, closer to zero and not statistically significant.[40]

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) implies a 20.1% upper bound on his effect, corresponding to a move from zero penetration to full penetration. Our -8% point estimate of the officer differential in Table 2 column (2) is equivalent to about 7% of police department and 0.5% of city government spending respectively.[41]

[40] In unreported analysis we find that these results are robust to controlling for city crime rates.

[41] Police departments spend an average of $105,515 (2000$) per officer during our sample period. The -8% officer differential translates into (2.07/1,000)×214,647×0.08, or 35.55 officers. Thus, the implied difference in police department spending between the two government forms is $3,751,058. This represents about 7% of police department spending (mean = $56,328,770) and 0.5% of city spending (mean = $683,043,100).

While the lower number of officers in manager cities may be consistent with managers' weaker patronage motivations (Enikolopov 2010) the lack of a similar pattern for civilians seems to belie this explanation. 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. The officer and civilian estimates are, however, consistent with voter-induced differences in pandering incentives, which according to our model should be present in popular policy issues, such as police officers (Proposition 1) but absent in neutral policy issues, such as police civilians (Proposition 2).

The negative officer differential is also consistent with a selection model where bad manager types biased against officers are screened out less quickly relative to bad mayor types (Proposition 4). The two models, however, generate opposing comparative statics allowing us to assess in Tables 8 and 9 below which model better explains our finding.

Institutional, Political and Informational Controls. Manager charter reforms often came on the heels of other reforms in city institutions. The three major prior reforms were: non-partisan ballots, at-large elections, and civil service rules. The literature sometimes packages these reforms together with manager reform, drawing a distinction between "traditional" and "reformed" cities (Stucky 2005). Our summary statistics in Table 1 confirm some of these stylized facts.

In column (1) of Table 3 the officer employment differential from Table 2 remains unaffected when including these institutional confounds, indicating an independent effect of manager government. Column (3) also upholds our previous finding of no statistically significant civilian employment differential. The independent effects of the three related institutions sometimes go in the same direction and other times in the opposite direction of manager government, although they are not statistically significant.

If interest groups are able to influence the choice of city political institutions the results thus far may simply reflect interest group preferences. We consider police unions to be the relevant interest group with preferences potentially opposite to voters'. We control for interest group strength using two variables: the presence of a state law allowing collective bargaining for police employees, and a more local measure, the fraction of police unionized as of 1968.[42] Voter crime perceptions is another key theoretical variable that we capture

[42] City-level police unionization data is very limited. The *Municipal Year Book* reports the presence/absence of a police union in its 1961 and 1962 editions, and the fraction of unionized police employees in its 1968 edition. We choose the latter variable since it is more informative.

using sales of local newspapers. Local media can strengthen the voter's crime perceptions while being related to government form.

Columns (2) and (4) of Table 3 show that the baseline results above are robust to the inclusion of interest group and media controls. The manager coefficient remains negative and statistically significant at the 5% level for police officers. It is also interesting to note that strong police unions are associated with low officer employment, whereas local media is associated with high officer employment. These findings resonate with two of our key theoretical assumptions: bad policymakers are biased against the popular policy; and stronger crime perceptions increase the voter's preference for high police.

Policymaker Salary and Characteristics. The summary statistics in Table 1 show that managers earn higher salaries than mayors, at least in the decade of our sample for which salary data is available. In Proposition 3 we established that a higher salary generates more popular policies by increasing the proportion of biased policymakers and thus triggering a pandering equilibrium. Thus salary is a potential confound of the manager-mayor policy differential.[43] While gender and race have been shown to affect policy at the U.S. city level (e.g. Ferreira and Gyourko 2011, Nye, Rainer and Stratmann 2010) they display less correlation with government form in our sample.

Table 4 accounts for these factors using a reduced sample due to limited data availability. The smaller officer differential in column (1) is consistent with the confounding effect of salary predicted by Proposition 3 becoming more pronounced over time.[44] In columns (2) and (4) we control for policymaker salary and continue to find a significantly negative officer differential and an insignificant positive civilian differential. Including further controls for policymaker gender and race in columns (3) and (6) does little to alter these findings.[45]

Instrumental Variables. While the results thus far show evidence of a police employment differential, OLS estimates may still be biased by reverse causality. Despite the infrequency of actual change 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

[43] Ferraz and Finan (2011) and Gagliarducci and Nannicini (2012) find that higher salaries increase the productivity of Brazilian city councilmembers and Italian mayors by attracting more competent types.

[44] According to salary averages published in ICMA's *Municipal Year Book*, for cities over 250,000 residents the ratio of manager to mayor salaries has been stable around 1.17 in the 1960-1980 period, after which it started to grow reaching 1.44 in 1990 and 1.79 in 2000.

[45] Salary may be endogenous to policy outcomes. For example, Enikopolov (2011) finds that growing cities reward managers with a higher salary. In Table A7 of the Online Appendix we find little evidence that policymakers who reduce crime are rewarded with higher salaries.

potential source of bias in the OLS estimates. 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 (Coate and Knight 2011).

To address concerns with OLS bias 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 they facilitated the ascension of engineers into top executive office. Floods caused by extreme precipitation were one such crisis. The demand for engineering-trained city executives was stronger when flood control was a local responsibility, i.e. before the federal government establishes the Army Corps of Engineers in 1936.[46] We thus instrument for manager government using the frequency of precipitation shocks in the local flood control era, *LFC precipitation shocks*.[47]

Our identification strategy requires that (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 Online Appendix we provide supportive evidence for these identifying assumptions.[48]

Table 5 presents IV estimates of the employment differentials controlling for typical local climate conditions.[49] In columns (2) and (4) 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,

[46] Table A1 in the Online Appendix shows that the relationship between extreme precipitation and flood incidence is markedly stronger in the local flood control period (1900-1936).

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

[48] In support of instrument validity we find that (i) LFC precipitation shocks are not correlated with early city characteristics (see Table A2). In support of the exclusion restriction we find that (ii.a) in contrast to LFC precipitation shocks, federal flood control (FFC) precipitation shocks and LFC temperature shocks are not related to police employment today (see Table A3), and (ii.b) LFC precipitation shocks are not related to other present-day city institutions (see Table A4).

[49] Formally, the first stage model is:

$$\begin{aligned} Manager_{i,t} &= \alpha_1 LFC_Precipitation_Shocks_i + \alpha_2 Century_Precipitation_Shocks_i + \\ &+\alpha_3 Median_Precipitation_i + \alpha_4 X_{i,t} + \tau_t + \epsilon_{i,t} \end{aligned} \quad (3)$$

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 precipitation shocks during the 20th century, $Median_Precipitation_i$ is median annual precipitation during the 20th century, and the remaining variables are from equation (2). We include $Century_Precipitation_Shocks_i$ to make our exclusion restriction credible as previous research has found evidence that climate affects economic growth (Dell, Jones and Olken 2008), crime (Jacob, Lefgren and Moretti 2007), conflict (Miguel, Satyanath and Sergenti 2004), and the origins of trust (Durante 2010).

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 OLS estimates.[50]

Whether the identified effects have external validity is an interesting question. If the costs of subsequently changing government form are heterogeneous, the identified effects will be local to high-cost cities and potentially different from the population-wide treatment effect. For example, Acemoglu, Robinson, and Torvik (2011) argue that voters dismantle exogenously imposed checks and balances when politician rents are low and special interests are strong. Table A5 in the Online Appendix shows that the IV results are robust to alternative instrument definitions, strengthening their external validity.[51]

Robustness. Thus far we have scaled the number of police employees by population, however a more relevant measure may be the crime level if large cities have more crime. A high officer-crime ratio may also better capture deviations from policy optimality, by providing a measure of "excess police." The theory model implies that hierarchical accountability allows voters more flexibility to optimally incentivize a policymaker, resulting in less excess police. The estimates are presented in Table 6. In both OLS and IV specifications manager cities have fewer officers per crime than mayor cities, but not fewer civilians per crime, strengthening support for (H1).

Table 7 reports sensitivity checks for the parsimonious OLS and IV estimates, i.e. the model that does not include geography and Census division fixed effects. 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 the post-1994 period, when the federal government intervened more forcefully in local police employment through COPS grants, dropping cities with many 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 of Table 7 varies the inference procedure. We examine three sets of alternative standard errors: clustering on both city and year, clustering on weather station (as this

[50] Baqir (2002) and Whalley (2010) also find little evidence that institutional endogeneity contaminates OLS results in the context of U.S. local governments.

[51] An additional external validity question is whether the identified effects apply to policy issues other than police employment. Finding other local policy issues susceptible to pandering incentives is an interesting question for future research.

is the unit of observation for the weather variables), and accounting for spatial correlation (Conley, 1999).[52] 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.

Political and Informational Mechanisms. To peer into the theoretical mechanisms we add to equation (2) an interaction of $Manager_{i,t}$ with one theoretically-characterized factor at a time. Table 8 reports the interaction estimates. Mayors of larger cities employ significantly more police per capita, both officers and civilians. The officer differential grows twice as big in large cities (p-value 0.077), consistent with (H2), while the civilian differential shrinks to practically zero.

Locally-measured police unionization is associated with a reduction of the officer differential by 15 percentage points (p-value 0.058), consistent with (H2), while having little effect on the civilian differential. These estimates support the model's prediction that, in order to preempt interest group influence on the policymaker, the voter limits manager discretion by transmitting pandering incentives through the city council, thus reducing observable policy differences between managers and mayors. Columns (4) and (8) of Table 8 present estimates of employment differentials as a function of voter crime perceptions, measured by high local newspaper sales. The results line up with the theoretical prediction that strong voter perceptions make it optimal for the voter to transmit pandering incentives to the manager. Strong voter perceptions are associated with a reduction of the officer differential by 17 percentage points and no significant change in the civilian differential.

The police unionization and local newspaper effects run counter to the predictions of a pure selection model where the officer differential would be amplified by these factors. Thus Table 8 is more supportive of the incentives-based hypotheses in (H2). Columns (3) and (4) also provide some evidence that managers respond to local factors like interest groups and voter perceptions. This suggests that city managers do respond to local accountability rather than being completely neutral autonomous professionals driven by peer accountability from the national professional association.

