Bureaucratic Sabotage and Policy Inefficiency*

Kun Heo† and Elisa M. Wirsching‡

April 12, 2024

Abstract

Poor public service provision creates electoral vulnerability for incumbent politicians. Under what conditions can bureaucrats exploit this to avoid reforms they dislike? We develop a model of political accountability in which a politician must decide whether to enact a reform of uncertain value, and a voter evaluates the incumbent based on government service quality, which anti-reform bureaucrats can sabotage. We find that bureaucratic sabotage leads to two types of policy inefficiency depending on voters' perceptions of the reform's merit. Sabotage either deters politicians from enacting beneficial reforms due to electoral risks (under-reform) or prompts them to implement excessive reforms by providing bureaucrats as a scapegoat (over-reform). This result arises because obfuscation by sabotage affects voter inference differently based on their prior beliefs.

*We appreciate help from Benjamin Blumenthal, Eric Dickson, Sanford Gordon, Catherine Hafer, Alexander Hirsch, Dimitri Landa, Justin Melnick, Ken Shotts, Tara Slough, Carolina Torreblanca, Hye Young You, and Antoine Zerbini. We also thank participants at the NYU PIPER lab, the NYU PE workshop, and APSA 2023 for helpful comments and suggestions.

†Ph.D. Candidate, New York University

‡Ph.D. Candidate, New York University

1 Introduction

In 2021, protests erupted among municipal workers in several cities over vaccine mandates for their employees. Consequently, garbage accumulation became noticeable in various neighborhoods across the country. For instance, preceding the implementation of New York City's COVID-19 vaccine mandate, sanitation workers in Staten Island and South Brooklyn left trash uncollected for over a week (ABCNews, 2021). The city's sanitation commissioner, Grayson, attributed this service lapse to the impending vaccine mandate, acknowledging that municipal garbage trucks were completing their routes with half-empty loads (Gross, 2021). This raised concerns of a deliberate slowdown by sanitation workers to push back against vaccination requirements.

Similarly, recent research indicates that local police forces adapt their service provision to express dissent against police reforms and influence city politics. Officers of the San Francisco police department, for example, strongly opposed the progressive policies of District Attorney Chesa Boudin. Making police accountability his central policy issue, Boudin charged several officers in a historical excessive-force prosecution and pushed for criminal justice reforms to slim the carceral state. During the recall campaign, San Francisco residents repeatedly raised concerns to city officials and the media that police weren't responding to crime and justified their lack of engagement with the District Attorney's reluctance to press charges (Knight, 2021; Swan, 2021). In an interview, Chesa Boudin complained that "we've seen, on body-worn camera footage, police officers telling victims there's nothing they can do and, 'Don't forget to vote in the upcoming recall election.'" (Pearson, 2022) This blame-shifting by police might have resonated with voters in a high-crime environment, who recalled the progressive District Attorney by a significant margin. Immediately after the "unfriendly" attorney was successfully removed, police notably intensified their effort in making stops and arrests again (Kyriazis, Schechter and Yogev, 2023). Similarly, in New York City, police punished city officials who supported significant cuts to the department's budget in 2020 by disproportionately slowing response times to 911 calls in these "non-aligned" council districts (Wirsching, 2023). As part of this political strategy, law enforcement unions employed various tactics to ensure voters hold political representatives accountable for poor public service provision. These tactics span from publicly shaming city officials for their policies and blaming them for crime incidents in "non-aligned" districts to instigating fear about rising crime rates if progressive city officials remain in office (Blumgart, 2020; Wirsching, 2023).

While examples of strategic work slowdowns by city bureaucrats abound, the logic, conditions, and consequences of such *bureaucratic sabotage* remain puzzling and largely unexplored. Why would bureaucrats engage in actions that disrupt public services when voters know they can do so for political reasons? And if this sabotage affects how voters view reform policies, why would politicians ever push for reforms that bureaucrats oppose? In this paper, we study how and when politicians' electoral vulnerability motivates bureaucrats to sabotage service provision and how the possibility of sabotage affects an incumbent's choice to enact reforms. With this framework, we shed light on a unique source of political power for bureaucrats and its consequences for public policy.

In canonical models of electoral accountability, voters rely on incomplete information from the incumbent's policy decisions and government performance to make their voting decisions. It is well-established that politicians' desire to impress the voter can affect their policy choices, sometimes acting against voters' interests (Canes-Wrone, Herron and Shotts, 2001; Gersen and Stephenson, 2014).[1]

Yet, we know less about how bureaucrats' political interests affect these dynamics. We integrate bureaucratic sabotage into a model of electoral accountability. A purely office-motivated incumbent chooses between a reform of unknown value and the status quo after observing a private signal about the reform's value. The voter observes the incumbent's policy choice together with a noisy signal about government service quality and decides whether to retain the incumbent for a second period or to elect the challenger. Importantly, service quality is affected by both the reform's inherent value and the bureaucrats' performance. Bureaucrats who have a fixed yet unknown degree of distaste for the reform can privately decide to engage in costly sabotage of public service provision, e.g., by refusing to work diligently. This complexity obscures the voter's interpretation since he is unable to assign responsibility for poor service provision. For example, when a community experiences a decline in safety after police reform (e.g., a budget cut), it becomes challenging for a resident to determine whether the decrease in security is due to the reform itself or a change in the behavior of police officers post-reform. Even if the reform could improve government services relative to the status quo, voters may observe lower service quality due to bureaucratic sabotage. Hence, bureaucrats can exploit their intermediary role in government to advance their policy preferences by lowering the incumbent's reelection chances in favor of the anti-reform challenger through sabotage. We show that, in equilibrium, incumbents implement reform if they are sufficiently confident about its value, bureaucrats sabotage if they are sufficiently anti-reform, and voters re-elect their representative if government performance is sufficiently high.

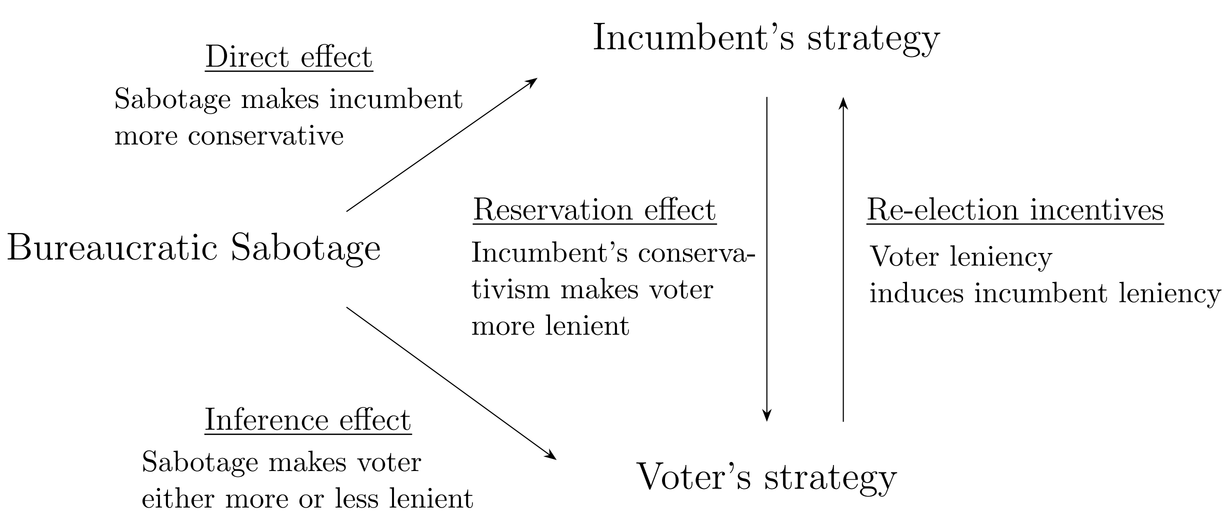
Our model produces several key insights. First, we show how bureaucratic sabotage affects the political equilibrium between the incumbent and the voter. A naive conclusion may

[1]For a review, see Ashworth (2012).

be that sabotage strictly discourages the incumbent from introducing reform by increasing the probability that the voter observes poor service provision. However, we find that the possibility of sabotage can either incentivize or deter incumbents from implementing reform, contingent upon the voter's prior beliefs. If reform is initially unpopular with the voter, sabotage leads to *under-reform* compared to the optimal level: incumbents fear bureaucratic sabotage and are hesitant to implement reform. Conversely, if reform is popular, sabotage leads to *over-reform*. In this case, sabotage increases the incumbent's electoral incentive to introduce reform, although the status quo is preferable.

This result arises because sabotage has two countervailing effects on the voter's observation and action. On the one hand, sabotage directly lowers the quality of public services, thus dissuading office-motivated politicians from pursuing reform (*direct effect*). On the other hand, when assessing public service delivery and adjusting his reelection intentions, the voter factors in the potential for sabotage (*inference effect*) and politicians' strategic response to it (*reservation effect*). As Figure 1 prefaces, we find that these effects do not necessarily point in the same direction, and the inference effect can dominate the direct and reservation effects in equilibrium. Consequently, sabotage can both encourage and discourage reform.

Figure 1: Mechanisms of Sabotage's Influence on Player Strategies



Second, to shed light on the exact mechanism that produces our main result, we unpack the inference effect for the voter. When deciding whether to reelect a reform-minded incumbent after observing a particular government outcome, voters seek to optimize the accuracy of their predictions. They want to correctly identify successful reforms (true positives) while avoiding reforms that appear promising but are bound to fail (false negatives). Sabotage affects this trade-off by (1) decreasing true positives and (2) increasing false negatives. However, the impact of this obfuscation on the voter's choice is ex-ante unclear because the relative size of these effects depends on the voter's prior on the reform's value. If the reform is already popular with the voter, making them forgiving toward the reforming incumbent, the voter is primarily concerned about false negatives, and sabotage further increases their willingness to reelect the incumbent despite poor services. However, somewhat surprisingly, if the reform is unpopular with the voter (i.e., they demand high-quality services to re-elect a reforming incumbent), sabotage strengthens their stringency, even though the voter is well aware that sabotage originates from bureaucrats. This occurs because the voter is especially weary of false positives, and, in the presence of sabotage, the voter tends to attribute high government performance to luck (false positives) rather than a successful policy (true positives).

Third, we show that bureaucrats' incentive to sabotage is non-monotonic with respect to the voter's prior belief about the reform's value. After incumbents introduce the reform, bureaucrats sabotage public services to reduce politicians' re-election probability by affecting voters' inference about the reform. The incentive to sabotage, therefore, depends on whether the voter is susceptible to information that bureaucrats mediate. When voters highly favor the reform, bureaucrats' incentive to sabotage is low because bureaucrats only have limited ability to sway voters' support for the incumbent who initiated the reform. However, as the reform becomes *less* popular, this effect weakens, and sabotage becomes more likely. Conversely, if voters are already pessimistic about the reform's benefit, bureaucrats have little incentive to resort to costly sabotage to tarnish the reputation of politicians because the voter is already likely to perceive the reform as a failure. This effect weakens as the reform becomes *more* popular, thus increasing the probability of sabotage. As a result, bureaucrats are most incentivized to sabotage when voters are torn between the reform and the status quo and, therefore, more open to interpreting poor public service provision as informative regarding the reform's merits.

Fourth, we explore how the voter's concern for politicians' private ideology can impact these policy inefficiencies. In our baseline model, all incumbents are fully office-motivated. In an extension of the model, we allow for the possibility that incumbents are reform zealots, i.e., strictly prefer reform over no reform. Counterintuitively, when the pool of incumbents worsens (i.e., the probability of reform zealots increases), the policy inefficiencies actually *improve*. This result arises because, with the presence of reform enthusiasts, pragmatist incumbents are less able to blame bureaucrats for poor service provision. In other words, the existence of reform enthusiasts diminishes the capacity of a pragmatic incumbent to mimic the behavior of such reform enthusiasts.

2 Related Literature and Contributions

We make multiple contributions to existing scholarship on bureaucratic politics, interest group influence, political accountability, and policy sabotage.

First, our theory addresses a fundamental debate in bureaucratic politics between the public choice school—particularly Tullock (1965), Downs (1967), and Niskanen (1971)—on one side, and theories of bureaucratic control and delegation on the other (see Moe (2012) for a detailed review of these streams of work). Building on the idea that bureaucrats are rational actors primarily driven by self-interest, Niskanen (1971) famously argued that bureaucrats' private information about the true cost of government production empowers them to influence policy. In Niskanen's narrative, the interaction between the incumbent and bureaucrats is characterized by bargaining—bureaucrats, armed with private information, extract informational rent by presenting a take-it-or-leave-it offer to the incumbent, whose alternatives are less favorable than accepting the offer on average. In this framework, bureaucrats are the primary strategic actor, while politicians take the backseat.

In contrast, influential theories of legislative control of bureaucracies (Miller and Moe, 1983; McCubbins, 1985; McCubbins, Noll and Weingast, 1987; Banks and Weingast, 1992) criticized Niskanen's framework for ascribing out-sized power to bureaucrats. Instead, they highlighted that the incumbent also acts strategically, employing various tools to fight back when bureaucrats try to extort their position. Specifically, they conceptualized the relationship between the incumbent and the bureaucrats as a top-down principal-agent relationship where the incumbent designs incentives, rules, and monitoring mechanisms to minimize agency loss while still leveraging bureaucratic expertise.

We reconcile these two longstanding ideas on bureaucratic politics by synthesizing a principal-agent perspective on strategic politicians with the notion of politically powerful bureaucrats. We highlight that the principal-agent relationship between the incumbent and the bureaucrats is nested within a more complex accountability mechanism, where the incumbent depends on voters whose decisions can be influenced by bureaucrats. Hence, we propose a *bottom-up* principal-agent model where bureaucrats can sway the incumbent's policy decisions by leveraging their private information, exploiting the incumbent's electoral vulnerability, and adjusting their work effort.

Second, we join a recent stream of literature in political economy that distinguishes the roles of politicians and bureaucrats in public goods provision (Ujhelyi, 2014; Yazaki, 2018; Foarta, 2023; Slough, 2024). Yet, existing research primarily focuses on issues of moral hazard in bureaucracies. A central assumption is that the successful execution of a policy requires considerable, often strenuous effort from bureaucrats. The central tension lies between the incumbent, who desires to appear competent to voters through successful policy implementation, and bureaucrats, who might fail to exert the necessary effort due to its costs or low bureaucratic quality. Consequently, incumbents adjust their policy decisions and delegation to bureaucrats based on factors influencing bureaucrats' motivation.

In contrast, we assume that bureaucrats incur a cost from politically motivated sabotage, e.g., because it goes against their objective to serve the public and complicates their own work.[2] The primary tradeoff for bureaucrats lies between the cost of sabotage and their influence in the election and policy-making. Sabotage can enhance bureaucrats' payoff in two main ways. By engaging in sabotage, they can make an unwanted policy look bad to the voter even if it is optimal for voter welfare, thus damaging the reputation of a politician who advocated reform. Additionally, the anticipation of sabotage may dissuade incumbents from introducing the unwanted policy in the first place. Hence, unlike bureaucrats in a moral hazard situation, bureaucrats in our framework strategically ruin the policy outcome to affect who their principals are and their choices in office.

Third, we speak to growing work on the role of bureaucrats as interest groups within government. We are not the first scholars to zoom in on bureaucrats' agenda power that stems from the electoral vulnerability of their elected principals. Moe (2006) argued that because bureaucrats can exercise power through the electoral process, they can influence who their principals are and what policies they choose in office, i.e., the electoral vulnerability of politicians turns them into "agents of the agents." Moe (2006) and an extensive subsequent empirical literature on the political influence of bureaucrats, particularly their public sector unions, indeed shows that bureaucrats are one of the most influential interest groups on all levels of government. Yet, importantly, this literature almost exclusively considers *direct* routes of political influence for bureaucrats through collective bargaining (Moe, 2009, 2011; Anzia and Moe, 2015; Paglayan, 2019; Zoorob, 2019), union endorsements (Moe, 2006; Hartney and Flavin, 2011; Hartney, 2022), electoral mobilization of their members (Leighley and Nagler, 2007; Anzia, 2014; Flavin and Hartney, 2015), political contributions (Moe, 2011; DiSalvo, 2015), or direct lobbying (Anzia, 2022).[3] In contrast, we explain how bureaucrats can exert policy influence through their role in government, i.e., *by the mere virtue of being bureaucrats*. Hence, we make an epistemic argument that the bureaucrats' intermediary role is one of the fundamental aspects of bureaucracy and interest group influence to which scholars should pay attention.

Fourth, we describe and micro-found a novel explanation for why bureaucratic agen-

[2]We discuss this modeling choice in detail in Section 3.7.