Electoral Effects. Finally, the theory model predicts that pandering incentives are strongest

[52] 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 fixed 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.

right before an election because the voter relies more on recent information about policy choices. We measure election proximity by coding the year before the election date as an election year.[53] To estimate electoral effects we add to equation (2) 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 off-election years, while $\beta_{Manager} + \beta_{Manager \times Election}$ measures the manager-mayor policy differential in election years. Incentives-based hypothesis (H3) states that the off-election year differential is zero, and the election year differential is negative. The pure selection model makes the reverse prediction.

Column (1) of Table 9 Panel A shows that mayors employ 0.6% more officers in election years, suggesting pandering ahead of an election, although the estimate is statistically weak. Based on (H2), however, we expect pandering incentives to be stronger in larger cities with a stronger retention motivation. Columns (2) to (4) restrict the sample to cities with population over 50,000 in 1960. The coefficient pattern from column (1) reappears in columns (2) and (3) and is now also statistically significant at conventional levels.[54] The further addition of a policymaker gender control in column (4) has little effect on the results. The interaction coefficient is negative and highly statistically significant, lending support to an incentives model (Proposition 5) as opposed to a pure selection model (Proposition 6) that predicts a positive interaction coefficient. The point estimates for civilians in Table 9 Panel B are all close to zero and never statistically significant. Overall these results imply that mayor cities display an electoral cycle in officer employment that is not present in the average manager city.[55]

6 Conclusion

This paper sheds light on the incentive effects of hierarchical accountability, a prevalent and increasingly common accountability form in many governments such as counties, cities and school districts. Our theory uses a principal-monitor-agent model where informational asymmetries play a central role. The model shows that the voter, while not able to harness the full informational advantage of the monitor, uses the monitor to credibly commit to allow

[53] 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. For manager governments we use the election date for the mayor office, if there is one, otherwise for the council.

[54] Levitt (1997) exploits the correlation between police hiring and exogenously-set election dates to identify the effect of police on crime, using data from 59 large U.S. cities with directly-elected mayors. In a follow-up analysis McCrary (2002) uses the same sample but recodes election dates.

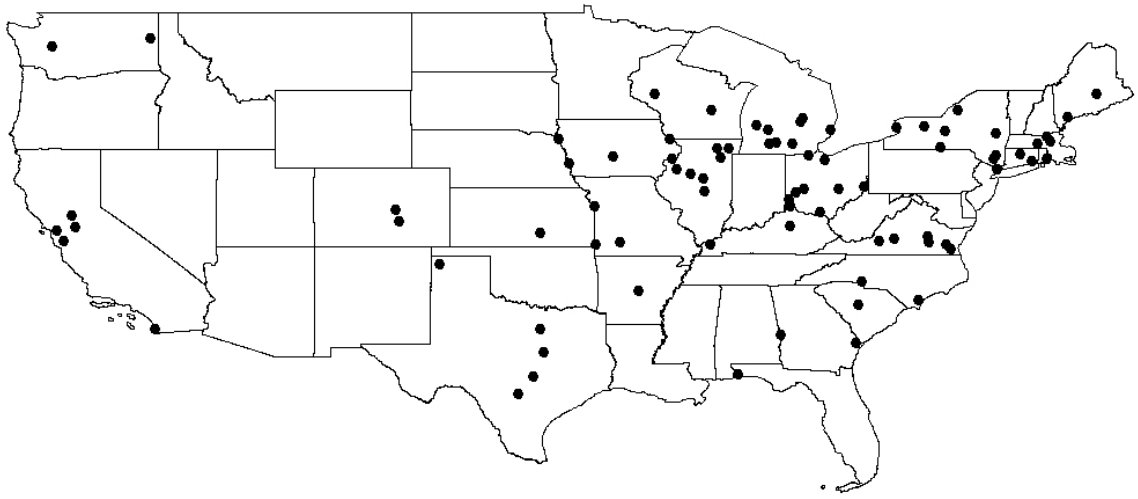
[55] In unreported analysis we have also estimated models with police hiring $\Delta \log(Police_{i,t})$ as the dependent variable. The results are stronger and are available upon request.

the policymaker discretion when interest groups and voter policy perceptions are weak.

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 theoretical mechanisms support our agency model of incentives and survive a number of robustness checks, including possible endogeneity of city government form.

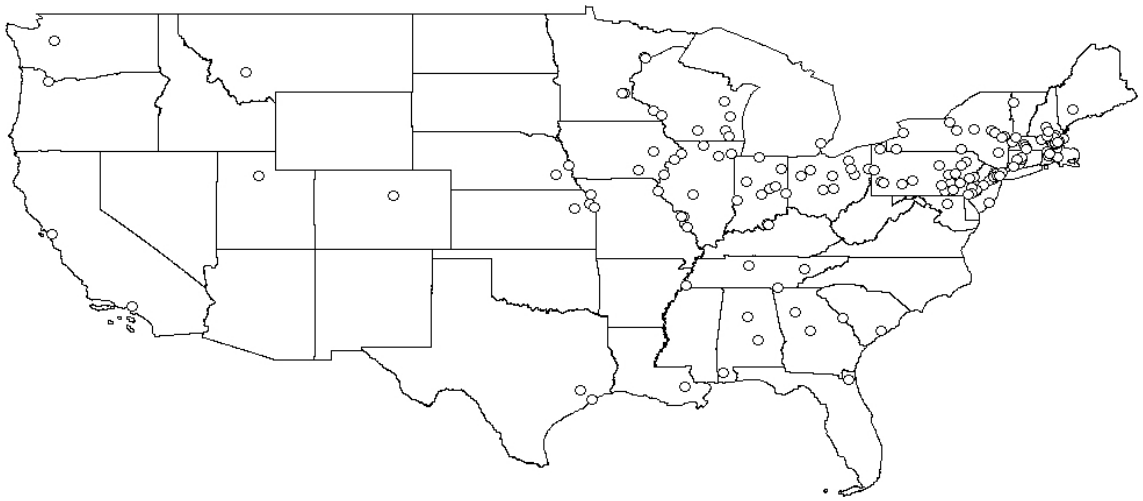
Our analysis raises a number of theoretical and empirical issues that can be addressed in further work. Our model emphasizes the importance of informational asymmetries, yet information is assumed to be exogenous to the institutional environment. One can imagine that voter and policymaker motivations in acquiring information depend on the accountability form. On the one hand, voters might be less interested in acquiring information if they cannot directly control the policymaker. On the other hand, indirectly-elected policymakers, who enjoy more discretion, have an incentive to acquire more policy information. In addition, if hierarchical agency opens the door to collusion, which contract theory models suggest may weaken incentives, our conclusions may overstate the case for this accountability form. It would also be instructive to know if hierarchical incentives work similarly in a model with competence types. Empirically, it would be interesting to test the model's predictions in alternative governmental environments, such as county executives or central bank governors, where policy popularity takes different forms. Another avenue would be to test for additional implications of the model, for instance how voter and monitor strategies adjust to the political and informational environment in order to optimally incentivize an indirectly-elected policymaker.

FIGURE 1: Manager Cities, 1960



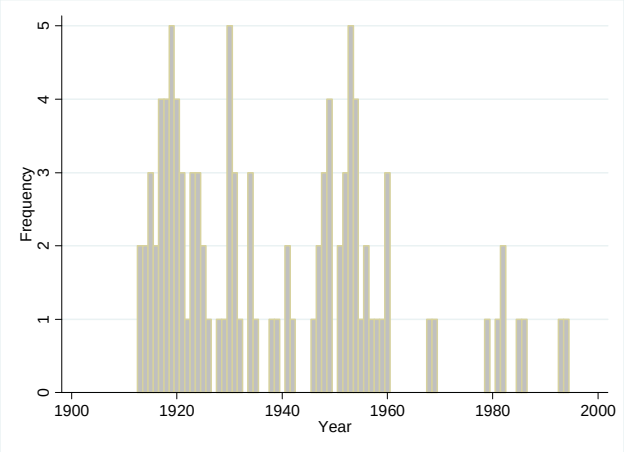
Notes: Authors' calculations for the sample of the 248 largest self-governing cities in 1900. Manager government locations are plotted based on 1960 form of government. The map reflects the jurisdictional boundaries of each state.

FIGURE 2: Mayor Cities, 1960



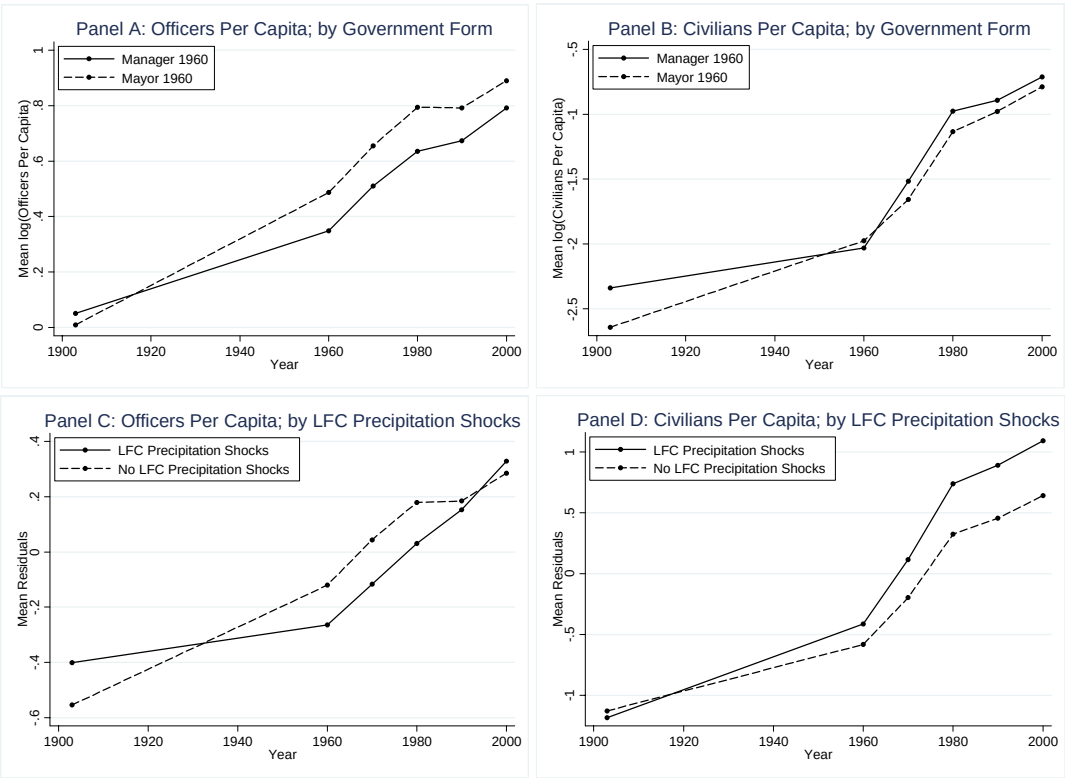
Notes: Authors' calculations for the sample of the 248 largest self-governing cities in 1900. Mayor government locations are plotted based on 1960 form of government. The map reflects the jurisdictional boundaries of each state.

FIGURE 3: Timing of Manager Charter Adoptions



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

FIGURE 4: Police Employment, by Indicated Cohort



Notes: Authors' calculations for the sample of the 248 largest self-governing cities in 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 1: Descriptive Statistics, By City Government Form

Sample:	Full	Manager Cities	Mayor Cities	Difference t-stat [p-value]
	(1)	(2)	(3)	(4)
Panel A: Police Employment (per 1,000), 1960-2000				
Officers Per Capita	2.07	1.85	2.20	-4.41
	(0.97)	(0.55)	(1.12)	[0.000]
Civilians Per Capita	0.37	0.38	0.37	0.31
	(0.32)	(0.24)	(0.36)	[0.760]
Panel B: City Geographic, Demographic, Economic, and Crime Characteristics, 1960-2000				
Distance to Nearest River	32	40	28	2.01
	(42)	(55)	(31)	[0.046]
Population	214,647	146,198	252,521	-1.79
	(584,486)	(184,417)	(712,626)	[0.074]
Fraction Non-White	0.20	0.19	0.20	-0.13
	(0.18)	(0.15)	(0.20)	[0.900]
Fraction College Graduate	0.13	0.15	0.13	2.34
	(0.08)	(0.08)	(0.08)	[0.020]
Household Income	28,398	28,653	28,241	0.70
	(7,493)	(7,477)	(7,498)	[0.487]
Violent Crime Rate	3.76	3.35	4.00	-1.71
(1975-2000)	(3.70)	(2.82)	(4.09)	[0.089]
Property Crime Rate	77.57	85.36	73.21	3.21
(1975-2000)	(38.06)	(35.94)	(39.03)	[0.001]
Panel C: City Government Characteristics				
Non-Partisan Elections 1960	0.65	0.85	0.53	6.27
	(0.48)	(0.35)	(0.50)	[0.000]
Fraction At-Large Seats 1960	0.65	0.83	0.55	5.86
	(0.42)	(0.34)	(0.43)	[0.000]
Early Civil Service	0.70	0.69	0.70	-0.14
	(0.46)	(0.46)	(0.46)	[0.886]
Collective Bargaining Law	0.57	0.47	0.62	-3.37
	(0.50)	(0.50)	(0.49)	[0.001]
High Percent Police in Union in	0.40	0.48	0.36	1.85
1968	(0.49)	(0.50)	(0.48)	[0.066]
High Local Newspaper Sales	0.50	0.55	0.47	1.22
over 1990-2000	(0.50)	(0.50)	(0.50)	[0.224]
City Government Spending	2.17	2.11	2.21	-0.60
(1972-2000)	(2.65)	(2.58)	(2.68)	[0.548]
Panel D: Policymaker Characteristics				
Annual Salary (survey reported)	81,210	100,096	66,687	9.14
(1992-1993,1995-2000)	(32,125)	(27,760)	(27,360)	[0.000]

Male (survey reported)	0.85	0.88	0.84	1.12
(1992-1993,1995-2000)	(0.35)	(0.33)	(0.37)	[0.262]
White (survey reported)	0.78	0.80	0.77	0.63
(1992-1993,1995-2000)	(0.41)	(0.40)	(0.42)	[0.526]
Male (name-based imputation)	0.93	0.92	0.93	-0.76
(1960-2000)	(0.26)	(0.27)	(0.25)	[0.450]
Panel E: Police Department Characteristics (per 1,000), as of 1900				
Officers Per Capita	1.10	1.10	1.10	-0.08
	(0.42)	(0.34)	(0.46)	[0.935]
Civilians Per Capita	0.07	0.08	0.06	1.79
	(0.09)	(0.11)	(0.06)	[0.074]
Arrests Per Capita	0.06	0.07	0.05	2.30
	(0.05)	(0.05)	(0.04)	[0.022]
Number of Observations	10,168	3,622	6,546	

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900. Each entry in columns (1) - (3) presents the mean for the indicated variable, with the standard deviation in parentheses. The entries in column (4) are test statistics for the hypothesis that the means of the variables in columns (2) and (3) are the same, with standard errors clustered at the city level. The entries in square brackets in column (4) are the p-values for the hypothesis of equal means. The number of observations reflects the maximum sample size across all the reported variables.

TABLE 2: City 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)
Manager	-0.16***	-0.08**	-0.06**	-0.06**	0.09	0.02	0.05	0.05
	(0.03)	(0.03)	(0.02)	(0.02)	(0.07)	(0.07)	(0.06)	(0.06)
Additional Controls:								
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography & Division	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Demographic	No	No	Yes	Yes	No	No	Yes	Yes
Policymaker Gender	No	No	No	Yes	No	No	No	Yes
Sample	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850
Number of Clusters	248	248	248	248	248	248	248	248

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. The main entries in the columns (1)-(8) report an OLS estimate of β_1 from equation (2) in the text. Standard errors clustered at the city level reported in parentheses. The models in columns (1) and (5) do not include any additional controls. The models in columns (2) and (6) also include 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 the coast*, *percentage of clay in the soil* and Census division fixed effects. The models in columns (3) and (7) also include the following demographic controls: *population, fraction of population non-white, fraction of population college graduate*, and *median household income.* The models in columns (4) and (8) also control for *policymaker male* (name-based imputation) and an indicator variable for *missing policymaker male* (name-based imputation). * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 3: City Government Form and Police Employment: OLS Estimates Controlling for Institutional and Political Factors

Dependent Variable =	Log (Officers Per Capita)		Log (Civilians Per Capita)	
	(1)	(2)	(3)	(4)
Manager	-0.08**	-0.07**	-0.03	-0.04
	(0.03)	(0.03)	(0.07)	(0.08)
Non-Partisan Elections 1960	0.07	0.07	0.00	-0.02
	(0.04)	(0.04)	(0.09)	(0.09)
Fraction At-Large Seats 1960	-0.06	-0.05	0.13	0.16*
	(0.04)	(0.04)	(0.09)	(0.09)
Early Civil Service	0.04	0.04	-0.01	-0.02
	(0.04)	(0.04)	(0.08)	(0.08)
Collective Bargaining Law		-0.08***		-0.08
		(0.03)		(0.06)
High Police Unionization 1968		-0.03		-0.12
		(0.04)		(0.08)
High Local Newspaper Sales 1990-2000		0.07**		0.03
		(0.03)		(0.07)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Geography & Division	Yes	Yes	Yes	Yes
Sample	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,974	9,974	9,850	9,850
Number of Clusters	248	248	248	248

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. The main entries in the columns (1)-(4) report an OLS estimate of β_1 from equation (2) in the text. Standard error clustered at the city level reported in parentheses. All models in the table include 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 the coast*, *percentage of clay in the soil* and Census division fixed effects. The models in columns (1) and (3) also include a missing value indicator for *Non-Partisan Elections in 1960*, *Fraction At-Large Seats 1960* and *Early Civil Service*. The models in columns (2) and (4) also include a missing value indicator for *High Police Unionization 1968*. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 4: City Government Form and Police Employment: OLS Estimates Controlling for Policymaker Salary and Characteristics

Dependent Variable =	Log (Officers Per Capita)			Log (Civilians Per Capita)		
	(1)	(2)	(3)	(4)	(5)	(6)
Manager	-0.03	-0.07**	-0.07***	0.08	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.08)	(0.09)	(0.09)
Log(Policymaker Salary)		0.09***	0.09***		0.11	0.11
		(0.02)	(0.02)		(0.08)	(0.08)
Policymaker Male			0.07*			0.00
			(0.04)			(0.09)
Policymaker White			-0.03			0.02
			(0.03)			(0.07)
Additional Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geography & Division	Yes	Yes	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes	Yes	Yes
Sample	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000
Number of Observations	1,164	1,164	1,164	1,156	1,156	1,156
Number of Clusters	236	236	236	235	235	235

Notes: Authors' calculations with city data from 1992-1993, 1995-2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900 with non-missing policymaker characteristics data. Each column reports the results from one regression. The main entries in the columns (1)-(6) report an OLS estimate of β_1 from equation (2) in the text. Standard errors clustered at the city level reported in parentheses. All models in the table include 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 the coast*, *percentage of clay in the soil* and Census division fixed effects, as well as demographic controls: *population, fraction of population non-white, fraction of population college graduate*, and *median household income*. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 5: City Government Form and Police Employment: IV Estimates

Dependent Variable =	Log (Officers Per Capita)		Log (Civilians Per Capita)	
	(1)	(2)	(3)	(4)
	Panel A: Second Stage *Dependent Variable = Log (Police Employment Per Capita)*			
Manager	-0.28**	-0.50**	0.69**	0.19
	(0.14)	(0.21)	(0.35)	(0.43)
Century Precipitation Shocks	-0.27	-1.78	7.11***	2.23
	(0.87)	(1.54)	(1.82)	(2.36)
Median Precipitation 1900-2000	0.55**	0.38	-1.41**	-1.41
	(0.22)	(0.61)	(0.60)	(1.12)
	Panel B : First Stage *Dependent Variable = Manager*			
Excluded Instrument:				
LFC Precipitation Shocks	11.67***	9.79***	11.69***	9.85***
	(2.85)	(2.54)	(2.85)	(2.54)
Controls:				
Century Precipitation Shocks	-6.90***	-9.86***	-6.95***	-9.87***
	(2.07)	(2.29)	(2.07)	(2.30)
Median Precipitation 1900-2000	-0.54	1.41*	-0.53	1.42*
	(0.45)	(0.81)	(0.45)	(0.81)
Excluded Instrument F-Statistic	16.76	14.88	16.83	14.99
[p-value]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Geography & Division	No	Yes	No	Yes
Sample	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,974	9,974	9,850	9,850
Number of Clusters	248	248	248	248