[3]An exception is Wirsching (2023), who considers effort shirking by police officers as a crucial way of political influence for police unions after significant budget cuts in New York City in 2020.

cies might undermine the very programs and services they provide. Several scholars have attempted to characterize recent surges of bureaucratic sabotage at the federal level, especially during the Trump administration. Some have argued that agencies sabotage their own work because, in an environment where securing legislation from Congress is difficult, US presidents pursue retrenchment by asking the administrative state to undermine itself (Noll, 2022). Others have considered the expressive benefits of "guerrilla" forms of government (O'Leary, 2020) and found that bureaucratic resistance is a result of bureaucrats navigating the moral dilemma between norms of professionalism and personal beliefs about policy (Kucinskas and Zylan, 2023). Notably, the voters are absent from these accounts. In contrast, we focus on how voters' dependence on bureaucrats to learn about policy outcomes can result in bureaucratic sabotage as a strategic choice.

Our model is closely connected to the models that study the relationship between voter information and electoral accountability. As in the many canonical models of electoral accountability following Fearon (1999), the voter faces two types of agency problems: moral hazard and adverse selection, with elections as the only tool to address both. Moral hazard arises because the incumbent has a private signal about the value of the reform. To resolve the moral hazard problem, the voter should commit to an election rule that incentivizes the incumbent to introduce the reform only if it improves the government service quality in expectation based on the observed service quality. However, since the voter has to choose between the status quo and the reform, he cannot commit to such a rule, except in a knife-edge case where the two rules coincide. We add a strategic third player (i.e., bureaucrats) to this standard setting and highlight how their unobservable signal obfuscation (i.e., sabotage) affects the political equilibrium. As we show, this obfuscation has substantial effects by changing the voter's equilibrium response to service quality as an informative signal about the incumbent's performance. Namely, the meaning of the signal crucially depends on its context, which determines the voter's reaction to the signal (Ashworth, Bueno De Mesquita and Friedenberg, 2018) and the incumbent's action (Ashworth, Mesquita and Friedenberg, 2017; Bueno de Mesquita and Tyson, 2020; Bils and Izzo, 2023). In contrast to existing models where the change in the information environment is exogenous, the change operates through the endogenous actions of strategic players (i.e., bureaucrats) in our model.

This paper is also closely connected to research studying the strategic obstructions in political accountability settings (Patty, 2016; Fong and Krehbiel, 2018; Hirsch and Kastellec, 2022; Gieczewski and Li, 2021). There are two key differences between our argument and existing work. First, the voter observes neither the saboteur's preferences nor actions. Second, existing models focus on the saboteur's incentive and its implications for voter perception. In contrast, we study the implications of bureaucratic sabotage for *policy-making*.

3 Baseline Model

Consider a two-period ($t = 1, 2$) political agency model with an incumbent politician (she), a representative voter (he), bureaucrats (they), and a challenger. There is an election after $t = 1$ where the voter chooses between the incumbent and the challenger as a new officeholder for $t = 2$. Players do not discount their future payoffs.[4]

3.1 Policy-Making

In $t = 1$, the incumbent must decide whether to introduce a reform or not. $t = 1$ is a window for reform, and the officeholder cannot introduce it in $t = 2$; it can only be repealed then.[5]

The reform's original value to government outcome for each period, $x \in \{0, 1\}$, is unknown to the public. x can be interpreted as the reform either succeeding ($x = 1$) or failing ($x = 0$). The common prior is $\Pr[x = 1] = 1/2$.[6] If the incumbent chooses not to introduce a reform, the status quo policy is implemented. The value of the status quo policy to government outcome is known as $q \in (0, 1)$. $1 - q$ can be interpreted as the probability that reform outperforms the status quo.

Before making the decision, the incumbent *privately* observes a noisy but informative signal $r \in [0, 1]$ about the reform's value x. r is drawn from a conditional density $f(r|x)$ such that satisfies the monotone likelihood ratio property with respect to x; $f(r|1)/f(r|0) : [0, 1] \to \mathbb{R}_+$ is an increasing bijection; a high r signals the higher likelihood for $x = 1$ relative to $x = 0$, $r = 0$ means that $x = 0$, and $r = 1$ means that $x = 1$. We further assume that $f(r = 1/2|x = 1) = f(r = 1/2|x = 0) = 1/2$.[7]

3.2 Government Outcome and Sabotage

The incumbent's policy decision is public and observed by other players. In addition to the incumbent's choice, the voter observes the government outcome g. g is affected by three factors. First, it is affected by the incumbent's policy decision. Let $a \in \{0, 1\}$ indicate the incumbent's decision to introduce the reform.

Second, it is affected by bureaucrats' sabotage. Such sabotage can include a variety of measures, including dragging their feet in delivering services, overlooking service infractions, misusing their authority, or mismanaging funds. Bureaucrats get disutility $-\kappa < 0$ from

[4] This is a simplifying assumption that does not affect the results qualitatively.

[5] We discuss the justifications for and the implications of this assumption in 3.7.

[6] This is for simplification. Any interior prior works in a similar way.

[7] This assumption purchases us the normative anchor to evaluate the incumbent's equilibrium behavior.

having the reform in place in each period. κ is the bureaucrats' private information, and other players only know that $\kappa \in [0, 1]$ is drawn from a uniform distribution.[8]

If the incumbent introduces reform, bureaucrats privately observe the real value of reform x and then privately decide whether to sabotage it or not.[9] Let $s \in \{0, 1\}$ indicate the bureaucrats' choice to sabotage the reform.

Sabotage is costly for them, so they take a known cost of $c \in [0, 1]$ if they sabotage. Sabotage is costly for bureaucrats because it goes against their motivation to produce government output. This motivation can be intrinsic in the sense that bureaucrats receive utility from contributing effort to government production regardless of their political preferences, akin to what scholars call "public service motivation" (Besley and Ghatak, 2005; Forand, Ujhelyi and Ting, 2022). c might also represent extrinsic motivation to contribute to high-quality public goods because better public service facilitates bureaucrats' jobs. For instance, lower crime rates reduce the need for constant policing, and police can be assumed to benefit from a sufficient level of effort. Additionally, effective sabotage likely requires costly collective action among bureaucrats, thus inducing costs c for their unions and collective organizations.[10]

Lastly, government outcome g is affected by an i.i.d. idiosyncratic shock ϵ drawn from a log-concave density $h(\cdot)$ that has full support on $\mathbb{R}$ and is symmetric around 0. Let $H(\cdot)$ denote its associated CDF.

Formally, government outcome g in period 1 is given by

$$g = a(1 - s)x + (1 - a)q + \epsilon.$$

After observing g, the voter chooses between the incumbent and the challenger for $t = 2$. Let $e \in \{0, 1\}$ indicate the voter's decision to reelect the incumbent.

3.3 Politicians' Preferences

Politicians get 1 by winning the election and 0 otherwise. They may also get a payoff from policy-making while they are in office, depending on their preference.

In the baseline model, the incumbent only cares about reelection and does not get any payoff from policy-making per se. On the other hand, the challenger is known to be an "anti-reform zealot," who prefers the status quo. The challenger gets $Z > 1$ if and only if she chooses the status quo and 0 otherwise.

[8] This is a simplifying assumption. A log-concave distribution on $[0, 1]$ works.

[9] Bureaucrats' knowledge about x does not affect the qualitative results.

[10] We further discuss this modeling choice in Section 3.7.

3.4 Second Period

The new officeholder in $t = 2$ can only choose to repeal the reform or not if the incumbent introduced it in $t = 1$. The government outcome in period $t = 2$ is produced by

$$\tilde{g} = \begin{cases} \tilde{a}(1-\tilde{s})x + (1-\tilde{a})q + \tilde{\epsilon} & \text{if } a = 1 \\ q + \tilde{\epsilon} & \text{if } a = 0. \end{cases}$$

where $\tilde{a} \in \{0, 1\}$ indicates the new officeholder's decision to retain the reform, $\tilde{s} \in \{0, 1\}$ is the bureaucrats' decision to sabotage in the second period at the cost of c, and $\tilde{\epsilon}$ is the shock in the second period drawn from $h(\cdot)$.

Simply speaking, $\tilde{g} = x + \tilde{\epsilon}$ only if both officeholders choose the reform over the status quo and bureaucrats choose not to sabotage in the second period. If not, $\tilde{g} = q + \tilde{\epsilon}$.

3.5 Payoffs

The voter gets

$$g + \tilde{g}.$$

The incumbent gets

$$e$$

The challenger, who is zealously anti-reform, gets

$$e[1 + (1-\tilde{a})Z].$$

The bureaucrats get

$$-a\kappa - cs - a\tilde{a}\kappa - c\tilde{s}.$$

3.6 Timing

1. The incumbent privately observes an informative signal, r, about the value of the reform.

2. The incumbent publicly chooses whether to introduce the reform ($a = 1$) or not ($a = 0$).

3. If the incumbent chooses to introduce the reform, then bureaucrats observe their disutility from reform κ and its original value x.

4. The bureaucrat privately chooses whether sabotage the reform ($s = 1$) or not ($s = 0$).
5. The government outcome g is produced, and the voter observes it.
6. The voter chooses between the incumbent and the challenger as the new officeholder in the election.
7. If the reform is in place, then the new officeholder chooses whether to repeal the reform or not.
8. Government outcome for Period 2 is produced.
9. Payoffs are realized, and the game ends.

3.7 Modeling Choices

A few clarifications are warranted before solving the model.

- **Political control of bureaucracy**: We depart from standard models of bureaucracy in that politicians cannot directly control bureaucrats' behavior. Many formal models of political delegation have explored how politicians can restrict bureaucratic behavior, including ex-ante limits on bureaucratic discretion and ex-post monitoring of their behavior (Brehm and Gates, 1997; Epstein and O'Halloran, 1999; Huber and Shipan, 2002; Huber and Mccarty, 2004). Yet, we omit the possibility of politicians monitoring bureaucrats in our model for two main reasons. First, recent work suggests that bureaucrats can undermine political control, for example, by selectively sharing information during Congressional oversight hearings (Ban, Park and You, 2024). Second, and more importantly, the institutional dynamics enabling politicians to monitor bureaucrats mainly apply to interactions between the US Congress and the federal bureaucracy. In contrast, our theory primarily applies to street-level bureaucracies, like police officers or waste collectors, who regularly interact with voters and enjoy a high degree of independence from political control. Street-level bureaucrats benefit from strong job protections provided by civil service regulations and robust public sector unions. Additionally, their roles often require significant autonomy and discretion due to the ambiguity of their task environment, which demands bureaucrats' individual and immediate decision-making (Wilson, 1978; Lipsky, 1980).
- **Costly sabotage**: We also differ from standard settings where bureaucrats choose the amount of *costly effort*. Most models on bureaucracy assume that bureaucrats have an inherent benefit from shirking their duties and show how the equilibrium

level of bureaucrats' effort responds to politicians' threat of oversight and punishment (Brehm and Gates, 1997; Epstein and O'Halloran, 1999; Huber and Shipan, 2002), their investment in bureaucratic quality (Slough, 2024), or their personnel policies (Ting, 2021). In contrast, in our setting, politically motivated shirking (rather than working) is costly for the bureaucrat. As mentioned earlier, the costs associated with sabotage may stem from various factors. First, bureaucrats may act contrary to their intrinsic motivation to serve the public interest, a concept known as public service motivation (e.g., Forand, Ujhelyi and Ting (2022)). Second, poor service provision can complicate their daily tasks—for instance, crime prevention becomes more challenging for police in areas with high crime rates. Last, there's the added challenge for bureaucrats and their unions of overcoming costly collective action to mobilize political sabotage.

It is important to note that while we abstract away from the standard effort shirking by bureaucrats, we do not claim that the moral hazard of bureaucrats is irrelevant to explain public service provision. Rather, we essentially normalize moral hazard problems in the bureaucracy and focus on how the political motivations of bureaucrats can lead to *additional* shirking. That is, one can think of bureaucratic sabotage in our model as an equilibrium perturbation of bureaucrats' optimal level of effort given the usual incentives to work. To see this point, consider the motivating example of police sabotaging Chesa Boudin's tenure in office in the days prior to the recall election. Absent the political incentives to oust Boudin, it is reasonable to assume that police exerted a level of effort that minimized the usual costs of working while maximizing its benefits (e.g., avoiding monitoring and possible punishment for shirking by city officials and the public). Our model sheds light on how and why political dynamics led to further reductions of effort *away from that equilibrium* before the recall election. Indeed, recent evidence suggests that police significantly reduced the number of arrests and stops immediately prior to the recall election, while their enforcement returned to the "baseline" level after Boudin was replaced (Kyriazis, Schechter and Yogev, 2023).

- **Policy continuity**: Since the incumbent does not inherently benefit from the policy they choose, incumbents have no incentive to revert their policy choice in the second period. There are two reasons why we make this assumption. First, in reality, it is rare that politicians flip their position on a specific policy reform and therefore, it is natural to assume that the voter expects policy continuity when he reelects the incumbent.

  Secondly, as much as the incumbent not changing the policy upon her reelection is an artifact of our setup, the incumbent's incentives to flip her policy in the second period would be an artifact of an alternative two-period model with policy-motivated incumbents where the game ends without an election in the second period (at least within the current setup where the reform's value, x, is period-invariant). If $r^* > 0$, the incumbent introduces the reform if and only if the expected value of the reform given a private signal r is larger than some cutoff value. Therefore, the voter updates about the reform's value based on whether the incumbent has introduced the reform or not in the first period. If the incumbent wins the election with (or without) introducing the reform, the voter's posterior valuation of the reform is higher (or lower) than the prior. Thus, if there is another election after the second period, it is unreasonable for the voter to expect that the incumbent will revert her policy decision upon reelection because if it is incentive compatible for the incumbent to introduce/not to introduce the reform in the first period due to reelection concerns, it should also be incentive compatible in the second period for the same reason. Hence, if there is a second election in the game, incumbents' flipping their policy is not sequentially rational.[11]

4 Analysis

The solution concept is weak Perfect Bayesian Equilibrium (Equilibrium, henceforth). We study pure strategy equilibrium where the voter reelects the incumbent if and only if $g \geq g^*$, the bureaucrats sabotage the reform if and only if $\kappa \geq \kappa^*$, and the incumbent introduces the reform if and only if $r \geq r^*$.

4.1 Normative Benchmark

To facilitate our analysis of how bureaucratic sabotage induces policy inefficiencies, we first establish a normative benchmark for politicians' equilibrium behavior. Given r, the incumbent's posterior belief that reform succeeds is

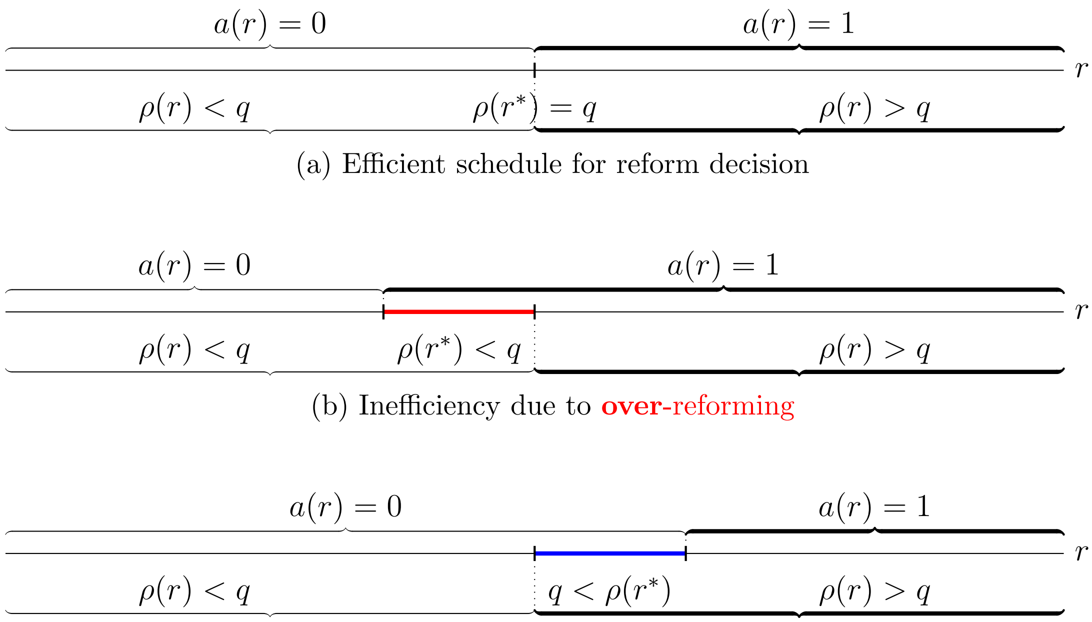
$$\mathrm{E}[x|r] = \Pr[x = 1|r] = \rho(r) \equiv \frac{f(r|1)}{f(r|1) + f(r|0)},$$

which is monotonically increasing in r by f's monotone likelihood ratio property.