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. The main entries in the first row of columns (1)-(4) report an IV estimate of β_1 from equation (2) in the text with the first stage given by equation (3) and estimated in Panel B. Standard errors clustered at the city level reported in parentheses. The excluded instrument is *LFC Precipitation Shocks*. The models in columns (1) and (3) do not include any additional controls. The models in columns (2) and (4) also include 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 the coast*, *percentage of clay in the soil* and Census division fixed effects. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 6: City Government Form and Police-Crime Ratios

Dependent Variable =	Log(Officers Per Crime)		Log(Civilians Per Crime)	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
	Panel A: Violent Crime			
Manager	-0.14**	-0.60	-0.01	0.12
	(0.07)	(0.40)	(0.08)	(0.45)
Sample	1975-2000	1975-2000	1975-2000	1975-2000
Number of Observations	6,126	6,126	6,101	6,101
Number of Clusters	248	248	248	248
	Panel B: Property Crime			
Manager	-0.20***	-0.46**	-0.07	0.26
	(0.03)	(0.23)	(0.06)	(0.31)
Sample	1975-2000	1975-2000	1975-2000	1975-2000
Number of Observations	6,143	6,143	6,117	6,117
Number of Clusters	248	248	248	248
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Geography & Division	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes

Notes: Authors' calculations with city data from 1975 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. The main entries in the columns (1)-(4) report an estimate of β_1 from equation (2) in the text. Standard errors clustered at the city level reported in parentheses. The models include 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 the coast*, *percentage of clay in the soil* and Census division fixed effects, as well as demographic controls: *population, fraction of population non-white, fraction of population college graduate*, and *median household income.* * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 7: City Government Form and Police Employment: Alternative Samples and Inferences

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 Large Cities	-0.13***	-0.28**	0.14*	0.68*
	(0.04)	(0.14)	(0.07)	(0.35)
(2) Drop Dependent Variable Outliers	-0.08***	-0.21*	0.18**	0.79**
	(0.03)	(0.11)	(0.07)	(0.34)
(3) Drop Low Government Concentration Census Divisions	-0.15***	-0.19*	0.12	0.80**
	(0.04)	(0.11)	(0.08)	(0.31)
(4) Drop post-COPS Program Years	-0.17***	-0.35**	0.10	0.70**
	(0.04)	(0.15)	(0.07)	(0.35)
(5) Drop High Missing Precipitation Cities	-0.18***	-0.26*	0.05	0.95**
	(0.04)	(0.15)	(0.07)	(0.43)
(6) Drop Far from Weather Station Cities	-0.17***	-0.28*	0.13*	0.79**
	(0.04)	(0.16)	(0.08)	(0.40)
Panel B: Alternative Standard Error Construction and Inferences				
(7) Cluster on both City and Year	-0.16***	-0.28**	0.09	0.69**
	(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
	(0.04)	(0.14)	(0.07)	(0.41)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Geography & Division	No	No	No	No

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900. Each entry reports the results from one regression. The main entries in the columns (1) and (3) report an OLS estimate of β_1 from equations (2) and (3) in the text. Standard errors clustered at the city level reported in parentheses (unless otherwise indicated). Panel A contains results for alternative subsamples as indicated in the text. Panel B contains results for alternative inference procedures as indicated in the text. The specifications reported in the Conley Standard Errors row do not include year fixed effects. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 8: City Government Form and Police Employment: Mechanisms

Dependent Variable =	Log (Officers Per Capita)				Log (Civilians Per Capita)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager	-0.10**	-0.18***	-0.25***	-0.25***	0.25**	-0.08	0.09	0.08
	(0.04)	(0.05)	(0.06)	(0.06)	(0.10)	(0.09)	(0.12)	(0.11)
Manager × High Population	-0.12*				-0.27**			
	(0.07)				(0.13)			
Manager × Collective Bargaining Law		0.06				0.24**		
		(0.05)				(0.11)		
Manager × High Police Unionization 1968			0.15*				-0.07	
			(0.08)				(0.15)	
Manager × High Local Newspaper Sales 1990-2000				0.17**				0.02
				(0.07)				(0.14)
High Population	0.12***				0.56***			
	(0.05)				(0.09)			
Collective Bargaining Law		0.04				-0.40***		
		(0.04)				(0.08)		
High Police Unionization 1968			-0.11*				0.04	
			(0.06)				(0.11)	
High Local Newspaper Sales 1990-2000				-0.09*				0.03
				(0.05)				(0.10)
Additional Controls:								
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography & Division	No	No	No	No	No	No	No	No
Sample	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850
Number of Clusters	248	248	248	248	248	248	248	248

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. The main entries in the columns (1)-(8) report OLS estimates of equation (2) in the text augmented with an additional variable and an interaction term. Standard errors clustered at the city level reported in parentheses. All models except columns (3) and (7) only include year fixed effects as additional controls. The models in columns (3) and (7) include year fixed effects and an indicator variable for missing *High Police Unionization 1968* and its interaction with *Manager*. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE 9: City Government Form and Electoral Cycles in Police Employment: OLS Estimates

Sample Cities =	Full	Large Cities		
	(1)	(2)	(3)	(4)
	Panel A: Dependent Variable = Log (Officers Per Capita)			
Election	0.006	0.010**	0.009**	0.009**
	(0.004)	(0.004)	(0.004)	(0.004)
Manager	-0.078	-0.064	-0.044	-0.045
	(0.056)	(0.084)	(0.047)	(0.047)
Manager × Election	-0.009*	-0.015**	-0.013***	-0.013***
	(0.005)	(0.007)	(0.005)	(0.005)
Sample	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,974	7,033	7,033	7,033
Number of Clusters	248	174	174	174
	Panel B: Dependent Variable = Log (Civilians Per Capita)			
Election	0.008	0.006	0.005	0.005
	(0.008)	(0.009)	(0.008)	(0.008)
Manager	0.009	-0.040	-0.034	-0.035
	(0.095)	(0.110)	(0.085)	(0.084)
Manager × Election	-0.001	-0.009	-0.006	-0.006
	(0.013)	(0.015)	(0.014)	(0.014)
Additional Controls:				
City Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographic	No	No	Yes	Yes
Policymaker Gender	No	No	No	Yes
Sample	1960-2000	1960-2000	1960-2000	1960-2000
Number of Observations	9,850	7,017	7,017	7,017
Number of Clusters	248	174	174	174

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year. The sample is the largest 248 self-governing cities in the United States in 1900 in column (1) and 174 cities that also have at least 50,000 residents in 1960 in columns (2)-(4). Each column reports the results from one regression. The main entries in columns (1)-(4) report OLS estimates of equation (2) in the text augmented with *Election* and *Manager×Election*. Standard errors clustered at the city level reported in parentheses. The models in columns (1) and (2) only include year and city fixed effects as additional controls. The model in column (3) also includes demographic controls: *population, fraction of population non-white, fraction of population college graduate*, and *median household income*. The model in column (4) also includes *policymaker male* (name-based imputation) and an indicator variable for *missing policymaker male* (name based imputation). * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

Theory Appendix

Lemma 1 *In hierarchical accountability equilibrium: (i) Strategic policymakers are always responsive to a monitor prescription; (ii) Strategic monitors are always responsive to a voter prescription; (iii) Retiring policymakers act on their preferences; (iv) Retiring monitors prescribe their own preferred policy in every contingency.*

Proof. Let P^θ denote a type θ policymaker's, and $M^\theta(\pi)$ a type θ monitor's, stationary equilibrium continuation value. (i) Suppose to the contrary that the policymaker is in a contingency where his preference conflicts with the monitor's prescription and in equilibrium he prefers to follow his preference: $P^\theta < 1$. The equilibrium continuation value is then 1 in such contingencies, but is $1 + P^\theta$ in the other contingencies, and so on average it is at least 1. Contradiction. (ii) In a sequentially rational strategy, once policy has been chosen, the monitor compares the value of compliance $M^\theta(\pi)$ to the value of non-compliance, which is zero. Since $M^\theta(\pi) \geq \left(1-\sigma^P\right)(1-bp) > 0$, the monitor is better off complying. (iii) Retiring policymakers are not affected by the firing rule, so they are truthful. (iv) For a retiring monitor any firing rule is sequentially rational. Thus he will choose the one that maximizes his equilibrium payoff, which, by Part (i), is prescribing his preferred policy. ■

Lemma 2 *In a contingency where a type θ monitor has a choice between two sequentially rational strategies that: (i) induce the same retention probability, he chooses the one with higher average payoff, (ii) induce different retention probabilities, he chooses the one with higher retention probability, as $\sigma^P \to 1$.*

Proof of Lemma 2. (i) A monitor strategy with average payoff $m^\theta = \sum_{(s,l)} m^\theta(s,l)\mathbb{P}\{(s,l)\}$ and retention probability $\rho^\theta = \sum_{(s,l)} \rho^\theta(s,l)\mathbb{P}\{(s,l)\}$ yields an equilibrium continuation value $M^\theta(\pi) = \pi\frac{\left(1-\sigma^P\right)\bar{m}^\theta+\sigma^P-\sigma^P\sigma^M\left(1-m^{\theta|G}\right)}{1-\sigma^M+\sigma^P\sigma^M(1-\rho^{\theta|G})} + (1-\pi)\frac{\left(1-\sigma^P\right)\bar{m}^\theta+\sigma^P-\sigma^P\sigma^M\left(1-m^{\theta|B}\right)}{1-\sigma^M+\sigma^P\sigma^M(1-\rho^{\theta|B})}$, where $\bar{m}^G = \pi + (1-\pi)(1-bp)$ and $\bar{m}^B = \pi(1-bp) + (1-\pi)$ are a good, respectively bad, monitor's expected payoff from a retiring policymaker. Note that $M^\theta(\pi)$ is increasing in both payoffs and retention probabilities. (i) If the monitor has a choice between two strategies with the same retention probabilities, then he chooses the one that maximizes average payoffs. (ii) Suppose the monitor has a choice between two strategies with different retention probabilities. In a given contingency with frequency $\mathbb{P}\{(s,l)\}$ one strategy insures retention but no current payoff, whereas the other a current payoff but no retention. Forgoing a current payoff is profitable if: $\pi\frac{\partial}{\partial\mathbb{P}\{(s,l)\}}\frac{\left(1-\sigma^P\right)\bar{m}^\theta+\sigma^P-\sigma^P\sigma^M\left[1-m^{\theta|G}+\mathbb{P}\{(s,l)\}\right]}{1-\sigma^M+\sigma^P\sigma^M\left[(1-\rho^{\theta|G}-\mathbb{P}\{(s,l)\}\right]} +$

$(1-\pi)\frac{\partial}{\partial \mathbb{P}\{(s,l)\}}\frac{\left(1-\sigma^P\right)\bar{m}^\theta+\sigma^P-\sigma^P\sigma^M\left[1-m^{\theta|B}+\mathbb{P}\{(s,l)\}\right]}{1-\sigma^M+\sigma^P\sigma^M\left[(1-\rho^{\theta|B}-\mathbb{P}\{(s,l)\}\right]} > 0$ which as $\sigma^P \to 1$ is satisfied when $m^{\theta|G}+\rho^{\theta|G}>1$ and $m^{\theta|B}+\rho^{\theta|B}>1$. Any monitor strategy has these properties.[1] ■

Proof of Propostion 1. Direct Incentives. The voter has four possible pure Markov strategies: two prescriptive and two unconditional. Let $\bar{v}(\pi|b,p) = \pi + (1-\pi)(1-bp)$ denote the voter's expected period payoff from a truthful policymaker of reputation π.

Consider the *prescriptive* firing rule "keep iff popular policy." In the (high state, high appeal) contingency both incumbent types' preferences are in line with the firing rule, so they comply with it. In contingencies where incumbent preferences go against the firing rule the stationary payoffs from complying $P^G = 1+\frac{\sigma^P p}{1-\sigma^P}$ and $P^B = 1+\frac{\sigma^P p(1-b)}{1-\sigma^P}$ are strictly greater than the payoff 1 from following own preferences. Thus in equilibrium both policymaker types always comply with the firing rule. The voter's equilibrium continuation value is $V(\pi_{t+1}) = \frac{\left(1-\sigma^P\right)\bar{v}(\pi_{t+1}|b,p)+\sigma^P p}{1-\beta}$ which is strictly increasing in the next policymaker's reputation π_{t+1}. Thus the voter's equilibrium strategy is to retain a policymaker whose posterior reputation is at least as good as the challenger's prior reputation. Since the policymakers' strategies are pooling in all three contingencies, time t on-equilibrium-path voter beliefs are $\tilde{\pi}_t = \pi_t$.[2] One can analogously show that the prescriptive firing rule "keep iff unpopular policy" is sequentially rational and makes policymakers pool on the unpopular policy. Equilibrium voter welfare in this case is $V(\pi) = \frac{\left(1-\sigma^P\right)\bar{v}(\pi|b,p)+\sigma^P(1-p)}{1-\beta}$. Since $p > \frac{1}{2}$ voter welfare is strictly higher when prescribing the popular policy.

When faced with an *unconditional* firing rule the policymaker follows his preferences. That means that in equilibrium a popular policy improves the policymaker's reputation $\tilde{\pi}_t(x_t=1) = \frac{\pi_t p}{\pi_t p+(1-\pi_t)p(1-b)} > \pi_t$ and an unpopular policy hurts the policymaker's reputation $\tilde{\pi}_t(x_t=0) < \pi_t$. The voter's equilibrium continuation value is strictly increasing in the next policymaker's reputation $V(\pi_{t+1}) = \frac{\bar{v}(\pi_{t+1}|b,p)}{1-\beta}$ making the voter want to keep after $x_t=1$ and fire after $x_t=0$. Thus an unconditional rule cannot be sequentially rational.

Hierarchical Incentives. The voter has sixteen possible pure Markov strategies: four prescriptive, four unconditional and eight half-prescriptive.

The *prescriptive* voter strategy "retain monitor iff keeps popular policymaker and fires unpopular policymaker" forces both monitor types to play the prescriptive strategy "keep policymaker iff popular policy," by Lemma 1(ii), and induces both policymaker types to

[1] If $\sigma^P \to 0$ monitor retention depends less and less on monitor strategies, which makes maximizing average payoff preferable to maximizing average retention probability.

[2] And therefore equilibrium policymaker reputation is stationary $\pi_t = \pi$ for all t. Off-equilibrium-path the following beliefs support an equilibrium: after $x_t = 1$ and some prior unpopular policy choices, beliefs are $\tilde{\pi}_t > \pi_t$ (= iff $\pi_t = 1$); after $x_t = 0$, beliefs are $\tilde{\pi}_t < \pi_t$ (= iff $\pi_t = 0$).

choose the popular policy, by Lemma 1(i). On-equilibrium-path voter beliefs are $\tilde{\pi}_t = \pi_t$ and $\tilde{\mu}_t = \mu_t$. The voter's equilibrium continuation value is strictly increasing in next period's policymaker and monitor reputations $V(\mu_{t+1}, \pi_{t+1}) = \frac{\left(1-\sigma^P\right)\bar{v}(\pi_{t+1}|b,p)+\sigma^P\left\{\left(1-\sigma^M\right)\bar{v}(\mu_{t+1}|b,p)+\sigma^M p\right\}}{1-\beta}$.[3] The analysis of the voter strategy prescribing the monitor to "keep unpopular and fire popular" is analogous and it induces both policymaker types to choose the unpopular policy.

The *prescriptive* voter strategy "retain monitor iff keeps policymaker" forces both monitor types to play the prescriptive strategy "keep policymaker no matter what," by Lemma 1(ii), and induces policymakers to be truthful. On-equilibrium-path voter beliefs are $\tilde{\pi}_t \neq \pi_t$ and $\tilde{\mu}_t = \mu_t$. The voter's equilibrium continuation value is strictly increasing in next period's policymaker and monitor reputations $V(\mu_{t+1}, \pi_{t+1}) = \frac{\left[\left(1-\sigma^P\right)+\sigma^P\sigma^M\right]\bar{v}(\pi_{t+1}|b,p)+\sigma^P\left(1-\sigma^M\right)\bar{v}(\mu_{t+1}|b,p)}{1-\beta}$.[4] This equilibrium is stationary iff $p < 1 - bp$ because otherwise if the policymaker's reputation drops too much the voter may prefer to switch to a prescriptive strategy that induces pandering. The analysis of the voter strategy prescribing the monitor to "fire policymaker no matter what" is analogous and induces both policymaker types to be truthful. This equilibrium is always stationary since $\tilde{\pi}_t = \pi_t = \pi$ for all t.

The *unconditional* strategies "retain monitor no matter what" and "replace monitor no matter what" are not sequentially rational for the voter. Each monitor type would prefer the prescriptive strategy that induces the policymaker to adopt the monitor's preferred policy, if that strategy were sequentially rational. If the voter replaces no matter what, this prescriptive strategy is clearly sequentially rational. If the voter retains no matter what, the policymaker is under unconstrained monitor control. By Lemma 1(i) a strategic policymaker is responsive to a monitor prescription. Thus monitor on-equilibrium-path beliefs about policymaker type are $\hat{\pi}_t = \pi_t$. The monitors' equilibrium continuation values are $M^G(\pi_{t+1}) = \frac{(1-\sigma^P)\bar{v}(\pi_{t+1}|b,p)+\sigma^P}{1-\sigma^M}$ and $M^B(\pi_{t+1}) = \frac{(1-\sigma^P)[\pi_{t+1}(1-bp)+1-\pi_{t+1}]+\sigma^P}{1-\sigma^M}$, the first strictly increasing and the second strictly decreasing in π_{t+1}. Since in this equilibrium policymakers pool according to the monitor's prescription, there are off-equilibrium-path beliefs that support it. On-equilibrium-path voter beliefs about monitor type are $\tilde{\mu}_t(x_t = 0, y_t = keep) = \frac{\mu_t(1-p)}{\mu_t(1-p)+(1-\mu_t)(1-p+bp)} < \mu_t$ and $\tilde{\mu}_t(x_t = 1, y_t = keep) = \frac{\mu_t p}{\mu_t p+(1-\mu_t)p(1-b)} > \mu_t$ and since the voter's equilibrium continuation value $V(\mu_{t+1}, \pi_{t+1}) = \frac{\left(1-\sigma^P\right)\bar{v}(\pi_{t+1}|b,p)+\sigma^P\bar{v}(\mu_{t+1}|b,p)}{1-\beta}$ is strictly increasing in μ_{t+1} the unconditional retention rule "retain monitor no matter what" is not sequentially rational because after $(x_t = 0, y_t = keep)$ the voter strictly prefers to replace;

[3] The off-equilibrium-path voter beliefs that support this equilibrium are: After popular-keep and unpopular-fire, but some prior deviations $\tilde{\mu}_t > \mu_t$ (= iff $\mu_t = 1$); After popular-fire and unpopular-keep $\tilde{\mu}_t < \mu_t$ (= iff $\mu_t = 0$).

[4] The off-equilibrium-path voter beliefs that support this equilibrium are: After keeping the policymaker, but some prior deviations $\tilde{\mu}_t > \mu_t$ (= iff $\mu_t = 1$); After firing the policymaker $\tilde{\mu}_t < \mu_t$ (= iff $\mu_t = 0$).

similarly, the unconditional retention rule "replace monitor no matter what" is not sequentially rational because after $(x_t = 1, y_t = keep)$ the voter strictly prefers to retain.

The *unconditional* strategy "retain monitor iff popular policy" makes all four monitor strategies are sequentially rational, as the voter does not prescribe a monitor response to either policy. The monitor prescription "keep iff popular" induces a strategic policymaker to choose the popular policy, by Lemma 1(i). This guarantees monitor retention. Thus, it will be used by both monitor types in the (high state, high appeal) contingency. By Lemma 2(ii), this prescription will be used in the other contingencies as well since the other monitor strategies reduce monitor retention probability for at least one policymaker type. The argument for the voter strategy "retain monitor iff unpopular policy" is analogous. It induces both monitor types to "keep iff unpopular policy" and induces both policymaker types to choose the unpopular policy.