Figure 2 illustrates the normative benchmark. If the incumbent maximizes voter welfare, she implements the reform if and only if $\rho(r) \geq q$, so $\rho(r^*) = q$. If $\rho(r^*) < q$, there is a range of r such that the incumbent implements reform even if she deems it undesirable for the voter. We refer to this case as *over-reforming.* On the other hand, if $\rho(r^*) > q$, there is a range of r where the incumbent does not introduce the reform even if it is optimal for the voter. We refer to this case as *under-reforming.*

[11]See Andreottola (2021) for another formal explanation for policy continuity.

Figure 2: Normative Criterion and Policy Inefficiencies



(a) Efficient schedule for reform decision

(b) Inefficiency due to **over**-reforming

(c) Inefficiency due to **under**-reforming

4.2 Second-Period Behavior

If the challenger wins the election, then she reverts back to the status quo policy since $Z > 0$. If the incumbent wins the election, she keeps the policy from the previous period. Bureaucrats do not sabotage the reform because they incur $-c$ if they sabotage.

4.3 The Voter's Inference and Election Decision

We start by analyzing how the voter updates his belief about the reform policy and casts his vote. Given the second-period behavior of the players, the election outcome has the following implications for the government outcome in $t = 2$, $\tilde{g}$.

Remark 1 *If the incumbent wins without introducing the reform or the challenger wins, then the status quo policy will be implemented in $t = 2$ and $\tilde{g} = q + \tilde{\epsilon}$. If the incumbent wins after introducing the reform, then the reform will be implemented in $t = 2$ and $\tilde{g} = x + \tilde{\epsilon}$.*

If the incumbent chooses the status quo, the voter is indifferent between the two candidates. We assume that the incumbent gets reelected with probability $1/2$ if she chooses the status quo since the voter is ex-ante indifferent between the incumbent and the challenger, holding the policy fixed.[12]

[12]This assumption helps us anchor the normative benchmark where the expected value of the reform based on r is the same as that of the status quo.

Assumption 1 *The incumbent gets reelected with probability* $1/2$ *if she chooses the status quo.*

If the incumbent introduces reform, then the voter reelects the incumbent if and only if g signals that the expected value of the reform is larger than q.

Suppose that the incumbent introduces reform if and only if $r \geq r'$ and bureaucrats sabotage if and only if $x = 1$ and $\kappa > \kappa'$. Then, the voter infers that the conditional expectation of the reform's value x from g is $\mathrm{E}[x|r \geq r', g]$, which is given by

$$\Pr[x = 1|g, r \geq r'] = \frac{1}{1 + \hat{L}(g, r', \kappa')}$$

where

$$\hat{L}(g, r', \kappa') = \frac{\Pr[x = 0] \Pr[g, r \geq r'|x = 0]}{\Pr[x = 1] \Pr[g, r \geq r'|x = 1]} = \frac{1 - F(r'|0)}{1 - F(r'|1)} \frac{h(g)}{h(g) + \kappa'\Big(h(g-1) - h(g)\Big)}.$$

$\mathrm{E}[x|r \geq r', g] \geq q$ if and only if $\frac{1}{1+\hat{L}(g,r',\kappa')} \geq q \iff \hat{L}(g, r', \kappa') \leq \log \frac{1-q}{q}$.

Further inspection of $\hat{L}$ provides the following insights about the voter's inference and behavior:[13]

Lemma 1 *If* $r' = 1$ *or* $\kappa' = 1$, $E[x|g] = 1/2$.

If $r' < 1$ *and* $\kappa' > 0$,

1. $E[x|r \geq r', g]$ *is increasing in* g. *Therefore, there exists a unique* $\hat{g}^*$ *that solves*

$$\hat{L}(g, r', \kappa') = \log \frac{1-q}{q}; \tag{1}$$

2. $E[x|r \geq r', g]$ *is increasing in* r'*;*

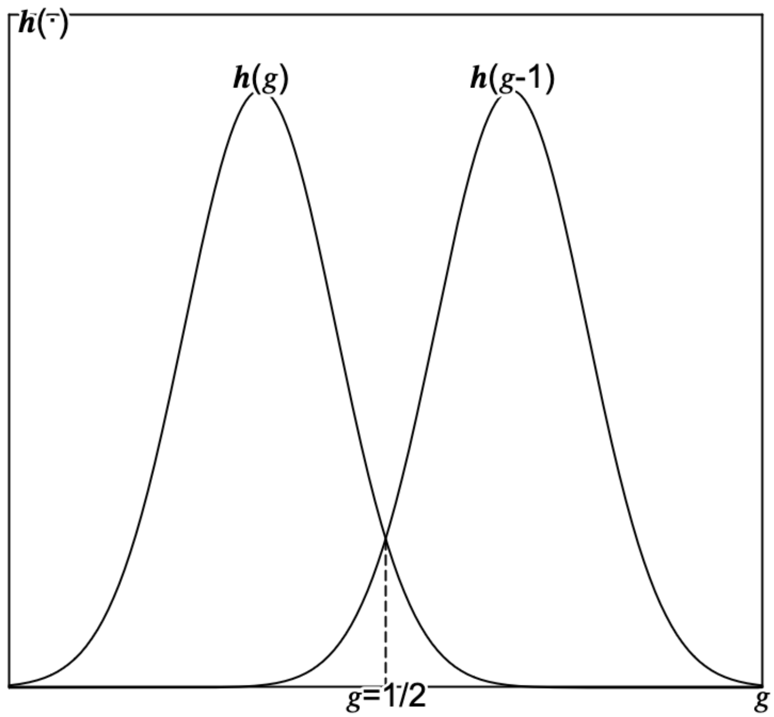
3. $E[x|r \geq r', g]$ *is decreasing in the probability of sabotage* $(1 - \kappa')$ *if and only if* $g > 1/2$, *and increasing in* $1 - \kappa'$ *if and only if* $g < 1/2$.

Intuitively, Lemma 1 states that at $g = 1/2$, the reform is equally likely to succeed or fail ($\mathrm{E}[x|g = 1/2] = 1/2$) because the likelihood that $g = 1/2$ is drawn from the density $h(g)$ is exactly the same as the likelihood that it is drawn from the density $h(g-1)$. If $g > 1/2$, then g is more likely to be drawn from $h(g-1)$, so the voter infers that the reform is more likely to succeed ($x = 1$) and not have been sabotaged ($s = 0$) than either to fail ($x = 0$) or have been sabotaged ($s = 1$). In contrast, if $g < 1/2$, then the voter's inference works in

[13] All proofs are relegated to the Appendix A.

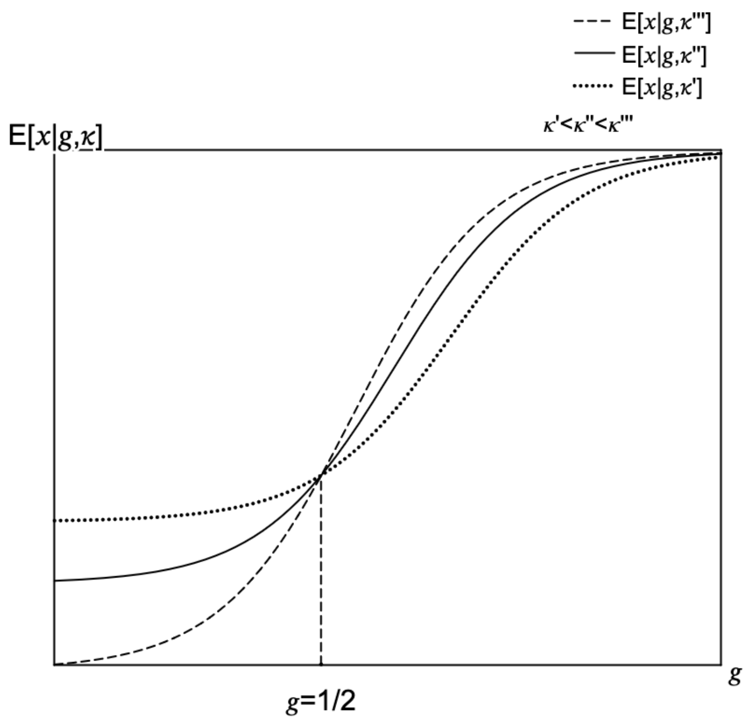
the opposite way. Figure 3 illustrates this logic. (Observe that $h(g) > h(g-1)$ if and only if $g < 1/2$.).

Figure 3: Likelihood Comparison



Additionally, as the incumbent becomes more stringent about when to implement reform (i.e., requires a more favorable signal r to implement reform, Lemma 1-3.), the voter increases her trust in the reform's success. Bureaucratic sabotage, in contrast, has two opposing effects on voter inference depending on the level of the government outcome g (Lemma 1-4.): As sabotage becomes more likely ($1 - k'$ increases), voters expect bureaucrats to interfere more, and the policy performance can be obfuscated more by bureaucratic action. Figure 4 shows how this impacts the voter's posterior expectations about the reform's value. Consider a high likelihood of sabotage (dotted line, low k'). The voter is less inclined to ascribe poor government outcomes (low g) to a failed reform and becomes less stringent with the incumbent. At the same time, he has less confidence in the reform's value when observing high-quality services (high g), thus becoming more stringent with the incumbent. The next section unpacks the mechanism for these two countervailing effects, which we call *inference effects of sabotage*.

Figure 4: Sabotage's Countervailing Effects on Voter Inference



4.3.1 Sabotage and Voter Learning: Understanding Inference Effects

As the incumbent can strategically choose whether to introduce reform or not and bureaucrats can sabotage reform, g is a *obfuscated* signal of the reform's true value of x. To understand the effect of strategic obfuscation on the voter's learning, consider the benchmark case where neither player intervenes with g, and the voter observes $g = x + \epsilon$.

Suppose that, for an arbitrary cutoff g', the voter concludes that the reform will work if he observes a "positive" signal $g \geq g'$ and it will not work if he observes a "negative" signal $g < g'$. Then, we can define four events, shown in Table 1.

The voter faces a Goldilocks problem in choosing the optimal g', i.e., he cannot be either too lenient or too stringent. If he is too lenient and chooses a low g', then a positive signal $g \geq g'$ does not necessarily mean that the reform outperforms the status quo. Thus, he wants to pick a high enough g' so that the positive predictive value (PPV), i.e.

$$\Pr[x = 1 | g \geq g'] = \frac{\Pr[TP]}{\Pr[TP] + \Pr[FP]}$$

is large enough. This ensures that the reform is a better choice than the status quo in expectation, given $g \geq g'$.

On the other hand, if the voter is too stringent so that g' is too high, he risks not choosing

Table 1: Confusion Matrix for Voter Inference

		Prediction: $g < g'$	Prediction: $g > g'$
Actual condition	$x = 1$	FN	TP
	$x = 0$	TN	FP
		False omission rate (FOR) $\frac{FN}{TN+FN}$	Positive predictive value (PPV) $\frac{TP}{TP+FP}$

Notes: FN denotes false negatives; TN denotes true negatives; TP denotes true positives; FP denotes false positives.

the reform when it is better than the status quo. So, he wants to pick a low enough g' such that the false omission rate (FOR), i.e.

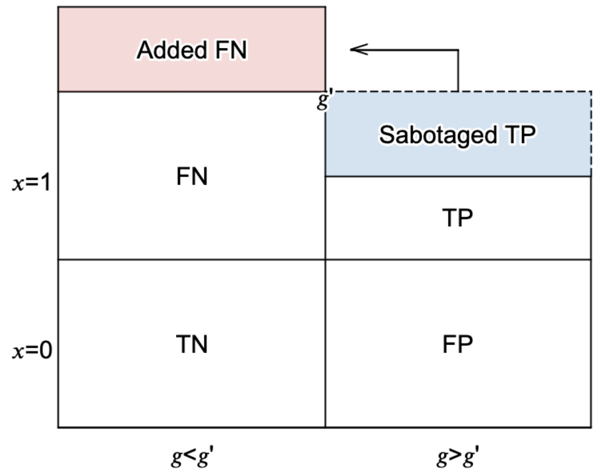
$$\Pr[x = 1|g < g'] = \frac{\Pr[FN]}{\Pr[TN] + \Pr[FN]}$$

is small. This ensures that the reform is expected to perform worse than the status quo given $g < g'$. Evidently, at the cutoff g', the voter is indifferent between the risk of true positives and false negatives.

Consider the impact of including the incumbent. Note that with a cutoff r', the incumbent introduces reform only if its expected value is high enough since r is an informative signal about x. Hence, if the incumbent chooses a cutoff r', failed reforms are filtered with some probability. In effect, the incumbent's strategy truncates the conditional distribution of reform's value from below. This truncation affects the voter's strategy. Particularly, the voter lowers g' since the truncation from below decreases $Pr[FP]$ and $Pr[TN]$ and, therefore, increases $Pr[x = 1|g]$. Thus, to maintain indifference at the cutoff, the voter lowers g' as the incumbent filters more failed reform by increasing r'.

Finally, consider the additional obfuscation through bureaucratic sabotage. Assume bureaucrats sabotage reform that would otherwise be successful and supported by voters (i.e., $x = 1$ and $g > g'$). Hence, with sabotage, some of the true positives turn into false negatives with probability $(1 - \kappa')$. This change has two countervailing effects. Figure 5 provides the intuition for this result. Firstly, it *decreases* $\Pr[x = 1|g \geq g']$ by lowering $\Pr[TP]$ (the blue shaded area "Sabotaged TP"). Intuitively, knowing that sabotage lowers the likelihood that the voter observes $g > g'$ when it is indeed valuable (i.e. when $x = 1$), the voter is inclined

Figure 5: The Effect of Sabotage on Voter Learning



The blue shaded area "Sabotaged TP" illustrates the PPV effect. The red shaded area "Added FN" illustrates the FOR effect.

to attribute a high $g > g'$ to mere luck rather than its actual value (i.e., a false positive). Formally, for the probability of sabotage $1 - \kappa'$,

$$\Pr[x = 1|g \geq g'] = \frac{\kappa' \Pr[TP]}{\kappa' \Pr[TP] + \Pr[FP]} < \frac{\Pr[TP]}{\Pr[TP] + \Pr[FP]}.$$

We call this the *PPV effect*.

Secondly, the change from TP to FN *increases* $\Pr[x = 1|g < g']$ by increasing $\Pr[FN]$ (the red shaded area "Added FN"). Namely, when the voter takes into account the fact that some of the negative signals that he observes are due to sabotage, his evaluation of the reform given a negative signal will increase as sabotage becomes more likely. That is,

$$\Pr[x = 1|g < g'] = \frac{\Pr[FN] + (1 - \kappa') \Pr[TP]}{\Pr[FN] + (1 - \kappa') \Pr[TP] + \Pr[TN]} > \frac{\Pr[FN]}{\Pr[FN] + \Pr[TN]}.$$

We call this the *FOR effect*.

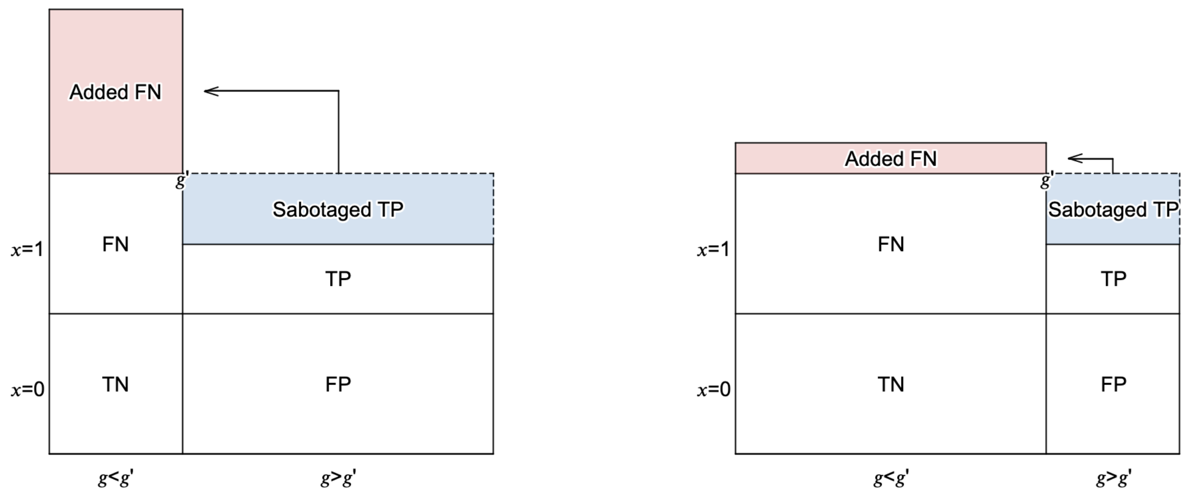
Which effect dominates depends on the initial level of g'. See Figure 6 for an illustration. If g' is high enough so that $g \geq g'$ is rare, the voter is more worried about false positives than false negatives—the FOR effect is low and dominated by the PPV effect.[14] In contrast, if g' is low, the voter faces higher risks of false negatives—the FOR effect is more likely to dominate the PPV effect.[15] Taken together, the effect of sabotage on voter behavior depends on what

[14]See Appendix B for calculations of these quantities based on Figure 6.