The *positive half-prescriptive* voter strategies are not sequentially rational. Here we show the proof for the strategy "retain monitor unless popular policy and policymaker fired." The argument for the other three strategies is analogous. By Lemma 1(ii) this voter strategy forces a strategic monitor to keep a popular policymaker. However, monitors have freedom to keep/fire unpopular policymakers. Thus only two strategies may be sequentially rational for the monitor: "keep iff popular policy" and "keep no matter what." By Lemma 1(ii) in equilibrium the monitor is always retained. By Lemma 2(i) the monitor chooses the strategy with the highest average payoff. In contingency $(s_t, l_t) = (0, 0)$ both monitors prefer the unconditional strategy so that in equilibrium the policymaker chooses the unpopular policy and is kept. In contingency $(s_t, l_t) = (1, 0)$ a good monitor prefers the prescriptive strategy, while the bad monitor prefers the unconditional strategy. In contingency $(s_t, l_t) = (1, 1)$ both monitor types are indifferent between the prescriptive and unconditional strategies. It can be verified that these strategies are sequentially rational for the monitor.[5] On-equilibrium-path voter beliefs are $\tilde{\mu}_t(x_t = 0, y_t = keep) = \frac{\mu_t(1-p)}{\mu_t(1-p)+(1-\mu_t)[1-p+bp(1-\pi_t)]} < \mu_t$. The voter's equilibrium continuation value is $V(\mu_{t+1}, \pi_{t+1}) = \frac{\left(1-\sigma^P\right)\bar{v}(\pi_{t+1}|b,p)+\sigma^P\left\{\left(1-\sigma^M\right)\bar{v}(\mu_{t+1}|b,p)+\sigma^M\left[\mu_{t+1}+\left(1-\mu_{t+1}\right)\bar{v}(\pi_{t+1}|b,p)\right]\right\}}{1-\beta}$, which is strictly increasing in μ_{t+1}. Thus the voter wants to replace after $(x_t = 0, y_t = keep)$.

The *negative half-prescriptive* voter strategies induce policymaker behavior equivalent to the prescriptive strategies. Here we show the proof for the strategy "replace monitor

[5] Monitor strategies that induce policymaker pooling are sequentially rational. The bad monitor's strategy in the disagreement contingency induces policymaker separation. If the policymaker turns out to be good, he chooses the popular policy and the bad monitor needs to keep him to survive. If the policymaker turns out to be bad, he chooses the unpopular policy and the bad monitor wants to keep him.

unless popular policy and policymaker kept." The argument for the other three strategies is analogous. By Lemma 1(ii) a strategic monitor may have only two sequentially rational strategies: "keep iff popular policy" and "keep no matter what." Note that if the monitor uses the unconditional strategy the policymaker (being kept no matter what) can strategically influence monitor retention. The prescriptive strategy guarantees monitor retention in all contingencies; it will be used by both monitor types in the (high state, high appeal) contingency. By Lemma 2(ii) it is preferred by both monitor types in the other contingencies as well, since it improves retention probability relative to the unconditional strategy.

Voter Feedback. Suppose that with probability φ, where $0 \leq \varphi \leq 1$, the voter learns the state s_t before the period-t election. Suppose that if informed the voter plays "retain iff optimal." In a pandering equilibrium the policymaker's incentive to adopt popular policies is then counteracted in the low state and reinforced in the high state. The policymaker will remain interested in pandering if the informed incentive is dominated by the uninformed incentive: $(1-2\varphi)P^G, (1-2\varphi)P^B \geq 1$ where $P^G = \frac{\left(1-\sigma^P\right)+\sigma^P p}{1-\sigma^P(\varphi p+1-\varphi)}$ and $P^B = \frac{\left(1-\sigma^P\right)+\sigma^P p(1-b)}{1-\sigma^P(\varphi p+1-\varphi)}$. Since $P^G > P^B$ both policymaker types pander when $\varphi \leq \frac{\sigma^P p(1-b)}{2[(1-\sigma^P)+\sigma^P p(1-b)]+\sigma^P(1-p)}$. In a truthful equilibrium the policymaker is unconstrained by an uninformed voter, whereas an informed voter pushes him to adopt optimal policies. A good policymaker's incentives are consistent with his preferences. The same is true for a bad policymaker except when $(s_t, l_t) = (1, 0)$; in that contingency truthful behavior - choosing the unpopular policy - requires that the informed incentive be weak: $\varphi P^B \leq 1$, where $P^B = \frac{1}{1-\sigma^P(1-\varphi bp)}$, which implies $\varphi \leq \frac{1-\sigma^P}{1-\sigma^P bp}$. ■

Proof of Proposition 2. The proof follows the steps in Lemma 1, Lemma 2 and Proposition 1 while setting $b = \mathbb{P}\{l_t = 1|s_t = 0\} = \mathbb{P}\{l_t = 0|s_t = 1\}$ and $p = \frac{1}{2}$. ■

Proof of Proposition 3. In a pandering equilibrium both policymaker types are reelected. A good citizen becomes a candidate iff $w^G \leq w + p$. A bad citizen becomes a candidate iff $w^G \leq w + p(1-b)$. Then, the fraction of good policymakers is $\pi(\Pi, w) = \frac{\Pi}{\Pi+(1-\Pi)\frac{w+p(1-b)}{w+p}\frac{W^G}{W^B}}$. This ratio is increasing in Π and decreasing in w. Denote $\hat{\pi} = \frac{\Pi}{\Pi+(1-\Pi)\frac{W^G}{W^B}}$. If $\bar{v}(\hat{\pi}|b, p) > p$ then a truthful equilibrium is played and $\pi(\Pi, w) = \hat{\pi}$, which does not depend on the public office wage w. If $\bar{v}(\hat{\pi}|b, p) < p$ then a pandering equilibrium is played and $\pi(\Pi, w) = \frac{\Pi}{\Pi+(1-\Pi)\frac{w+p(1-b)}{w+p}\frac{W^G}{W^B}}$. ■

Proof of Proposition 4. Direct Selection. When policymakers play truthfully, the voter's equilibrium response is "keep iff popular policy." That is because $\tilde{\pi}_t(x_t = 1) =$

$\frac{\pi_t p}{\pi_t p+(1-\pi_t)p(1-b)} > \pi_t$ and $\tilde{\pi}_t(x_t=0) = \frac{\pi_t(1-p)}{\pi_t(1-p)+(1-\pi_t)[1-p(1-b)]} < \pi_t$ and the voter's equilibrium continuation value is strictly increasing in π_{t+1}, as we will show below. The fraction of good policymakers at $t+1$ is $\pi_{t+1} = \pi_t(1+bp) - \pi_t^2 bp > \pi_t$. The voter's equilibrium continuation value is $V(\pi_{t+1}) = (1-bp+bp\pi_{t+1}) + \beta\,(1-bp+bp\pi_{t+2}) + ...$ and since the sequence (π_t) is strictly increasing, $V(\pi_{t+1})$ is strictly increasing in π_{t+1}.

Hierarchical Selection. When policymakers are truthful the voter's equilibrium response is "retain monitor iff keeps popular policymaker and fires unpopular policymaker." That is because $\tilde{\mu}_t\,(x_t=1, y_t=keep) = \frac{\mu_t[\pi_t p+(1-\pi_t)p(1-b)]}{\mu_t[\pi_t p+(1-\pi_t)p(1-b)]+(1-\mu_t)[\pi_t p(1-b)+(1-\pi_t)p(1-b)]} > \mu_t$, $\tilde{\mu}_t\,(x_t=0, y_t=keep) = \frac{\mu_t(1-p)}{\mu_t(1-p)+(1-\mu_t)[\pi_t(1-p)+(1-\pi_t)(1-p+bp)]} < \mu_t$, $\tilde{\mu}_t\,(x_t=1, y_t=fire) = 0 < \mu_t$, and $\tilde{\mu}_t\,(x_t=0, y_t=fire) = 1 > \mu_t$. The fraction of good policymakers at $t+1$ is $\pi_{t+1} = \pi_t\,[1+(2\mu_t-1)\,bp] - \pi_t^2\,(2\mu_t-1)\,bp > \pi_t$. The voter's equilibrium continuation value $V(\pi_{t+1})$ is strictly increasing in μ_{t+1} because the sequences (π_t) and (μ_t) are strictly increasing. The hierarchical selection rate is $\frac{\pi_{t+1}-\pi_t}{\pi_t} = bp(1-\pi_t)\,(2\mu_t-1)$, slower than the direct selection rate $\frac{\pi_{t+1}-\pi_t}{\pi_t} = bp(1-\pi_t)$, since $\mu_t < 1$.

The officer differential is $[\pi_H + (1-bp)\,\pi_H] - [\pi_D + (1-bp)\,\pi_D] = (\pi_H - \pi_D)\,bp < 0$. ■

Proof of Proposition 5. The proof follows the steps in Lemma 1, Lemma 2 and Proposition 1 while adjusting the equilibrium continuation values to reflect the midterm resetting of incumbent policymaker and monitor types. ■

Proof of Proposition 6. Follows directly from Proposition 4. ■

References

Acemoglu, D., J. Robinson, and R. Torvik (2011) "Why Do Voters Dismantle Checks and Balances?" NBER working paper 17293.

Acemoglu, D., G. Egorov and K. Sonin (2011) "A Political Theory of Populism," NBER Working paper 17306.

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

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

Ashworth, S., E. Bueno de Mesquita, A. Friedenberg (2010) "Creating Incentives and Selecting Good Types Revisited," working paper, University of Chicago.

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.

Baron, D. (1994) "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review* 88(1): 33-47.

Barro, R. (1973) "The Control of Politicians: An Economic Model," *Public Choice* 14: 19-42.

Besley, T. (2004) "Paying Politicians: Theory and Evidence," *Journal of the European Economic Association* 2(2-3): 193-215.

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

Booth, R. and E. Vespa (2011) "The Dynamics of Municipal Spending: Theory and Evidence," working paper, New York University.

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: 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: 532-550.

Carter, M. and A. Sapp (1992) "A Comparative Analysis of Clauses in Police Collective Bargaining Agreements as Indicators of Change in Labor Relations," *American Journal of Police* 12(2): 17-46.

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.

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

Dahlberg, M. and E. Mork (2011) "Is there an Election Cycle in Public Employment? Separating Time Effects from Election Year Effects," working paper, Uppsala University.

Dell, M., B. Jones, and B. Olken (2008) "Climate Shocks and Economic Growth: Evidence from the Last Half Century," NBER Working Paper 14132.

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.

Enikopolov, R. (2010) "Politicians, Bureaucrats and Targeted Redistribution: The Role of Career Concerns," working paper, New Economic School.

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

Ferejohn, J. (1986) "Incumbent Performance and Electoral Control," *Public Choice* 50: 5-26.

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.

Fishback, P., W. 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: 36-71.

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

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 (2012) "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association*, forthcoming.

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.

George, L. M. and J. Waldfogel (2006) "The New York Times and the Market for Local Newspapers," *American Economic Review* 96(1): 435-477.

Heath, L. and K. Gilbert (1996) "Mass Media and Fear of Crime," *American Behavioral Scientist* 39: 379-386.

Iaryczower, M., G. Lewis and M. Shum (2010) "To Elect or to Appoint? Bias, Information, and Responsiveness of Bureaucrats and Politicians," working paper, Cal Tech.

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.

Judd, D.R. and T. Swanstrom (2010) *City Politics: The Political Economy of Urban America*, New York: Longman.

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

Laffont, J.J. (1999) "Political Economy, Information and Incentives," *European Economic Review* 43: 649-669.

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.

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.

McCrary, J. (2002) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *American Economic Review* 92(4): 1236-1243.

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.

North, D.C. and B.R. Weingast (1989) "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England," *Journal of Economic History* 49(4): 803-832.

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

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.

Prat, A. (2002) "Campaign Spending with Office-Seeking Politicians, Rational Voters, and Multiple Lobbies," *Journal of Economic Theory* 103(1): 162-189.

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

Rynecki, S. and M. Morse (1981) *Police Collective Bargaining Agreements: A National Management Survey*, Police Executive Research Forum, Washington, DC.

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, London School of Economics.

Stillman, R.J. (1974) *The Rise of the City Manager: A Public Professional in Local Government*, Albuquerque: Univ. of New Mexico Press.

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

Stone, H.A., D.K. Price, and K.H. Stone (1940) *City Manager Government in the United States: A Review After Twenty-Five Years*, Chicago: Social Science Research Council.

Strausz, R. (1997) "Delegation of Monitoring in a Principal-Agent Relationship," *Review of Economic Studies* 64 (3): 337-57.

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

Stucky, T.D. (2005) "Local Politics and Police Strength," *Justice Quarterly* 22(2): 139-169.

Trejo, S.J. (1991) "Public Sector Employment and Municipal Unions," *Industrial and Labor Relations Review* 45(1): 166-180.

Valletta, R.G. (1993) "Union Effects on Municipal Employment and Wages: A Longitudinal Approach," *Journal of Labor Economics* 11(3): 545-574.

Vlaicu, R. (2008) "Executive Performance under Direct and Hierarchical Accountability Structures," working paper, University of Maryland.

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. (2010) "Elected Versus Appointed Policymakers: Evidence from City Treasurers," NBER Working paper 15643.

Online Appendix

A.1 Data Sources and Measurement

Police and Crime

Officers Per Capita: Number of sworn police officers in city police department. *Sources:* Uniform Crime Reports (1960-2000).
Civilians Per Capita: Number of non-sworn employees in city police department. *Sources:* Uniform Crime Reports (1960-2000).
Violent Crime Rate: Total number of murder, rape, and robbery crimes reported in city in a given year, per 1000 population. *Sources:* Uniform Crime Reports (1975-2000).
Property Crime Rate: Total number of motor vehicle theft, larceny, burglary, and assault crimes in a city reported in a given year, per 1000 population. *Sources:* Uniform Crime Reports (1975-2000).
Collective Bargaining Law: Indicator variable equal to 1 in a year when state legislation provides the right for city police to bargain collectively, and 0 otherwise. *Sources:* NBER Public Sector Collective Bargaining Law Data Set (http://www.nber.org/publaw/), as updated by Vespa and Booth (2010).
High Percent Police in Union 1968: Indicator variable equal to 1 for cities with above median fraction of the police department unionized in 1968, and 0 otherwise. *Source:* Municipal Year Book (1968).

Geography

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

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

Presence of Swamp: County has a swamp.

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

Percentage of Clay in the Soil: Based on contemporary surveys by USDA soil scientists.

Soil Indicated Flood Frequency Index: The index is based on the 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.

Demographics

Population: 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).

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).

City Institutions

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

Non-Partisan Elections 1960: 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 1960: Fraction of city council seats that 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).

Election: Indicator variable equal to 1 in an election year, and 0 otherwise. If the election takes place before July 31 of the year, the previous year is coded as an election year. Election years are coded based on mayor elections for the mayor cities and based on city council elections for manager cities *Sources:* Municipal Year Book (various years), World Almanac (various years), www.ourcampaigns.com, city charters, newspaper articles.
City Government Spending: Total expenditure by city government, in thousands of 2000$. *Sources:* Census of Governments, City Government Finances (1972-2000).
High Local Newspaper Sales 1990-2000: Indicator variable equal to 1 for cities with an above median fraction of local newspaper sales per capita during the 1990 to 2000 period, and 0 otherwise. *Source:* George and Waldfogel (2006).

Policymaker Salary and Characteristics

Annual Salary (survey reported): Salary of manager or mayor in 2000$. *Sources:* ICMA Salaries of Municipal Officials (1992-1993,1995-2000).
Male (survey reported): Manager or mayor is male. *Sources:* ICMA Salaries of Municipal Officials (1992-1993,1995-2000).
White (survey reported): Manager or mayor is white. *Sources:* ICMA Salaries of Municipal Officials (1992-1993,1995-2000).
Male (name-based imputation): Manager or mayor is male. For imputation procedure see section A.3. *Sources:* World Almanac and Book of Facts (1960-2000).

1900 City Outcomes

This city-level data are from U.S. Census Bureau (1905) and U.S. Census Bureau (1906).
Officers Per Capita: Number of sworn police employees per resident.
Civilians Per Capita: Number of non-sworn police employees per resident.
Arrests Per Capita: Number of police department arrests per resident.
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: Number of city residents.

Climate

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 of years in the Local Flood Control Era, from 1900 to 1936, with annual city precipitation in the top 1 percent of the national precipitation distribution.

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

Median Precipitation 1900-2000: Median annual city precipitation from 1900 to 2000.

Fraction of Years in Top 1 Percent 1900-2000: Fraction of years from 1900 to 2000 with annual city precipitation in the top 1 percent of the national precipitation distribution

Fraction of Years in Top 5 Percent 1900-2000: Fraction of years from 1900 to 2000 with annual city precipitation in the top 5 percent of the national precipitation distribution

Fraction of Years in Top 10 Percent 1900-2000: Fraction of years from 1900 to 2000 with annual city precipitation in the top 10 percent of the national precipitation distribution

FFC Precipitation Shocks: Fraction of years in the Federal Flood Control Era, from 1937 to 1960, with annual city precipitation in the top 1 percent of the national precipitation distribution.

LFC Drought Shocks: Fraction of years in the Local Flood Control Era, from 1900 to 1936, with annual city precipitation in the bottom 1 percent of the national precipitation distribution.

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

LFC Hot Shocks: Fraction of years in the Local Flood Control Era, from 1900 to 1936, with city annual high temperature in the top 1 percent of the national high temperature distribution.

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

Median High Temperature 1900-2000: Median annual high temperature in a city from 1900 to 2000.

LFC Cold Shocks: Fraction of years in the Local Flood Control Era, from 1900 to 1936, with city annual low temperature in the bottom 1 percent of the national high temperature distribution.

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

Median Low Temperature 1900-2000: Median annual low temperature in a city from 1900 to 2000.

A.2 Data References

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

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

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

Federal Bureau of Investigation, *Uniform Crime Reports*.

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.

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

A.3 Name-Based Imputation of Policymaker Gender

To obtain a measure of policymaker gender for the full sample period 1960-2000 we use the manager and mayor names reported in the *World Almanac*. To impute gender we use the ICMA *Salaries of Municipal Officials* dataset to identify the modal gender for each policymaker first name. We then use this predicted gender with the actual first name of the policymaker from the *World Almanac* to assign gender. The imputation of gender from the policymaker name may be imperfect. We are unable to impute a policymaker gender when (i) the policymaker name in the *World Almanac* does not exist in the ICMA salary dataset, (ii) the policymaker name in the *World Almanac* is reported as an initial, not full first name, and (iii) the policymaker name is not reported in the *World Almanac*. Ultimately, we are able to impute gender for the policymaker in this fashion for 69.78 percent of the city-year observations in our data. We did not use similar name-based imputation for policymaker race due to the low degree of correlation found.[1]

[1]While it is possible in principle to impute race from policymaker names in a similar fashion in practice many policymaker names are less informative about race then they are about gender. For example, the fraction of policymakers who are directly reported as male in the ICMA salary data is 0.86. This is very

A.4 Instrument Validity and Robustness

As explained in Section 4 of the paper we obtain climate measures by first aggregating daily weather data to the yearly level. We then define an extreme weather event as a year when a city's weather measure is in the upper nth percentile of the national weather-years 1900-2000 distribution. We then calculate the fraction of years that a city has an extreme weather event during the local flood control era (1900-1936) and for the full century. We create these cross-sectional variables for high and low precipitation, and hot and cold temperatures. In addition, to control for differences in typical weather across cities we calculate median precipitation, high temperature and low temperature for each city using data from 1900-2000. Tables A1 to A5 report results relevant to the validity and robustness of our IV approach.

Precipitation Shocks and Flood Risk. We use precipitation to measure the timing of flood crises because comprehensive local data on the occurrence of floods does not exist in the local flood control period and flood occurrence may partially reflect the choice of flood control technology. To provide a first check of 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 contemporary USDA soil scientists.[2] This index measure is based on local flood frequency class as determined by the interpretation of soil properties and other evidence gathered during soil survey fieldwork.

Table A1 examines the relationship between soil indicated flood frequency and three candidate measures of precipitation shock: the fraction of years that city precipitation is in the top 1st, 5th, and 10th percentile of the national precipitation distribution for the century.[3] In principle any of these measures could represent a promising instrument. To avoid concerns about weak instruments leading to finite sample bias in our IV estimates we seek the strongest possible predictor of flood frequency.

All Local Flood Control (LFC) precipitation shocks measures predict flood frequency.

close to the 0.93 fraction imputed to being male based on the policymaker name and the twocorrelation between the two measures is 0.53 for the years when both measures exist. In contrast, the fraction white from the name-based measure is 0.29, far lower than the 0.79 fraction white in the ICMA salary data. In addition, the correlation between the two policymaker race measures is only 0.03 for the years when both measures exist.

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

[3]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).

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 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.

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 A2 show that cities experiencing precipitation shocks in the local flood control era are very similar across a number of observables to cities that do not.