[15]The logic above is similar to that of the main results in Heo and Landa (2024). For further formal discussion on the decision problems with a stochastic process, see Patty and Penn (2023).

type of wrong inference the voter is most worried about. If the PPV effect dominates the FOR effect, the voter is better off being more stringent and choosing a higher g'. In contrast, if the FOR effect dominates the PPV effect, the voter is better off being more lenient and choosing a lower g'.

Figure 6: Sabotage's Effects on Voter Inference Conditional on g'



(a) When g' is low: FOR dominates PPV; sabotage decreases g'

(b) When g' is high: PPV dominates FOR; sabotage increases g'

It is noteworthy that this result depends on the assumption that bureaucrats can only change TP into FN by sabotaging the reform. For instance, even if bureaucrats do not know x when they make their decision on sabotage, as long as sabotage can affect g's distribution only when the reform actually works, the logic above holds.

4.4 Bureaucrats' Sabotage Incentive

We now study bureaucrats' optimal strategy given the voter's behavior. Suppose that the voter reelects the incumbent who introduces the reform if and only if $g \geq \hat{g}^*$. Then bureaucrats get $-\kappa\Big(1 - H(\hat{g}^* - 1)\Big)$ if they do not sabotage and $-\kappa\Big(1 - H(\hat{g}^*)\Big) - c$ if they do sabotage. Thus, they sabotage if and only if

$$\begin{aligned} &-\kappa\Big(1 - H(\hat{g}^* - 1)\Big) \leq -\kappa\Big(1 - H(\hat{g}^*)\Big) - c \\ &\hat{\kappa}^* \equiv \frac{c}{H(\hat{g}^*) - H(\hat{g}^* - 1)} \leq \kappa. \end{aligned} \tag{2}$$

Section 4.7 provides a detailed discussion of the implications of equation (2).

4.5 Incumbent's Decision and Unique Equilibrium

The incumbent's decision over the reform depends on its implication for her reelection prospect. The incumbent introduces the reform if and only if she is weakly more likely to get reelected when she introduces it than when she sticks with the status quo. Recall that the probability that the incumbent wins the election with the status quo is $1/2$.

Given the conditional probability that the reform works, $\rho(r)$, the incumbent gets

$$\rho(r)\Bigg(\kappa'\Big(1-H(\hat{g}^*-1)\Big)+(1-\kappa')\Big(1-H(\hat{g}^*)\Big)\Bigg)+[1-\rho(r)]\Big(1-H(\hat{g}^*)\Big)$$

if she introduces reform. Thus, she introduces reform if and only if

$$\rho(r)\Bigg(\kappa'\Big(1-H(\hat{g}^*-1)\Big)+(1-\kappa')\Big(1-H(\hat{g}^*)\Big)\Bigg)+[1-\rho(r)]\Big(1-H(\hat{g}^*)\Big)\geq\frac{1}{2} \qquad (3)$$

4.6 Equilibrium Characterization of Sabotage

By bringing together all three actors, we can now characterize the equilibrium. Let $\kappa'=\hat{\kappa}^*(\hat{g}^*)$ in (3) to get

$$\rho(r)c+1-H(\hat{g}^*)\geq\frac{1}{2} \iff \rho(r)\geq\frac{H(\hat{g}^*)-1/2}{c}$$

given equation (2). Thus, we can define

$$\hat{r}^*(g)\equiv\rho^{-1}\Big(\frac{H(g)-1/2}{c}\Big) \qquad \text{and} \qquad L(g,\hat{r}^*(g))=\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}. \qquad (4)$$

The following statement describes the unique equilibrium of the main model.

Proposition 1 *In the main model, there exists a unique pure strategy equilibrium defined by* $\Big(g^*(q,c),\kappa^*(q,c),r^*(q,c)\Big)$ *such that*

$$L(g^*,\hat{r}^*(g^*(q,c)))=\frac{1-q}{q}, \qquad \kappa^*(q,c)=\hat{\kappa}^*(q,c), \qquad r^*(q,c)=\hat{r}^*(g^*(q,c));$$

In equilibrium,[16] there is a minimum government service quality, $g^*(q,c)$, such that the voter reelects a reforming incumbent if and only if he observes at least this minimum quality, $g\geq g^*(q,c)$. In turn, there is also a minimum signal about the reform's value, $r^*(q,c)$, such that the incumbent introduces the reform if and only if $r\geq r^*(q,c)$. In addition, bureaucrats

[16]For the equilibrium of the benchmark case without sabotage, see Appendix A.3.2.

engage in sabotage if their grievance from the reform κ is larger than some minimum value $\kappa^*(q, c)$.

4.7 Comparative Statics

We now consider how the exogenous parameters q and c affect equilibrium outcomes. Importantly, by comparing equilibrium outcomes across various levels of c, we can describe in detail how bureaucratic sabotage leads to different types of policy inefficiencies.

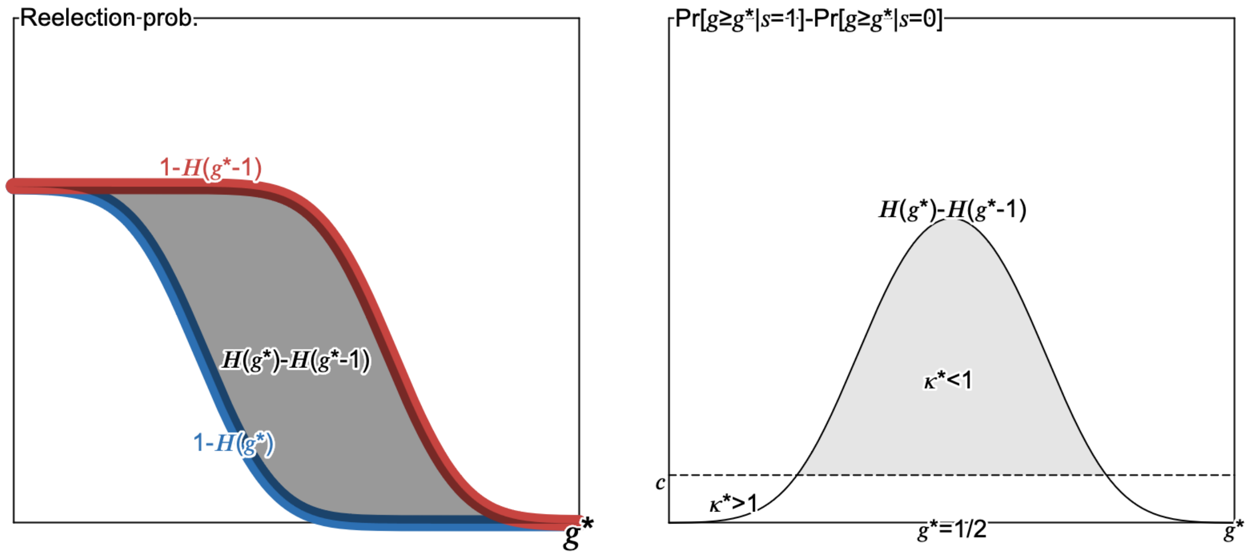
Proposition 2 (Bureaucrats Behavior) *q and c have the following effects on bureaucrats' equilibrium behavior:*

1. *As the status quo's value increases, bureaucrats' incentive to sabotage changes non-monotonically (a single-peaked curve); $\kappa^*(q, c)$ is U-shaped with respect to q;*
2. *There exists a unique $\bar{c} > 0$ such that sabotage occurs ($1 - \kappa^*(c, q) > 0$) only if $c < \bar{c}$.*
3. *Given $c < \bar{c}$, there exist unique $\underline{q}(c)$ and $\bar{q}(c)$ such that sabotage occurs only if $q \in (\underline{q}(c), \bar{q}(c))$.*
4. *$\bar{q}(c) - \underline{q}(c)$ decreases in c.*
5. *The ex-ante probability of sabotage $1 - \kappa^*$ is weakly decreasing in c.*

Why is the sabotage incentive non-monotonic with respect to q? Note that by sabotaging the reform that would have succeeded, bureaucrats horizontally shift the distribution of g left from $h(g - 1)$ to $h(g)$. This change reduces the probability that the voter observes a positive enough signal to reelect a reforming incumbent. Without sabotage, the probability that the voter observes $g \geq g^*$ is $1 - H(g^* - 1)$. Sabotage decreases this probability to $1 - H(g^*)$. We call this the *direct effect of sabotage*. As the gap between these two probabilities $H(g^*) - H(g^* - 1)$ increases, bureaucrats engage in sabotage with a smaller grievance κ and the ex-ante probability that they engage in sabotage $(1 - \kappa^*)$ increases.

See Figure 7 for intuition. When $g^* = 1/2$, the gap between the probability that the voter reelects the incumbent without and with sabotage $H(g^*) - H(g^* - 1)$ is largest. If the voter applies a lower or a higher cutoff than $1/2$, the probability that the incumbent gets reelected is less sensitive to sabotage. If $g^* < 1/2$, then the voter may still observe $g \geq g^*$ in spite of sabotage. As a result, bureaucrats engage in sabotage only if their grievances over the reform κ are high enough. On the other hand, if $g^* > 1/2$, it is unlikely that the voter observes $g \geq g^*$ anyway, even without sabotage, so the bureaucrats' incentive to sabotage is also smaller than when $g^* = 1/2$.

Figure 7: Sabotage's Marginal Effect on Re-election



(a) The X-axis is the voter's cutoff g^* and the Y-axis is the reelection probability. The red line is the probability of reelection as a function of g^* when $x = 1$ and $s = 0$ and the blue line is the same probability when $x = 0$ or $s = 1$. The grey area between the two lines captures the marginal effect of sabotage as a function of the voter's cutoff g^*.

(b) The X-axis is the voter's cutoff g^* and the Y-axis is sabotage's marginal effect on reelection probability. The line $H(g^*) - H(g^* - 1)$ is the sabotage's marginal effect as a function of the voter's cutoff g^* (The size of the grey area on Panel (a).). Notice that it is maximized at $g^* = 1/2$. The shaded area indicates the range of g^* where sabotage is incentive compatible.

As panel (b) of Figure 7 illustrates, the cost of sabotage truncates the marginal effect of sabotage on re-election probability from below, and sabotage is only incentive compatible for the bureaucrat if c is low enough relative to g^*. Intuitively, bureaucrats only sabotage if their costs to do so are not too high and are less likely to sabotage as these costs increase.

Proposition 3 (Voter Behavior) *q and c have the following effects on the voter's equilibrium behavior:*

1. *The voter applies a more stringent criterion for reelecting a reforming incumbent as the status quo's value increases; $g^*(q, c)$ is increasing in q.*

2. *$g^*(q, c)$ is weakly decreasing in c if and only if $q > q^\dagger > 1/2$.*

As the status quo policy's value q increases, the prior probability that the reform outperforms the status quo $(1 - q)$ decreases and the voter applies a more stringent criterion g^* to reelect a reforming incumbent. To understand the effect of changes in the costs of sabotage, we decompose the voter's equilibrium beliefs (4) into two components:

$$L(g, \hat{r}^*(g)) = \underbrace{\frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)}}_{\text{Reservation Effect, } \downarrow g^*} \cdot \underbrace{\frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}}_{\text{Inference Effect, } \uparrow g^* \text{ iff } g^*>1/2}$$

The second component captures the inference effect of sabotage. As discussed at length in section 4.3.1, the inference effect of sabotage makes a lenient voter even more lenient and a strict voter even stricter. However, in addition to sabotage's inference effect, we have to take into account how the voter reacts to the incumbent's equilibrium response to sabotage, represented by the first component of $L(g, \hat{r}^*(g))$.

As per the direct effect, sabotage makes reform riskier for the incumbent by increasing the likelihood of low-quality service[17] and, thus, induces her to raise the bar for the signal to introduce the reform. In turn, the incumbent's cautiousness convinces the voter to have a favorable perspective on the reform, even in light of low-quality service. We call this the *reservation effect of sabotage.*[18] Hence, whether the voter becomes more or less lenient as a result of changes in the cost of sabotage depends on the combination of these effects. Whereas the reservation effect amplifies the inference effect for low q, it weakens the inference effect for high q. As Figure 8 illustrates, sabotage, therefore, only induces the voter to be more strict with the incumbent for sufficiently high $q > q^\dagger > 1/2$.

Proposition 4 (Incumbent Behavior) *q and c have the following effects on the incumbent's equilibrium behavior:*

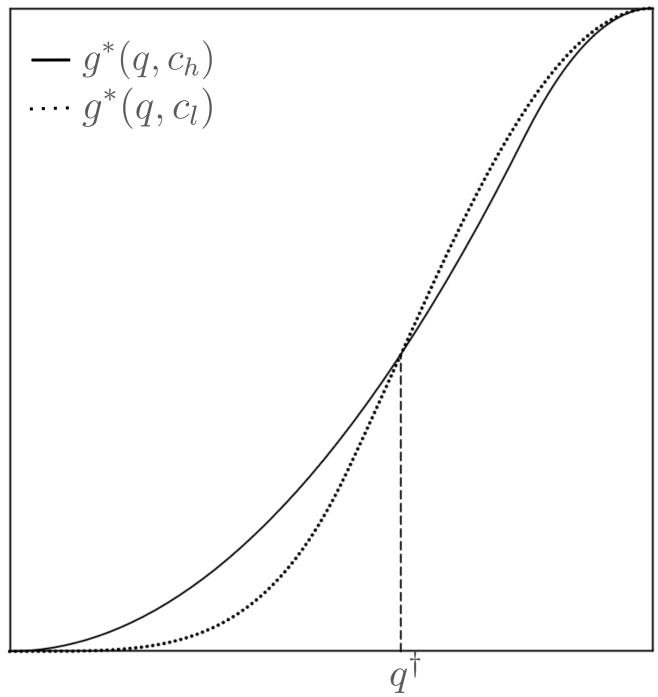
1. *As the status quo's value increases, the incumbent requires a higher signal to introduce reform; $r^*(q, c)$ is weakly increasing in q.*
2. *$r^*(q, c)$ is weakly increasing in c if and only if $q < q^{\dagger\dagger} < 1/2$.*
3. *There exists a unique $q_E(c) > 1/2$ such that $\rho(r^*(q, c)) \geq q$ if and only if $q \geq q_E(c)$.*
4. *$q_E(c)$ is increasing in c.*

Given the voter's tendency to become more stringent with large q, the probability that a reforming incumbent gets reelected decreases as q increases. In response, the incumbent

[17]To see, this notice that the incumbent becomes conservative toward the reform by raising the minimum signal that she must observe to introduce the reform *holding other factors fixed*; as c increases, $\hat{r}^*(g)$ decreases, $\hat{r}^*(g) = \rho^{-1}\left(\frac{H(g)-1/2}{c}\right)$, holding g fixed.

[18]In effect, the incumbent's decision to introduce the reform itself is an informative, positive signal about the reforms' value because the incumbent would not do so unless its value is sufficiently large. However, this logic may not hold if there is a type of incumbent who prefers the reform regardless of its value. We discuss the implications of such type in the Extension.

Figure 8: Sabotage's Equilibrium Effect on Voting Decision

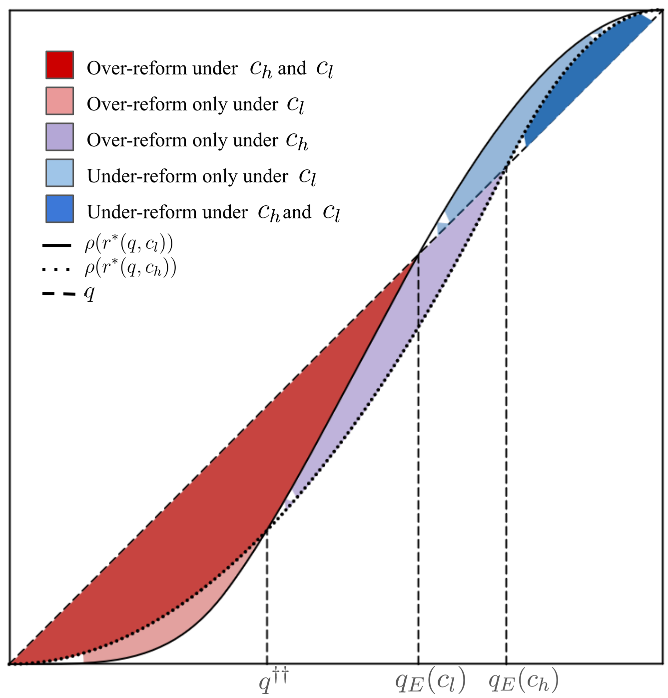


The voter's equilibrium cutoff g^* with high c_h and low c_l.

becomes more conservative and requires a higher signal r to introduce reform. Interestingly, a lower cost of sabotage can either increase or decrease the incumbent's willingness to reform, depending on the value of the status quo. Evidently, if the voter highly values the status quo (high q), bureaucratic sabotage leads the politician to be overly cautious with reform and thus induces *under-reform* relative to the normative benchmark. However, if the voter benefits little from the status quo and has large trust in the reform's effectiveness ex-ante, bureaucratic sabotage makes the incumbent more reckless and thus induces *over-reform* relative to the normative benchmark.

Figure 9 shows the implications of this result for policy efficiency by comparing the incumbent's equilibrium beliefs for high c_h and low c_l. Evidently, if the voter already values the reform ex-ante ($q < q^{\dagger\dagger}$), a high risk of sabotage (c_l) decreases efficiency by inducing the incumbent to *over-reform*, even if the status quo is normatively preferable. If reform is less popular, $q > q^{\dagger\dagger}$, the lower cost of sabotage makes the incumbent more conservative toward the reform. Importantly, the efficiency implication of the induced conservatism varies, depending on the reform's popularity. When the reform's popularity is intermediary high ($q^{\dagger\dagger} < q < q_E(c_l)$), lowering the cost of sabotage *increases* efficiency by reducing over-reforming inefficiencies. If the reform's popularity is intermediary low ($q_E(c_l) < q < q_E(c_h)$),

Figure 9: Sabotage's Equilibrium Effect on Incumbent's Policy-Making

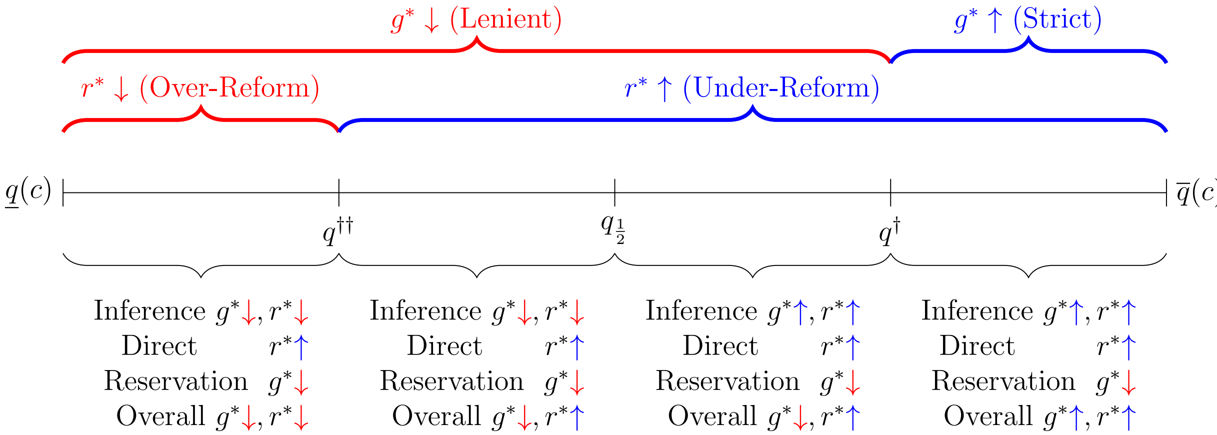


The X-axis is the value of the status quo (reform is ex-ante popular on the left is ex-ante unpopular on the right), and the Y-axis is the minimum expected value of reform that the incumbent introduces the reform as a function of the status quo and the sabotage's cost to bureaucrats. Notice that the square is divided into two areas across the 45-degree dashed line. Above the 45-degree line (northwest region), the expected value of the reform from the incumbent's signal is higher than q, so the reform is efficient, and the decision not to implement is inefficient due to under-reform. Below the 45-degree line (southeast region), the reform is worse than the status quo, and to implement it is inefficient due to over-reform.

lowering the cost of sabotage changes the type of policy inefficiency from over-reform to under-reform. Lastly, if the reform's popularity is very low ($q > q_E(c_h)$), the lower cost of sabotage decreases efficiency by causing too much conservatism on the incumbent's action (under-reforming).

What are the mechanisms for these differential effects of sabotage in equilibrium? Figure 10 illustrates how the overall effect of sabotage depends on the relative size of its constituent parts: The inference effect ($\kappa \to g^*$), the direct effect ($\kappa \to r^*$), and the reservation effect ($\kappa \to r^* \to g^*$).

Figure 10: Mechanisms of Equilibrium Effects of Sabotage



Across all values of q, the direct effect discourages the incumbent from introducing reform, and the reservation effect derived from it mitigates the incumbent's conservatism. In contrast, the direction of the inference effect depends on q. For the lowest levels of q ($\underline{q}(c) < q < q^{\dagger\dagger}$), sabotage makes the voter lenient through the large inference effect, which feeds into the incumbent's electoral incentives and dominates the direct effect of sabotage. In turn, sabotage leads to over-reform.

For intermediary low levels of q ($q^{\dagger\dagger} < q < q_{\frac{1}{2}}$), sabotage's effect on voter inference still makes him lenient toward the reform, but this inference effect is not strong enough to fully compensate the electoral risk from the direct effect of sabotage. Therefore, sabotage leads to under-reform.

For intermediary high levels of q ($q_{\frac{1}{2}} < q < q^{\dagger}$), the inference effect makes the voter more strict, but it is not large enough to cancel out the reservation effect. Therefore, overall, sabotage makes the voter more lenient. Again, voter leniency is too weak to mitigate the direct effect, and thus, sabotage leads to under-reform.

In the highest levels of q ($q^{\dagger} < q < \overline{q}(c)$), the inference effect makes the voter more strict by a large margin and dominates the reservation effect. Thus, the voter is overall more strict, and thus, the incumbent is also more conservative with sabotage

Taken together, the possibility of sabotage deters incumbents from implementing risky reform if the voter is already weary of reform failure. At the same time, the incumbent can leverage bureaucratic sabotage to gamble on reform if the voter is ex-ante optimistic about reform, even if it is doomed to fail.[19]

[19]Figure A2 shows the equilibrium predictions given these dynamics of over- and under-reforming.

5 Extension: Policy Motivated Incumbent and Electoral Selection

We now consider how our results change if we allow for different incumbent types. The incumbent is now known to be either a "reform zealot" or a pragmatic, office-motivated type as in the baseline model. The incumbent's type is her private information, and the voter knows she is a pragmatic type with probability $\tau \in [0, 1]$.

The zealous, pro-reform incumbent gets

$$Za + e(1 + Z\hat{a}).$$

We assume $Z > 1$, such that a zealous type incumbent introduces the reform with probability one and maintains it if she wins the reelection. Equation (4) then becomes

$$L^*_\tau(g) \equiv \log \left(\frac{\tau\Big(1 - F(\hat{r}^*(g)|0)\Big) + (1-\tau)}{\tau\Big(1 - F(\hat{r}^*(g)|1)\Big) + (1-\tau)} \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \right). \tag{5}$$

Observe the following:

Lemma 2 *$L(g, \hat{r}^*(g))$ is decreasing in τ.*

Proposition 5 *g^* and r^* are weakly decreasing in τ.*

Intuitively, as the probability of a reform zealot decreases, the incumbent's policy choice becomes more informative, and, for a given signal, the voter is more optimistic about the reform's value. He, therefore, is more lenient towards the incumbent, who he assumes to be a pragmatic type, and the pragmatic incumbent is more inclined to introduce reform as a result. This has interesting implications for policy inefficiencies: Over-reform becomes more prevalent because the fact that the voter expects the incumbent to be of a better type allows these better, pragmatic types to behave worse and free-ride on bureaucrats' involvement in government production. Incumbents' type and action are, therefore, strategic substitutes. Under-reform, however, becomes less common because pragmatic types, who are in better standing with the voter, fear bureaucratic sabotage less. Type and actions are strategic complements.

6 Empirical Examples

In this section, we provide examples of under- and over-reform and illustrate how our model helps to explain various dynamics in bureaucratic politics.

6.1 Examples of Under-Reforming

The deaths of unarmed Black Americans at the hands of police in recent years, including George Floyd, Daunte Wright, Breonna Taylor, and Tyre Nichols, have sparked a movement calling for sweeping police reform. In 2020, millions marched for police reform, and lawmakers across the aisle supported reform endeavors. Arguably, in light of the evidence of widespread racial disparities and misconduct by police across the country (Ba et al., 2021; Hoekstra and Sloan, 2022), reforms of law enforcement are desirable for US society. Yet, lawmakers' support for police reform faltered in recent years, and reform policies stalled (McCaskill, 2020; Pearson, 2022). Why?

Our model suggests how resistance by powerful police organizations and their threat to sabotage reform policies might have contributed to politicians' unwillingness to follow through with reforms aimed at police accountability and transparency. In particular, our results predict that incumbents shy away from desirable reforms because of bureaucrats' threat of sabotage if voters' are sufficiently weary about the effectiveness of reforms (q is high). The difficulties to eliminate "qualified immunity" for police officers are a clear case in point. In the aftermath of George Floyd's killing, federal and state lawmakers nationwide attempted to reverse a legal principle that effectively shields police officers from being sued for violating individuals' civil rights. Yet, the respective federal bill soon stalled in Congress, as bipartisan Senate negotiations failed, and by October 2021, at least 35 qualified-immunity bills had been withdrawn or died in state legislatures (Kindy, 2021). The outspoken opposition to these reforms by police organizations played an important part in this development. Police unions bought ads in local newspapers warning that officers might hesitate to pursue criminals due to concerns about potential lawsuits and asking readers to call state legislators in opposition to the reforms (Kindy, 2021). For example, a full-page advertisement in The Boston Globe in August 2020 by 13 Massachusetts police associations read, "We are your neighbors. The bill has been hastily thrust upon our legislative leaders without any involvement from people in law enforcement or any opportunity for public comment from people like you."[20] Similarly, in opinion pieces, they asserted that crime would surge uncontrollably (Kindy, 2021). In the context of actually rising crime rates after 2020, this strategy reduced lawmakers' willingness to pursue reforms that could portray them as soft on crime. In cases where police groups could not prevent immunity reforms completely, for example in New Mexico, they often managed to shift the narrative and ensured that victims could only seek retribution from cities and counties, rather than individual officers (Kindy, 2021). Hence, by leveraging citizens' fear of crime and credibly threatening a change in the quality of law

[20] https://bostonglobe.newspapers.com/browse/the-boston-globe_9077/2020/08/17/

enforcement, police made reforms of "qualified immunity" electoral risky and unattractive for incumbents.

6.2 Examples of Over-Reforming

Conversely, our model also explains how and when incumbents can *leverage* the possibility of sabotage for their electoral gains. If reforms are relatively popular with voters, incumbents can over-invest in policies that are doomed to fail while blaming bureaucratic sabotage for such failure.

A prominent example of this is the strategy of populist incumbents, most prominently Donald Trump, to blame the "deep state" for policy failures, i.e., claiming that bureaucrats are actively undercutting the president's constitutional authority and thwarting the will of the people by sabotaging Trump's policies. In 2017, almost half of the American public (48%) believed that a "deep state" exists, described as "military, intelligence, and government officials who try to secretly manipulate government policy." Only 35% called it a conspiracy theory.[21] At the start of his presidency, Trump inherited a bureaucracy that was both sparse and aging since the federal workforce as a percentage of the total American population had shrunk since the 1970s (Partnership for Public Service, 2019). Similarly, the compensation system for federal employees has not been reformed since 1949, and federal workers complain about rigid job classifications and excessive outsourcing of government work (Verkuil, 2017; Medina, 2021).

Instead of bolstering the bureaucracy, Trump leveraged "deep state" rhetoric to justify policy failures. For instance, facing problems in confronting a surge of migrants at the southwestern border, Trump claimed that his desire to enforce tougher screening of asylum seekers was purposefully obstructed and delayed by bureaucrats at the Department of Homeland Security (Shear and Kanno-Youngs, 2019). Similarly, in light of difficulties in addressing the COVID-19 pandemic and in an effort to hastily provide remedies against scientific advice, Trump leveled several attacks on the US Food and Drug Administration and cast scientifically dubious treatments as "breakthroughs." (Facher, 2020) He asserted that the agency was strategically delaying vaccines and treatments for the virus in order to undermine his 2020 election efforts, tweeting that "The deep state, or whoever, over at the FDA is making it very difficult for drug companies to get people in order to test the vaccines and therapeutics. Obviously, they are hoping to delay the answer until after November 3rd. Must focus on speed, and saving lives!" (Yen and Woodward, 2020). Hence, by claiming that the federal bureaucracy was working to undermine his administration, Trump weaponized fears

[21] https://abcnews.go.com/Politics/lies-damn-lies-deep-state-plenty-americans-poll/story?id=47032061

of a “deep state” bureaucracy among his supporters to legitimize drastic policies that did not succeed.

7 Conclusion & Discussion

Politicians inherently depend on bureaucrats to deliver policies to their voter base, and poor public service provision creates an electoral vulnerability for politicians. This raises the question: When and how can bureaucrats exploit this to affect policies they dislike? In this paper, we argue that bureaucrats' central position in government production, together with voters' difficulty in attributing responsibility for service provision, vests bureaucrats with a unique source of political power. To the best of our knowledge, we provide the first analysis into how this leads to strategic sabotage of public service provision by bureaucrats, affects voter learning from policy outcomes, and can impact politicians' policies and chances of re-election.

Using a three-player model with a politician, a bureaucrat, and a voter, we find that bureaucratic sabotage leads to complex and non-monotonic disruptions in electoral accountability relationships between voters and politicians. Depending on the voter's beliefs about the merit of reform policies and the observed quality of government, bureaucratic sabotage (1) can make the voter either more or less favorable to the incumbent, (2) happens more often if voters are more susceptible to government outcome, and (3) can lead to either under-reform or over-reform relative to the normative optimum.

Our model and analysis enrich our understanding of the degree of political motivations among bureaucrats and their consequences for voter learning and politicians' behavior. In doing so, we highlight an underappreciated mechanism of political influence for bureaucrats as interest groups and micro-found a reason for why bureaucrats act against the very programs and services they oversee. Additionally, we respond to recent calls to integrate interactions between politicians, bureaucrats, and voters within a single framework for studying political accountability (Grossman and Slough, 2022). Compared to conventional models of electoral accountability that examine the relationships between voters and politicians or between politicians and bureaucrats separately, this integration allows us to uncover new mechanisms influencing voter learning, service quality, and government responsiveness. It also demonstrates a novel pathway through which the information environment shapes the outcomes of accountability relationships (Ashworth, Bueno De Mesquita and Friedenberg, 2018).

This article opens several paths for future work. In our model, we focus on a simple two-period game and abstract away from potential dynamics. Particularly, we treat both the voter's perceptions about the reform's value relative to the status quo (q) and bureaucrats' perceived costs of sabotage (c) as exogenous. It appears fruitful for future theoretical research to explore how our results are affected by voters' dynamic adjustment of their beliefs about the cost of sabotage or the reform's value over time.

Our model can also inform future empirical work on the drivers, conditions, and consequences of bureaucratic sabotage in several ways. In particular, one could test the comparative statics described in Propositions 2, 3, and 4, i.e., the effect of changes in voter's beliefs about the reform's value (q) and bureaucrats' cost-benefit trade-off when sabotaging (c relative to κ) on the probability of reform ($1 - r^*$), the probability of reelection ($1 - g^*$), and the probability of sabotage ($1 - \kappa^*$). Similarly, scholars could empirically evaluate the impact of bureaucratic sabotage (i.e., variation in c) on voters' perceptions of reform merit ($\mathrm{E}[x|g, c]$), conditional on the realized government quality (g). Our results suggest that we should observe opposite effects of the propensity to sabotage on voters' preferences about reform (see Lemma A1.3). When selecting empirical cases for such analyses, scholars want to pay close attention to three issues. First, the cases should closely match the scope conditions of our theory—particularly, bureaucrats' distaste for reform, their discretion and independence from political control, and voters' difficulty in attributing the responsibility for the government outcome. The second and thornier issue concerns the source of the exogenous variation in either c or q for ceteris paribus comparisons. Particularly, it proves empirically challenging to identify valid instruments that affect one of these exogenous parameters while leaving the other unchanged. Take, for example, the case of police resistance to law enforcement reforms. Assume that a scholar sets out to study how sudden shifts in voters' attitudes toward the necessity for police reform (q) affect the degree of police sabotage, incumbents' policies, and their re-election chances. Instances of police brutality followed by widespread protests might seem like ideal shocks. However, it's important to recognize that such events have a direct impact on how police officers weigh the costs and benefits of engaging in sabotage. For example, a broader shift in the political climate following large-scale protests tends to increase police officers' concerns about potential consequences for their actions, strengthening their resistance to measures like the removal of qualified immunity (i.e., reducing c relative to κ). Hence, it is difficult to test model predictions with this design. However, other instruments, such as localized unionization of individual bureaucratic units through unionization elections (Goncalves, 2021), could be promising candidates to empirically study the effect of rapid changes in the cost of organized sabotage on its prevalence and consequences.