In Table A3 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 (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 (drought) shocks are also more likely to adopt manager government. 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. Columns (4), (9) and (5), (10) show little relationship between extreme temperature shocks and manager government adoption or police employment.[4]

[4]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.

Another potential concern with our IV approach is that LFC shocks might lead to crises with persistent effects on policy making, regardless of government form, in violation of our exclusion restriction. 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 A4 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 institutional changes have little to do with flood control policy.

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 we next examine whether our estimates change with alternative versions of the instrument. We consider three different types of instrument specifications.[5]

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 1 percent or top 5 percent of the national distribution.[6] 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.[7]

The results in Table A5 reveal that our central results above change little when we use 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 10%) and the point estimates are of similar magnitude to the estimates in our baseline analysis. In columns (6) to (10) while the estimates of the manager coefficient for civilian employment are less stable, none of the estimates indicate a statistically significant negative manager effect.

[5] 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 over 10, to mitigate finite sample issues with the IV estimates.

[6] In unreported analysis we have also estimated the baseline models above with precipitation shocks defined by the top 5 percent. We obtain very similar results with a point estimate on the manager coefficient of -0.77 and a city-clustered standard error of 0.34. However, the first stage F-statistic is only 7.39 and so weak instrument bias may be a concern.

[7] We obtain additional first stage strength by adding water availablility controls: *presence of very large river*, *presence of large river*, *presence of small river*, and *located on the coast*.

TABLE A1: Precipitation Shocks Timing and Soil Indicated Flood Frequency: OLS Estimates

Dependent Variable =	Soil Indicated Flood Frequency Index					
	(1)	(2)	(3)	(4)	(5)	(6)
LFC Precipitation Shocks:						
Fraction of Years in Top 1 Percent 1900-1936	31.79***			16.68**		
	(6.91)			(7.26)		
Fraction of Years in Top 5 Percent 1900-1936		8.67***			1.72	
		(2.08)			(2.63)	
Fraction of Years in Top 10 Percent 1900-1936			8.79***			3.09
			(1.48)			(1.99)
Century Precipitation Shocks:						
Fraction of Years in Top 1 Percent 1900-2000	-3.89			1.52		
	(5.09)			(5.83)		
Fraction of Years in Top 5 Percent 1900-2000		-3.35*			2.17	
		(1.79)			(2.42)	
Fraction of Years in Top 10 Percent 1900-2000			-4.85***			-0.38
			(1.36)			(2.06)
Median Precipitation 1900-2000	0.67	0.60	0.14	-2.55	-3.20*	-3.32
	(0.86)	(1.01)	(1.08)	(1.66)	(1.90)	(2.11)
Additional Controls:						
Geography & Division	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 the data appendix. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States in 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 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 the coast, percentage of clay in the soil* and Census division fixed effects. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE A2: 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	2.77	1.79	4.54	-11.22	-2.82	3.51	496.30*
Shocks	(1.92)	(5.04)	(4.85)	(7.54)	(5.93)	(4.13)	(192.60)
Median Precipitation 1900-	0.38	-0.35	-0.83	5.80***	1.38	-0.10	-102.85***
2000	(0.34)	(0.72)	(0.58)	(1.72)	(0.94)	(0.78)	(28.65)
Additional Controls:							
Geography & Division	No	No	No	No	No	No	No
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 the data appendix. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States in 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 A3: Weather Shocks, City Government Form, and Police Employment: OLS Estimates

Dependent Variable =	Manager					Log (Officers Per Capita)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LFC Precipitation Shocks	10.14***	10.18***				-4.91**	-4.62**			
	(3.04)	(3.03)				(2.11)	(2.01)			
FFC Precipitation Shocks		0.32					2.26			
		(2.39)					(1.68)			
Century Precipitation	-6.03***	-6.37**				2.68	0.26			
Shocks	(2.20)	(3.05)				(1.63)	(2.42)			
Median High Precipitation	-0.61	-0.61				0.75***	0.74***			
1900-2000	(0.46)	(0.46)				(0.24)	(0.24)			
LFC Drought Shocks			3.64*					0.01		
			(1.90)					(1.12)		
Century Drought Shocks			-4.46**					0.49		
			(2.14)					(1.28)		
Median Low Precipitation			-1.17***					1.09***		
1900-2000			(0.38)					(0.21)		
LFC Hot Shocks				-10.12**					2.91	
				(4.48)					(2.75)	
Century Hot Shocks				10.42					-2.87	
				(4.72)					(2.90)	
Median High Temperature				1.39					0.87***	
1900-2000				(0.51)					(0.27)	
LFC Cold Shocks					1.70					-0.19
					(1.03)					(0.36)
Century Cold Shocks					-1.95*					0.11
					(1.06)					(0.38)
Median Low Temperature					0.54					0.19
1900-2000					(0.58)					(0.34)

Additional Controls:										
Geography & Division	No	No	No	No	No	No	No	No	No	No
Number of Observations	248	248	248	248	248	243	243	243	243	243

Notes: Authors' calculations with the city data described in the data appendix. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States in 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 A4: 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 1900-2000	-0.61	-1.02***	0.27	-0.83**
	(0.46)	(0.35)	(0.41)	(0.34)
Additional Controls:				
Geography & Division	No	No	No	No
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 the data appendix. The unit of observation is a city for the sample of the largest 248 self-governing cities in the United States in 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 A5: City Government Form and Police Employment: Alternative IV Estimates

Dependent Variable =	Log (Officers Per Capita)					Log (Civilians Per Capita)				
Alternative Specification=	Precipitation Shock Definition		Reform Era (RE) Definition		Drought IV	Precipitation Shock Definition		Reform Era (RE) Definition		Drought IV
Definition=	Any Year Top 1 %	Any Year Top 5 %	RE 2: 1909-1936	RE 3: 1900-1929	Shock: Bottom 1 %	Any Year Top 1 %	Any Year Top 5 %	RE 2: 1909-1936	RE 3: 1900-1929	Shock: Bottom 1 %
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Second Stage: Dependent Variable = Log (Police Employment 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)	(0.24)	(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 1900-2000	-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)	(0.35)	(0.87)	(0.63)	(0.64)	(0.64)	(0.59)
	Panel B : First Stage: Dependent Variable = Manager									
RE Precipitation Shocks	0.50***	0.31***	12.04***	8.41***		0.50***	0.31***	12.03***	8.42***	
	(0.12)	(0.07)	(3.09)	(2.39)		(0.12)	(0.08)	(3.09)	(2.38)	
Century Precipitation Shocks	-0.10	-0.19**	-5.05***	-6.15***		-0.10	-0.19**	-5.08***	-6.20***	
	(0.08)	(0.08)	(1.78)	(2.15)		(0.08)	(0.08)	(1.78)	(2.15)	
Median High Precipitation 1900-2000	-0.98**	-1.28**	-0.74	-0.57		-0.98**	-1.28**	-0.74*	-0.57	
	(0.47)	(0.52)	(0.43)	(0.45)		(0.47)	(0.53)	(0.43)	(0.45)	
LFC Drought Shocks					5.01***					4.98***
					(1.54)					(1.54)
Century Drought Shocks					-5.98***					-5.96***
					(1.73)					(1.73)
Median Low Precipitation 1900-2000					-1.13***					-1.13***
					(0.36)					(0.36)

Excluded Instrument F-Statistic:	17.34 [0.0000]	17.61 [0.0000]	15.22 [0.0001]	12.44 [0.0005]	10.54 [0.0013]	17.17 [0.0000]	17.48 [0.0000]	15.15 [0.0001]	12.48 [0.0005]	10.45 [0.0014]
Additional Controls:										
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geography & Division	No	No	No	No	No	No	No	No	No	No
Number of Observations	9,974	9,974	9,974	9,974	9,974	9,850	9,850	9,850	9,850	9,850
Number of Clusters	248	248	248	248	248	248	248	248	248	248

Notes: Authors' calculations with city data from 1960 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States in 1900. Standard errors clustered at the city level reported in parentheses. The models in columns (1)-(4) and (6)-(9) only include year fixed effects as additional controls. The models in columns (5) and (10) also include indicator variables for *presence of very large river*, *presence of large river*, *presence of small river* and *located on the coast.* * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE A6: City Government Form and Crime Rates

Dependent Variable =	Log(Violent Crime Rate)	Log(Violent Crime Rate)	Log(Property Crime Rate)	Log(Property Crime Rate)
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Manager	0.08	0.26	0.14***	0.12
	(0.07)	(0.39)	(0.03)	(0.23)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Geography & Division	Yes	Yes	Yes	Yes
Demographic	Yes	Yes	Yes	Yes
Sample	1975-2000	1975-2000	1975-2000	1975-2000
Number of Observations	6,131	6,131	6,148	6,148
Number of Clusters	248	248	248	248

Notes: Authors' calculations with city data from 1975 to 2000 as described in the data appendix. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States in 1900. Each column reports the results from one regression. Standard errors clustered at the city level reported in parentheses. The models include 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 the coast, percentage of clay in the soil* and Census division fixed effects, as well as demographic controls: *population, fraction of population non-white, fraction of population college graduate,* and *median household income*. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

TABLE A7: Crime Rates and Policymaker Salary: OLS Estimates

Dependent Variable =	Log(Policymaker Salary)			
	(1)	(2)	(3)	(4)
Manager	0.53*	0.55*	0.53	0.57*
	(0.31)	(0.30)	(0.34)	(0.32)
Log(Violent Crime Rate)	0.01	0.01		
	(0.04)	(0.03)		
Log(Violent Crime Rate) × Manager	0.02	0.01		
	(0.03)	(0.03)		
Log(Property Crime Rate)			0.01	0.01
			(0.03)	(0.02)
Log(Property Crime Rate) × Manager			0.00	0.00
			(0.03)	(0.03)
Additional Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes
Demographic	No	Yes	No	Yes
Sample	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000	1992-1993, 1995-2000
Number of Observations	1,052	1,052	1,055	1,055
Number of Clusters	233	233	233	233

Notes: Authors' calculations with city data from 1992-1993, 1995-2000 as described in the data appendix. The unit of observation is a city-year for the sample of the largest 248 self-governing cities in the United States in 1900. Standard errors clustered at the city level reported in parentheses. The models include city and year fixed effects. The models in columns (2) and (4) also include the following demographic controls: *population, fraction of population non-white, fraction of population college graduate*, and *median household income*. * indicates significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.