References

ABCNews. 2021. "Work slowdown blamed as trash piles up on Staten Island.". https://abc7ny.com/staten-island-trash-garbage-work-slowdown/11173917/.

Andreottola, Giovanni. 2021. "Flip-Flopping and Electoral Concerns." *The Journal of Politics* 83(4):1669–1680.
URL: *https://doi.org/10.1086/715070*

Anzia, Sarah F. 2014. *Timing and Turnout: How Off-Cycle Elections Favor Organized Groups*. Chicago: The University of Chicago Press.

Anzia, Sarah F. 2022. *Local Interests: Politics, Policy, and Interest Group in US City Governments*. Chicago: The University of Chicago Press.

Anzia, Sarah F. and Terry M. Moe. 2015. "Public Sector Unions and the Costs of Government." *The Journal of Politics* 77(1):114–127.
URL: *https://doi.org/10.1086/678311*

Ashworth, By Scott, Ethan Bueno De Mesquita and Amanda Friedenberg. 2017. "Accountability and Information in Elections." *American Economic Journal : Microeconomics* 9(2):95–138.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15:183–201.

Ashworth, Scott and Ethan Bueno De Mesquita. 2006. "Monotone Comparative Statics for Models of Politics." *American Journal of Political Science* 50(1):214–231.

Ashworth, Scott, Ethan Bueno De Mesquita and Amanda Friedenberg. 2018. "Learning about Voter Rationality." *American Journal of Political Science* 62(1):37–54.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera and Modibo Sidibé. 2021. "Police Officer Assignment and Neighborhood Crime." NBER Working Paper 29243. https://www.nber.org/system/files/working_papers/w29243/w29243.pdf.

Bagnoli, Mark and Ted Bergstrom. 2006. Log-concave probability and its applications. In *Rationality and Equilibrium*, ed. Charalambos D. Aliprantis, Rosa L. Matzkin, Daniel L. McFadden, James C. Moore and Nicholas C. Yannelis. Berlin, Heidelberg: Springer Berlin Heidelberg pp. 217–241.

Ban, Pamela, Ju Yeon Park and Hye Young You. 2024. "Bureaucrats in Congress:Strategic Information Sharing in Policymaking." Working Paper. https://hyeyoungyou.files.wordpress.com/2024/02/bureaucrats_in_congress.pdf.

Banks, Jeffrey S. and Barry R. Weingast. 1992. "The Political Control of Bureaucracies under Asymmetric Information." *American Journal of Political Science* 36(2):509–524.

Besley, Timothy and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review* 95(3):616–636.
URL: *https://www.aeaweb.org/articles?id=10.1257/0002828054201413*

Bils, Peter and Federica Izzo. 2023. "Policymaking in Times of Crisis." Working Paper. http://peterbils.com/wp-content/uploads/2023/02/Policy_making_in_times_of_crisis-24.pdf.
URL: *www.newyorker.com/magazine/2020/11/02/barack-obama-new-book-excerpt-promised-land-obamacare.*

Blumgart, Jake. 2020. "'They rely on you being intimidated': Local elected officials in the US describe how police unions bully them." CityMonitor, August 18 2020. https://citymonitor.ai/government/they-rely-you-being-intimidated-local-elected-officials-us-describe-how-police-unions-bully.

Brehm, John and Scott Gates. 1997. *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor: University of Michigan Press.

Bueno de Mesquita, Ethan and Scott A. Tyson. 2020. "The Commensurability Problem: Conceptual Difficulties in Estimating the Effect of Behavior on Behavior." *American Political Science Review* 114(1):385–391.

Canes-Wrone, Brandice, Michael C. Herron and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45(3):532–550.

DiSalvo, Daniel. 2015. *Government Against Itself: Public Union Power and Its Consequences*. New York: Oxford University Press.

Downs, Anthony. 1967. *Inside Bureaucracy*. Boston: Little, Brown.

Epstein, David and Sharyn O'Halloran. 1999. *Delegating Powers*. New York: Cambridge University Press.

Facher, Lev. 2020. "Trump has launched an all-out attack on the FDA. Will its scientific integrity survive?" STAT News, Aug. 27, 2020. https://www.statnews.com/2020/08/27/trump-has-launched-an-all-out-attack-on-the-fda-will-its-scientific-integrity-survive/.

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Adam Przeworski, Susan Carol Stokes and Bernard Manin. New York: Cambridge University Press pp. 55–97.

Flavin, Patrick and Michael T. Hartney. 2015. "When Government Subsidizes Its Own: Collective Bargaining Laws as Agents of Political Mobilization." *American Journal of Political Science* 59(4):896–911.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12163*

Foarta, Dana. 2023. "How Organizational Capacity Can Improve Electoral Accountability." *American Journal of Political Science* 67(3):776–789.

Fong, Christian and Keith Krehbiel. 2018. "Limited Obstruction." *American Political Science Review* 112(1):1–14.

Forand, Jean Guillaume, Gergely Ujhelyi and Michael M Ting. 2022. "Bureaucrats and Policies in Equilibrium Administrations." *Journal of the European Economic Association* 21(3):815–863.

Gersen, Jacob E. and Matthew C. Stephenson. 2014. "Over-accountability." *Journal of Legal Analysis* 6(2):185–243.

Gieczewski, Germán and Christopher Li. 2021. "Dynamic Policy Sabotage." *American Journal of Political Science* .

Goncalves, Felipe. 2021. "Do Police Unions Increase Misconduct?" Working Paper. https://static1.squarespace.com/static/58d9a8d71e5b6c72dc2a90f1/t/60622724b6a902732b636324/1617045285669/Goncalves_Unions.pdf.

Gross, Courtney. 2021. "Sanitation commissioner details trash delays, says some trucks returning half full." Spectrum News NY1. https://ny1.com/nyc/all-boroughs/news/2021/10/26/sanitation-slowdown--commissioner-says--miscommunication-.

Grossman, Guy and Tara Slough. 2022. "Government Responsiveness in Developing Countries." *Annual Review of Political Science* 25(1):131–153.
URL: *https://doi.org/10.1146/annurev-polisci-051120-112501*

Hartney, Michael and Patrick Flavin. 2011. "From the Schoolhouse to the Statehouse: Teacher Union Political Activism and U.S. State Education Reform Policy." *State Politics & Policy Quarterly* 11(3):251–268.
URL: *http://www.jstor.org/stable/41575825*

Hartney, Michael T. 2022. "Teachers' unions and school board elections: a reassessment." *Interest Groups & Advocacy* 11:237–262.

Heo, Kun and Dimitri Landa. 2024. "A Theory of Policy Justification." Working Paper.

Hirsch, Alexander V. and Jonathan P. Kastellec. 2022. "A theory of policy sabotage." *Journal of Theoretical Politics* 34(2):191–218.
URL: *https://doi.org/10.1177/09516298221085974*

Hoekstra, Mark and CarlyWill Sloan. 2022. "Does Race Matter for Police Use of Force? Evidence from 911 Calls." *American Economic Review* 112(3):827–60.

Huber, John D. and Charles R. Shipan. 2002. *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy*. New York: Cambridge University Press.

Huber, John D. and Nolan Mccarty. 2004. "Bureaucratic Capacity, Delegation, and Political Reform." *American Political Science Review* 98(3):481–494.

Kindy, Kimberly. 2021. "Dozens of states have tried to end qualified immunity. Police officers and unions helped beat nearly every bill." Washington Post, October 7, 2021. https://www.washingtonpost.com/politics/qualified-immunity-police-lobbying-state-legislatures/2021/10/06/60e546bc-0cdf-11ec-aea1-42a8138f132a_story.html.
URL: *https://www.washingtonpost.com/politics/qualified-immunity-police-lobbying-state-legislatures/2021/10/06/60e546bc-0cdf-11ec-aea1-42a8138f132a_story.html*

Knight, Heather. 2021. "More S.F. residents share stories of police standing idly by as crimes unfold: 'They didn't want to be bothered'." San Francisco Chronicle. https://www.sfchronicle.com/sf/bayarea/heatherknight/article/sf-police-crime-16931399.php.

Kucinskas, Jaime and Yvonne Zylan. 2023. "Walking the Moral Tightrope: Federal Civil Servants' Loyalties, Caution, and Resistance under the Trump Administration." *American Journal of Sociology* 128(6):1761–1808.

Kyriazis, Anna, Lauren Schechter and Dvir Yogev. 2023. "The Day After the Recall: Policing and Prosecution in San Francisco." Working Paper. https://github.com/lrschechter

/workingpapers/blob/c3607e05afb130ed4de5291954f44a6105186bf5/KyriazisSchechterYogev.pdf.

Leighley, Jan E. and Jonathan Nagler. 2007. "Unions, Voter Turnout, and Class Bias in the U.S. Electorate, 1964–2004." *The Journal of Politics* 69(2):430–441.
URL: *http://www.jstor.org/stable/10.1111/j.1468-2508.2007.00541.x*

Lipsky, Michael. 1980. *Street Level Bureaucracy: Dilemmas of the Individual In Public Services*. Russell Sage Foundation.

McCaskill, Nolan D. 2020. "Police reforms stall around the country, despite new wave of activism." Politico, September 23, 2020. https://www.politico.com/news/2020/09/23/breonna-taylor-police-reforms-420799#:~:text=Activists%20tracking%20bills%20in%20state,already%20adjourned%20for%20the%20year.
URL: *https://www.politico.com/news/2020/09/23/breonna-taylor-police-reforms-420799*

McCubbins, Mathew D. 1985. "The Legislative Design of Regulatory Structure." *American Journal of Political Science* 29(4):721–748.

McCubbins, Mathew D., Roger G. Noll and Barry R. Weingast. 1987. "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics, & Organization* 3(2):243–277.

Medina, Eduardo. 2021. "Across the U.S., Clashes Intensify Between City Officials and the Police Over Vaccination Issues." New York Times, Oct. 17, 2021. https://www.nytimes.com/2021/10/17/us/police-vaccine-mandate.html.

Miller, Gary J. and Terry M. Moe. 1983. "Bureaucrats, Legislators, and the Size of Government." *The American Political Science Review* 77(2):297–322.

Moe, Terry M. 2006. "Political Control and the Power of the Agent." *The Journal of Law, Economics, and Organization* 22(1):1–29.
URL: *https://doi.org/10.1093/jleo/ewj011*

Moe, Terry M. 2009. "Collective Bargaining and the Performance of the Public Schools." *American Journal of Political Science* 53(1):156–174.

Moe, Terry M. 2011. *Special Interest: Teachers Unions and America's Public Schools*. Washington D.C.: Brookings Institution Press.

Moe, Terry M. 2012. "Delegation, Control, and the Study of Public Bureaucracy." *The Forum* 10(2). Article 4.
URL: *https://doi.org/10.1515/1540-8884.1508*

Niskanen, William A. 1971. *Bureaucracy and Representative Government*. New York: Aldine-Atherton.

Noll, David L. 2022. "Administrative Sabotage." *Michigan Law Review* 120(5):753–824.

O'Leary, Rosemary. 2020. *The Ethics of Dissent: Managing Guerrilla Government*. Third ed. Washington D.C.: CQ Press.

Paglayan, Agustina S. 2019. "Public-Sector Unions and the Size of Government." *American Journal of Political Science* 63(1):21–36.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12388*

Partnership for Public Service. 2019. "Federal Workforce." Fed Figures 2019. https://ourpublicservice.org/wp-content/uploads/2022/03/FedFigures_FY18-Workforce-1.pdf.
URL: *https://ourpublicservice.org/wp-content/uploads/2022/03/FedFigures_FY18-Workforce-1.pdf*

Patty, John W. 2016. "Signaling through Obstruction." *American Journal of Political Science* 60(1):175–189.

Patty, John W. and Elizabeth Maggie Penn. 2023. "Algorithmic Fairness and Statistical Discrimination." *Philosophy Compass* 18(1):e12891.
URL: *https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/phc3.12891*

Pearson, Jake. 2022. "More Than Two Years After George Floyd's Murder Sparked a Movement, Police Reform Has Stalled. What Happened?" ProPublica, October 24, 2022. https://www.propublica.org/article/why-police-reform-stalled-elizabeth-glazer.
URL: *https://www.propublica.org/article/why-police-reform-stalled-elizabeth-glazer*

Saumard, Adrien and Jon A. Wellner. 2014. "Log-concavity and strong log-concavity: a review." *Statistics Surveys* 8:45–114.

Shear, Michael D. and Zolan Kanno-Youngs. 2019. "Trump Administration to Push for Tougher Asylum Rules." New York Times, April 9, 2019. https://www.nytimes.com/2019/04/09/us/politics/asylum-seekers-trump-administration.html.

URL: *https://www.nytimes.com/2019/04/09/us/politics/asylum-seekers-trump-administration.html?action=click%26module=Top%20Stories%26pgtype=Homepage*

Slough, Tara. 2024. "Bureaucratic Quality and Electoral Accountability." *American Political Science Review* pp. 1–20.

Swan, Rachel. 2021. "San Francisco police just watch as burglary appears to unfold, suspects drive away, surveillance video shows." San Francisco Chronicle. https://www.sfchronicle.com/bayarea/article/San-Francisco-police-only-watch-as-burglary-16647876.php.

Ting, Michael M. 2021. "The Political Economy of Governance Quality." *American Political Science Review* 115(2):667–685.

Tullock, Gordon. 1965. *The Politics of Bureaucracy*. Washington, D.C.: Public Affairs Press.

Ujhelyi, Gergely. 2014. "Civil service reform." *Journal of Public Economics* 118:15–25.
URL: *https://www.sciencedirect.com/science/article/pii/S0047272714001492*

Verkuil, Paul R. 2017. *Valuing Bureaucracy: The Case for Professional Government*. Cambridge University Press.

Wilson, James Q. 1978. *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, Massachusetts: Harvard University Press.

Wirsching, Elisa M. 2023. "Political Power of Bureaucratic Agents: Evidence from Policing in New York City." Working Paper. https://elisawirsching.github.io/research/policeresistance.pdf.

Yazaki, Yukihiro. 2018. "The effects of bureaucracy on political accountability and electoral selection." *European Journal of Political Economy* 51:57–68.
URL: *https://www.sciencedirect.com/science/article/pii/S0176268016300908*

Yen, Hope and Calvin Woodward. 2020. "AP FACT CHECK: Trump's baseless claim of 'deep state' at FDA." AP News, Aug. 24, 2020. https://apnews.com/article/election-2020-politics-religion-virus-outbreak-iran-nuclear-d976ab2b204ac14b220d5bf6dc565b2d.

Zoorob, Michael. 2019. "Blue Endorsements Matter: How the Fraternal Order of Police Contributed to Donald Trump's Victory." *PS: Political Science & Politics* 52(2):243–250.

Appendix: Supporting Information for *Bureaucratic Sabotage and Policy Inefficiency*

A Proofs

A.1 Marginal Effects on The Voter Inference

Proof for Lemma 1. Lemma 1-1.

$$sign \frac{\partial}{\partial g^*} \hat{L}(g, r', \kappa') = sign \frac{\partial}{\partial g^*} \frac{h(g)}{h(g) + \kappa' \Big(h(g-1) - h(g) \Big)}$$

A log-concave distribution satisfies the monotone likelihood ratio property with respect to horizontal shift (Saumard and Wellner, 2014), so $\frac{\partial}{\partial g^*} \frac{h(g)}{h(g)+\kappa'\big(h(g-1)-h(g)\big)} > 0$.

Since $h(g-1) = h(g)$ if and only if $g = 1/2$, $\hat{L} = 0$ if and only if $g = 1/2$. ■

Lemma 1-2.

$$\begin{aligned} sign \frac{\partial}{\partial r'} \hat{L}(g, r', \kappa') &= sign \frac{\partial}{\partial r'} \frac{1 - F(r'|0)}{1 - F(r'|1)} \\ &= sign \Bigg(- \Big([1 - F(r'|1)] + \Big) f(r'|0) + \Big([1 - F(r'|0)] + \Big) f(r'|1) \Bigg) \end{aligned}$$

Observe that this is not positive if and only if

$$\frac{f(r'|0)}{1 - F(r'|0)} \geq \frac{f(r'|1)}{1 - F(r'|1)}. \tag{6}$$

Consider an arbitrary r and $\tilde{r} > r$. By the monotone likelihood ratio property,

$$\frac{f(\tilde{r}|1)}{f(\tilde{r}|0)} \geq \frac{f(r|0)}{f(r|0)} \iff f(\tilde{r}|1) f(r|0) \geq f(r|1) f(\tilde{r}|0).$$

Then

$$\begin{aligned} & f(r|0) \int_r^1 dF(\tilde{r}|1) \geq f(r|1) \int_r^1 dF(\tilde{r}|0) \\ \iff & [1 - F(r|1)] f(r|0) \geq [1 - F(r|0)] f(r|1), \end{aligned}$$

which implies equation (6). ■

Lemma 1-3. Observe that

$$sign\frac{\partial}{\partial\kappa'}\hat{L}(g,r',\kappa') = sign\frac{\partial}{\partial\kappa'}\frac{h(g)}{h(g)+\kappa'\Big(h(g-1)-h(g)\Big)}$$
$$= sign[h(g)-h(g-1)].$$

Since $h(g)$ is symmetric around 0 and single-peaked, $h'(g) < 0$ if $g > 0$. Notice that this implies that $h(g-1)-h(g) = 0$ if $g = 1/2$ and > 0 if $g < 1/2$. ■ ■

Remark 1 *The probability of sabotage, $1-\hat{\kappa}^*(\hat{g}^*)$ given an arbitrary cutoff $\hat{g}^*$ is single-peaked with respect to $\hat{g}^*$ and has its maximum at $\hat{g}^* = 1/2$.*

Proof for remark 1. It is straightforward that $h(g) > h(g-1)$ if and only if $g < 1/2$. ■

A.2 Endogenizing Sabotage and Inference

Proof for Proposition 1. From equation (2), we can endogenize bureaucrats' action by finding their cutoff κ^* as a function of the voter's cutoff $\hat{g}^*$: $\kappa^* = \kappa^*(\hat{g}^*)$. Plug $\kappa' = \kappa^*$ in L to get

$$L(g,r') \equiv \hat{L}(g,r',\kappa^*) = \frac{1-F(r'|0)}{1-F(r'|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}. \tag{7}$$

Endogenizing κ^* preserves the qualitative results from Lemma 1. Observe the following.

Lemma A1 *For $L(g,r') \equiv \hat{L}(g,r',\kappa^*)$*

1. *$L(g,r')$ is decreasing in g.*
2. *$L(g,r')$ is decreasing in r'.*
3. *$L(g,r')$ is increasing in c if and only if $g < 1/2$.*

Most notably, L retains the monotonic properties of $\hat{L}$ with respect to g and r'. Therefore, there exists a unique $\hat{g}^*(q,r',c)$ such that $L(\hat{g}^*(q,r',c),r') = \log\frac{1-q}{q} \iff \mathrm{E}[x|g;r',c] = q$.

Corollary 1 *$\hat{g}^*(q,r',c)$ is increasing in q, decreasing in r'. It is increasing in c if and only if it is smaller than $1/2$.*

Recall that g^* is increasing in the probability of sabotage, $1-\kappa'$, if and only if it is larger than $1/2$. As c increases, sabotage becomes less likely, so g^* is increasing in c if and only if it is less than $1/2$.

Proof for Lemma A1. It is straightforward that $sign\frac{\partial L}{\partial r'} = sign\frac{\partial \hat{L}}{\partial r'}$.

To see $sign\frac{\partial L}{\partial g} = sign\frac{\partial \hat{L}}{\partial g}$, notice that

$$\frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} = \frac{1}{1 + c\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}}$$

is decreasing in g since

$$\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}$$

is increasing in g. Observe the following: $sign\frac{\partial}{\partial g^*}\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}$ is the same as

$$sign\Bigg(\Big(H(g)-H(g-1)\Big)\frac{h'(g-1)h(g)-h'(g)h(g-1)}{\Big(h(g)\Big)^2} - \Big(h(g)-h(g-1)\Big)\Big(\frac{h(g-1)}{h(g)}-1\Big)\Bigg).$$

First, $H(g) > H(g-1)$. Then, log-concavity of h ensures $h'(g-1)h(g) - h'(g)h(g-1) > 0 \iff h'(g-1)h(g-1) > \frac{h'(g)}{h(g)}$. Notice this this holds if $\frac{\partial}{\partial g^*}\frac{h'(g)}{h(g)} < 0 \iff h''(g)h(g) < \Big(h'(g)\Big)^2$, which is a property of a log-concave function (Bagnoli and Bergstrom, 2006). Lastly, $\Big(h(g)-h(g-1)\Big)\Big(\frac{h(g-1)}{h(g)}-1\Big) = -h(g) - \frac{\big(h(g-1)\big)^2}{h(g)} < 0$. Therefore, $\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}$ is increasing in g.

$L(g, r')$ depends on c as well. $L(g, r')$ is increasing in c if and only if

$$\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)} < 0 \iff h(g-1) < h(g) \iff g < 1/2.$$

■

Proof for Corollary 1. As q increases, $\log\frac{1-q}{q}$ decreases. Since $L(g^*, r')$ is monotonically decreasing in g^*, g^* is increasing in q.

Since L is monotonic with respect to g^*, if L is increasing/decreasing in a parameter, g^* is increasing/decreasing as well. See (Ashworth and Bueno De Mesquita, 2006). ■

Lemma A2 *$L(g, \hat{r}^*(g))$ is decreasing in g.*

Proof for Lemma A2. $\frac{\partial L(g,\hat{r}^*(g))}{\partial g} = \frac{\partial}{\partial g}\Big(\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}\Big)$ and

$$\begin{aligned}&\frac{\partial}{\partial g}\Big(\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}\Big)\\ &= \frac{\partial}{\partial g}\Big(\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\Big)\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} + \frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\frac{\partial}{\partial g}\Big(\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}\Big).\end{aligned}$$

Notice

$$\frac{\partial}{\partial g}\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} = \frac{\partial r^*}{\partial g}\frac{\partial}{\partial r^*}\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} \leq 0$$

since $\frac{\partial r^*}{\partial g} \geq 0$ and $\frac{\partial}{\partial r^*}\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} \leq 0$. Recall

$$\frac{\partial}{\partial g}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \leq 0.$$

Thus, $\frac{\partial L(g,\hat{r}^*(g))}{\partial g} \leq 0$. ■ By the property of the monotone real mapping (the intermediate value theorem), there exists unique $g^*(q)$ such that $\partial L(g^*(q), \hat{r}^*(g^*(q))) = \frac{1-q}{q}$. ■

A.3 Comparative Statics

A.3.1 The effects of the status quo's value

$g^*(q)$ that satisfies equation $L(g,\hat{r}^*(g)) = \frac{1-q}{q}$ is increasing in q. Recall that $\hat{r}^*(g)$ is increasing in g.

Recall that κ^* is a U-shaped function of g^* that attains its minimum at $g^* = 1/2$ and g^* is increasing in q.

A.3.2 Equilibrium Characterization without Sabotage

To understand how sabotage affects policy-making, we first evaluate an equilibrium benchmark without sabotage, i.e., $\kappa' = 1$. Then $\rho(r')\Big(H(\hat{g}^*) - H(\hat{g}^*-1)\Big) = H(\hat{g}^*) - 1/2$, so $\rho(r') = \frac{H(\hat{g}^*)-1/2}{H(\hat{g}^*)-H(\hat{g}^*-1)}$. Define

$$\hat{r}^*_B(g) \equiv \rho^{-1}\Big(\frac{H(g)-1/2}{H(g)-H(g-1)}\Big) \quad \text{and} \quad L_B(r',g) \equiv \frac{1-F(r'|0)}{1-F(r'|1)}\frac{h(g)}{h(g-1)}.$$

The following proposition describes the unique equilibrium under no sabotage.

Proposition A1 *In the benchmark case, without sabotage, There exists a unique pure strategy equilibrium defined by* $\Big(g^*_B(q), r^*_B(q)\Big)$ *such that*

$$L_B(g^*_B(q), \hat{r}^*_B(g^*_B(q))) = \frac{1-q}{q}, \qquad r^*_B(q) = \hat{r}^*_B(g^*_B(q)).$$

In equilibrium, there is a minimum government service quality, $g^*_B(q)$, such that the voter reelects a reforming incumbent if and only if he observes at least this minimum quality, $g \geq g^*_B(q)$. In turn, there is also a minimum signal about the reform's value, $r^*_B(q)$, such that the incumbent introduces the reform if and only if $r \geq r^*_B(q)$.

Proof for Proposition A1. First, observe that

$$\frac{\partial}{\partial g}\frac{H(g)-1/2}{H(g)-H(g-1)} = \frac{[2H(g)-1]h(g-1)-[2H(g-1)-1]h(g)}{2[H(g)-H(g-1)]^2}.$$

Notice that $H(0) = 1/2$, so, for $g \in [0,1]$, $H(g) \geq 1/2$ and $H(g-1) \leq 1/2$. Thus, $\frac{\partial}{\partial g}\frac{H(g)-1/2}{H(g)-H(g-1)} < 0$ for $g \in [0,1]$. Since ρ is increasing, $\hat{r}^*_B(g)$ is increasing in g.

Then $sign\frac{\partial}{\partial g}L_B(g^*_B, \hat{r}^*_B(g^*_B(q))) = sign\frac{\partial}{\partial g}\log L_B(g^*_B, \hat{r}^*_B(g^*_B(q)))$, and

$$\frac{\partial}{\partial g}\log L_B(g^*_B, \hat{r}^*_B(g^*_B(q))) = \underbrace{\frac{\partial}{\partial g}\log\frac{1-F(\hat{r}^*_B(g)|0)}{1-F(\hat{r}^*_B(g)|1)}}_{<0 \text{ since } \frac{\partial \hat{r}^*_B(g)}{\partial g}>0 \text{ \& } \frac{\partial}{\partial r'}\frac{1-F(r'|0)}{1-F(r'|1)}<0} + \underbrace{\frac{\partial}{\partial g}\log\frac{h(g)}{h(g-1)}}_{<0 \text{ by MLRP}} < 0.$$

Notice that $\frac{H(g)-1/2}{H(g)-H(g-1)} > 1$ if $g > 1$ and $\frac{H(g)-1/2}{H(g)-H(g-1)} < 0$ if $g < 0$. Recall that $\rho(r) \in [0,1]$, so $g^*_B \in [0,1]$.

Thus, $L_B(g, \hat{r}^*_B(g))$ is decreasing in g, and by the intermediate value theorem, there exists a unique $g^*_B(q)$ such that $L_B(g^*_B(q), \hat{r}^*_B(g^*_B(q))) = \frac{1-q}{q}$. ■

As discussed, the voter requires a higher service quality as the status quo's value increases, and as a result, the incumbent also requires a stronger signal.

Proposition A2 *In the benchmark case without sabotage, the value of the status quo q has the following effects:*

1. *The voter applies a more stringent criterion for reelecting a reforming incumbent as the status quo's value increases; $g^*_B(q)$ is increasing in q.*
2. *As the status quo's value increases, the incumbent requires a higher signal to introduce reform; $r^*_B(q)$ is weakly increasing in q.*

Proof for Proposition A2. Since $\frac{1-q}{q}$ is decreasing and $L_B(g, \hat{r}^*_B(g))$ is decreasing in g, $g^*_B(q)$ such that $L_B(g^*_B(q), \hat{r}^*_B(g^*_B(q))) = \frac{1-q}{q}$ is increasing in q.[1]

It is straightforward that $\hat{r}^*_B(g))$ is increasing in g. ■

The next statement depicts the equilibrium in terms of reform efficiency.

Proposition A3 *There exists a unique $q^\dagger_{BE} \in (1/2, 1)$ such that $\rho(r^*_B(q^\dagger_{BE})) = q^\dagger_{BE}$ and $\rho(r^*_B(q)) < q$ if $q \in (0, q^\dagger_{BE})$ and $\rho(r^*_B(q)) > q$ if $q \in (q^\dagger_{BE}, 1)$.*

[1] See Ashworth and Bueno De Mesquita (2006).

Proof for Proposition A3. Observe that

$$\frac{\partial^2}{\partial g^2}\frac{H(g)-1/2}{H(g)-H(g-1)} = [H(g)-(1/2)]\frac{h'(g-1)[H(g)-H(g-1)]-[h(g)-h(g-1)]h(g-1)}{[H(g)-H(g-1)]^3}$$
$$+[1/2-H(g-1)]\frac{h'(g)[H(g)-H(g-1)]-[h(g)-h(g-1)]h(g)}{[H(g)-H(g-1)]^3},$$

which is 0 at $g=1/2$ by the symmetry at the point. For $g\in[0,1/2]$, $h'(g-1)\geq 0\geq -h'(g)$, $H(g)-1/2\geq 1/2-H(g-1)$, and $h(g)\geq h(g-1)$. Therefore, $\frac{\partial^2}{\partial g^2}\frac{H(g)-1/2}{H(g)-H(g-1)}\leq 0$. Thus, $\rho(\hat{r}^*(g))$ is increasing in g concave at 0. By the opposite logic, it is convex at $g\geq 1/2$. Thus, $\frac{H(g)-1/2}{H(g)-H(g-1)}$ is a reverse S-shaped function on $g\in[0,1]$.

Since $\frac{H(g)-1/2}{H(g)-H(g-1)}$ is monotonically increasing in g, there exists a unique $g^*_{NB}(q)$ such that

$$\frac{H(g^*_{NB}(q))-1/2}{H(g^*_{NB}(q))-H(g^*_{NB}(q)-1)}=q$$

holds for any q. Notice that $g^*_{NB}(q)$ is an increasing function of q. Notice that this is if q satisfies the equation above, then the incumbent chooses the reform if and only if $\rho(r)\geq q$.

On the other hand, notice that there is a unique $q_{NB}(g)$ such that

$$\frac{h(g)}{h(g-1)}=\frac{1-q_{NB}(g)}{q_{NB}(g)}\frac{1-F(\rho^{-1}(q_{NB}(g))|1)}{1-F(\rho^{-1}(q_{NB}(g))|0)}.$$

Observe

$$\frac{\partial}{\partial q}\frac{1-q}{q}\frac{1-F(\rho^{-1}(q)|1)}{1-F(\rho^{-1}(q)|0)}$$
$$=-\underbrace{\frac{(1-q)[1-F(\rho^{-1}(q)|1)]\frac{\partial}{\partial q}F(\rho^{-1}(q)|0)}{q[1-F(\rho^{-1}(q)|0)]^2}}_{\geq 0}-\underbrace{\frac{(1-q)q\frac{\partial}{\partial q}F(\rho^{-1}(q)|1)+1-F(\rho^{-1}(q)|1)}{q^2[1-F(\rho^{-1}(q)|0)]}}_{\geq 0}\leq 0$$

Notice that F and ρ^{-1} are increasing, so $\frac{\partial}{\partial q}F(\rho^{-1}(q)|x)\geq 0$. Since $\frac{h(g)}{h(g-1)}$ is decreasing in g by the MLRP, $q_{NB}(g)$ is monotonically increasing in g. Suppose $g\in[0,1]$.

Then, $g^*_{NB}(q)$ has at least one fixed point by Brouwer's Fixed point theorem (it is are continuous mapping from a closed interval to itself). Furthermore, since $\frac{H(g)-1/2}{H(g)-H(g-1)}$ is reversed S-shaped in $[0,1]$, its inverse $g^*_{NB}(q)$ is S-shaped function of q. By the property of the S-shaped functions, $g^*_{NB}(q)$ has at most a unique interior fixed point $q^\dagger_{BE}\in(0,1)$ such that $g^*_{NB}(q^\dagger_{BE})=q^\dagger_{BE}$ and $g^*_{NB}(q)>q$ if $q\in(q^\dagger_{BE},1)$ and $g^*_{NB}(q)<q$ if $q\in(0,q^\dagger_{BE})$.

To see this, first, we have to check the corners $q_{NB}=0$ and $q_{NB}=1$. Let $q_{NB}=0$. Then $g^*_{NB}(q)=0$ but $\frac{h(0)}{h(-1)}<\lim_{q_{NB}\to 0}\frac{1-q_{NB}}{q_{NB}}\frac{1-F(\rho^{-1}(q_{NB})|1)}{1-F(\rho^{-1}(q_{NB})|0)}=\infty$, so $q_{NB}=0$ is not a fixed point of $g^*_{NB}(g)$. Since $h(g)/h(g-1)$ is monotonic, $g^*_{NB}(0)=g^*(0)=-\infty$, and again, by the monotonicity of $\frac{H(g)-1/2}{H(g)-H(g-1)}$, $\lim_{g\to-\infty} \frac{H(g)-1/2}{H(g)-H(g-1)} = -\infty$. However, $\rho(r) \in [0,1]$, so $r^*(0) = 0$. In contrast, let $q_{NB} = 1$. Then $g^*_{NB} = 1$ since $H(0) = 1/2$. But then $\frac{h(1)}{h(0)} > 0$. Thus, $q_{NB} = 1$ is not a fixed point of $g^*_{NB}(g)$ either. Notice that $h(g)/h(g-1) > 0$ and by its monotonicity, $g^*_{NB}(1) = \infty$, and again, by the monotonicity of $\frac{H(g)-1/2}{H(g)-H(g-1)}$, $\lim_{g\to\infty} \frac{H(g)-1/2}{H(g)-H(g-1)} = \infty$. But again, since $\rho(r) \in [0,1]$, $r^*(1) = 1$.

The shape of $g^*_{NB}(q)$ implies that it has a unique fixed point. To see this, notice that the derivative of $g^*_{NB}(q)$ is single-peaked since it is S-shaped, which implies that it has a unique inflection point. Thus, there exists at most a unique pair $\underline{q}_{NB}$ and $\overline{q}_{NB}$ such that $\frac{\partial}{\partial q_{NB}} g^*_{NB}(q_{NB}) > 1$ iff $q \in (\underline{q}_{NB}, \overline{q}_{NB})$. There exists a fixed point in between the two by the intermediate value theorem; For $q_{NB} \in (0, \underline{q}_{NB})$, $g^*_{NB}(q_{NB}) < q_{NB}$ since $g^*_{NB}(0) = 0$ and $g^*_{NB}(q_{NB})$ is convex and, for $q_{NB} \in (\overline{q}_{NB}, 1)$, $g^*_{NB}(q_{NB}) > q_{NB}$ since $g^*_{NB}(1) = 1$ and $g^*_{NB}(q_{NB})$ is concave.

Thus, $q^\dagger_{BE}$ such that $\rho(\hat{r}^*_B(g^*_{NB}(q^\dagger_{BE}))) = q^\dagger_{BE}$ at most uniquely exists in $(0,1)$, and there exists a unique pair $(g^*_{NB}, q^\dagger_{BE})$ defined by each other in $(0,1)^2$. Notice $\rho(r^*_B(g^*_{NB}(q)))) \geq q$ iff $q \geq q^\dagger_{BE}$.

Suppose $q^\dagger_{BE} \leq 1/2$ and $g^*_B(1/2) \leq 1/2$ so

$$\rho(r^*_B(g^*_B(1/2)))) = \rho\Big(\frac{H(g^*_B(1/2)) - 1/2}{H(g^*_B(1/2)) - H(g^*_B(1/2) - 1)}\Big) \leq \rho(1/2) = 1/2.$$

Notice that $\frac{H(g)-1/2}{H(g)-H(g-1)}$ attains $1/2$ iff $g = 1/2$, so $g^*_B(1/2)) = 1/2$ if $q^*_{BE} = 1/2$. However, $\frac{h(1/2)}{h(1/2)} = 1 < \frac{1/2}{1/2}\frac{1-F(1/2|1)}{1-F(1/2|0)}$. Thus, $g^*_B(1/2)$ such that $\frac{h(g^*_B)}{h(g^*_B-1)}\frac{1-F(\hat{r}^*_B(g^*_B)|0)}{1-F(\hat{r}^*_B(g^*_B)|0)} = 1$ is larger than $1/2$. This contradicts the assumption $g^*_B(1/2) \leq 1/2$. Thus, $q_{BE} > 1/2$.

To be more general, there exists a unique $\overline{P} > 1/2$ such that $\hat{r}^*_B(g^*_B(1/2)) \leq 1/2$ iff $\Pr[reelect|a = 0] \leq \overline{P}$. To see this, notice that $r^*_B(g^*_B(1/2)) \leq 1/2$ iff $\frac{H(g^*_B(1/2))-(1-\Pr[reelect|a=0])}{H(g^*_B(1/2))-H(g^*_B(1/2)-1)} \leq 1/2$. Observe that $\frac{H(g)-(1-\Pr[reelect|a=0])}{H(g)-H(g-1)} \leq \frac{1}{2}$ iff $1 - \Pr[reelect|a = 0] \geq \frac{H(g)+H(g-1)}{2} \iff 1 - \frac{H(g)+H(g-1)}{2} \geq \Pr[reelect|a = 0]$. Let $\overline{P} \equiv 1 - \frac{H(g^*_B(1/2))+H(g^*_B(1/2)-1)}{2} > 1/2$. Notice that $\Pr[reelect|a = 0] \leq \overline{P}$ iff $\hat{r}^*_B(g^*_B(1/2)) \leq 1/2$. ■

From this benchmark, it is clear that even without bureaucratic sabotage, the status quo's value, q, determines the type of policy inefficiency that arises in equilibrium. There exists a unique interior value of the status quo $q^\dagger_{BE} \in (0,1)$, such that the incumbent's decision is efficient; $\rho(r^*_B) = q^\dagger_{BE}$. If the status quo's value is less than this unique point, $q < q^\dagger_{BE}$, the voter's distaste for the status quo induces the incumbent to be overly zealous, leading to over-reform in expectation; $\rho(r^*) < q$. In contrast, if $q > q^\dagger_{BE}$, the voter's prior leads to policy inefficiency due to under-reform; $\rho(r^*) > q$.

A.3.3 Sabotage's Effect on Equilibrium

Proof for Propositions 2, 3, and 4.

Sabotage's Effect on Voter Inference
Observe the following:

$$\log L(\hat{r}^*(g), g) = \log \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)} + \log \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$$
$$\log L_B(\hat{r}^*_B(g), g) = \log \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} + \log \frac{h(g)}{h(g-1)}.$$

There exists $q^\dagger$ such that $\log L(\hat{r}^*(g^*(q^\dagger)), g^*(q^\dagger)) = \log L_B(\hat{r}^*_B(g^*_B(q^\dagger)), g^*_B(q^\dagger)) = \log \frac{1-q^\dagger}{q^\dagger}$. To see this, first, recall that

$$\frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \leq \frac{h(g)}{h(g-1)} \iff g \geq 1/2.$$

Furthermore, observe

$$\frac{\partial^2}{\partial c \partial g} \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} = -\frac{\left(1 - c\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}\right)\frac{\partial}{\partial g}\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}}{\left(1 + c\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}\right)^3}.$$

Recall that $\frac{\partial}{\partial g}\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)} > 0$. Thus, $\frac{\partial^2}{\partial c \partial g}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} > 0$ if and only if $1 > c\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}$. Since $\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}$ is increasing in g and 0 at $g = 1/2$, there exists a unique $g > 1/2$ such that this holds; there exists a unique $g_{cross-partial} > 1/2$ such that $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ is convexly decreasing in g, so that $\log \frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)} < 0$ and increases iff $g > g_{cross-partial}$.

Sabotage makes the incumbent more conservative for the reform, holding the voter's cutoff fixed.

$$\underbrace{H(g) - H(g-1) \geq c}_{\text{sabotage is IC, } k^* \leq 1} \iff \hat{r}^*(g) \geq \hat{r}^*_b(g) \iff \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} - \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)} \geq 0.$$

If $c < \bar{c} \equiv H(1/2) - H(-1/2) = 2H(1/2) - 1$, there exists $g_c > 0$ such that $H(g) - H(g-1) \geq c$ if and only if $g \in \left(\frac{1}{2} - g_c, \frac{1}{2} + g_c\right)$ since $H(g) - H(g-1)$ is single-peaked and symmetric around 1/2. Let $\underline{q} \equiv \frac{1}{1+L(\frac{1}{2}+g_c, \hat{r}^*(\frac{1}{2}+g_c))}$ and $\overline{q} \equiv \frac{1}{1+L(\frac{1}{2}-g_c, \hat{r}^*(\frac{1}{2}-g_c))}$.

Because $\rho(\hat{r}^*(g)) = \rho(\hat{r}^*_B(g))$ at $g = 0, \frac{1}{2} - g_c$, and $\frac{1}{2} + g_c$, $\rho(\hat{r}^*(g)) - \rho(\hat{r}^*_B(g)) > 0$ if and only if $g \in \left(\max\{0, \frac{1}{2} - g_c\}, \frac{1}{2} + g_c\right)$. Furthermore, $\rho(\hat{r}^*(g)) = \frac{H(g)-1/2}{c}$ is concave for $g \geq 0$ (Observe $h'(g) \leq \iff g \geq 0$.). In contrast, recall $\rho(\hat{r}^*_B(g)) = \frac{H(g)-1/2}{H(g)-H(g-1)}$ is reverse

S-shaped in $[0,1]$ with the inflection point at $g = 1/2$. Therefore, $\rho(\hat{r}^*(g)) - \rho(\hat{r}^*_B(g))$ is single-peaked and attains the maximum at $g > 1/2$ as $\frac{\partial^2 \rho(\hat{r}^*(g))}{\partial g^2} < 0$ and $\frac{\partial^2 \rho(\hat{r}^*(g))}{\partial g^2} > 0$ in $g \in \left(1/2, 1/2 + g_c\right)$, by the intermediate value theorem.

Because $\rho^{-1}(\cdot)$ is monotonically increasing, $\hat{r}^*(g) - \hat{r}^*_B(g)$ is also single-peaked and there exists a non-degenerate interval whose supremum is $1/2 + g_c$ where $\frac{\partial \hat{r}^*_B(g)}{\partial g}$ decreases and $\frac{\partial \hat{r}^*(g)}{\partial g}$ increases as g increases.

Again, because $\frac{1-F(\cdot|0)}{1-F(\cdot|1)}$ is monotonically decreasing, there exists $\delta_g \in (0, 1/2)$ such that $\log \frac{1-F(\hat{r}^*_B(g)|0)}{1-F(\hat{r}^*_B(g)|1)} - \log \frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} \in (0, \epsilon_\rho)$ if $g \in (\frac{1}{2} + g_c - \delta_g, \frac{1}{2} + g_c]$ for any $\epsilon_\rho > 0$. Therefore, there exists a unique $g''' \in \left(\max\{0, \frac{1}{2} - g_c\}, \frac{1}{2} + g_c\right)$ such $\log \frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} - \log \frac{1-F(\hat{r}^*_B(g)|0)}{1-F(\hat{r}^*_B(g)|1)} > 0$ is monotonically decreasing in g for $g \in \left(g''', \frac{1}{2} + g_c\right)$.

If $g \in \left(1/2 - g_c, 1/2\right]$, then

$$\log L_B(\hat{r}^*_B(g), g) - \log L(\hat{r}^*(g), g) > 0$$
$$\iff \underbrace{\log \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} - \log \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)}}_{\geq 0 \text{ and decreasing in } g \text{ iff } g > g'''} + \underbrace{\log \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)}}_{>0 \text{ and decreasing in } g} > 0.$$

Since L and L_B are decreasing, $g^*(q) < g^*_B(q)$ if $g^*_B(q) \leq 1/2$.

Suppose $g \in \left(1/2, 1/2 + g_c\right)$. Then $\log \frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)} < 0$ and monotonically increasing if $g > g_{cross-partial} > 1/2$ and hand, $\log \frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)} - \log \frac{1-F(\hat{r}^*_B(g)|0)}{1-F(\hat{r}^*_B(g)|1)} \geq 0$ is decreasing in g iff $g > g'''$. In short,

$$\log L_B(\hat{r}^*_B(g), g) - \log L(\hat{r}^*(g), g)$$
$$= \underbrace{\log \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} - \log \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)}}_{\geq 0 \text{ and decreasing in } g \text{ iff } g > g'''} + \underbrace{\log \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)}}_{\leq 0 \text{ and increasing in } g \text{ iff } g > g_{cross-partial}}.$$

Then there exists $g^\dagger > \max\{g_{cross-partial}, g'''\}$ such that $\log L_B(\hat{r}^*_B(g), g) - \log L(\hat{r}^*(g), g) = 0$ by the Intermediate Value Theorem. Define $q^\dagger$ such that $\frac{1-q^\dagger}{q^\dagger} = \log L(\hat{r}^*(g^\dagger), g^\dagger)$.

Notice that $q > q^\dagger$ iff $g^*(q) > g^*_B(q)$. Also, it is straightforward that $r^*(q) > r^*_B(q)$ if $q > q^\dagger$.

Normative Benchmark with Sabotage

Observe that there exists a unique $g^*_{NS}(q)$ such that

$$\frac{H(g^*_{NS}(q)) - 1/2}{c} = q$$

and $q^*_{NS}(g)$ such that

$$\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}=\frac{1-q^*_{NS}(g)}{q^*_{NS}(g)}\frac{1-F(\rho^{-1}(q^*_{NS}(g))|1)}{1-F(\rho^{-1}(q^*_{NS}(g))|0)}.$$

Since $\frac{H(g)-1/2}{c}$ is increasing in g and $\frac{1-q}{q}\frac{1-F(\rho^{-1}(q)|1)}{1-F(\rho^{-1}(q)|0)}$ is increasing in q, we can define $g^*_{NS}(q)$ and $q^*_{NS}(g)$ as increasing functions.

Notice that $\frac{H(g)-1/2}{c}$ is concave. Since $\frac{H(g)-1/2}{c}=0$ at $g=0$ and $H(1)-1/2=H(1)-H(0)<c$, so $\frac{H(g)-1/2}{c}<1$ at $g=1$. Therefore, there exists a unique g^*_{SE} such that $\frac{H(g^*_{SE})-1/2}{c}=g^*_{SE}$. Define $q^\dagger_{SE}\equiv q^*_{NS}(g^*_{SE})$.

Suppose that

$$\frac{h(g^*_{NS}(q^\dagger_{BE}))}{h(g^*_{NS}(q^\dagger_{BE}))+c\frac{h(g^*_{NS}(q^\dagger_{BE})-1)-h(g^*_{NS}(q^\dagger_{BE}))}{H(g^*_{NS}(q^\dagger_{BE}))-H(g^*_{NS}(q^\dagger_{BE})-1)}}=\frac{h(g^*_{NS}(q^\dagger_{BE}))}{h(g^*_{NS}(q^\dagger_{BE})-1)}=\frac{1-q^\dagger_{BE}}{q^\dagger_{BE}}\frac{1-F(\rho^{-1}(q^\dagger_{BE})|1)}{1-F(\rho^{-1}(q^\dagger_{BE})|0)}.$$

Recall that $q^\dagger_{BE}>1/2$, so $g^*(q^\dagger_{BE})=q^\dagger_{BE}>1/2$. Recall also $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}>\frac{h(g}{h(g-1)}$ if $g>1/2$ and both are decreasing in g. Therefore,

$$\frac{h(g^*_{NS}(q^\dagger_{BE}))}{h(g^*_{NS}(q^\dagger_{BE}))+c\frac{h(g^*_{NS}(q^\dagger_{BE})-1)-h(g^*_{NS}(q^\dagger_{BE}))}{H(g^*_{NS}(q^\dagger_{BE}))-H(g^*_{NS}(q^\dagger_{BE})-1)}}=\frac{h(g^*_{NS}(q^\dagger_{BE}))}{h(g^*_{NS}(q^\dagger_{BE})-1)}$$

implies that $g^*_{NS}(q^\dagger_{BE})>g^*_{NB}(q^\dagger_{BE})=q^\dagger_{BE}$. This further implies that $q^\dagger_{BE}>q^\dagger_{SE}$.

Suppose $q^\dagger_{SE}=1/2$, so $g^*_{NS}(1/2)=1/2$. But this leads to a contradiction since

$$\frac{h(1/2)}{h(1/2)+c\frac{h(-1/2)-h(1/2)}{H(1/2)-H(-1/2)}}=1<\frac{1/2}{1/2}\frac{1-F(\rho^{-1}(1/2)|1)}{1-F(\rho^{-1}(1/2)|0)}.$$

Thus, as in the case of $q^\dagger_{BE}$, $q^\dagger_{SE}>1/2$.

Sabotage's Effect on the Incumbent's Action

Now, consider $q^{\dagger\dagger}$ such that

$$\frac{H(g^*(q^{\dagger\dagger}))-1/2}{c}=\frac{H(g^*_B(q^{\dagger\dagger}))-1/2}{H(g^*_B(q^{\dagger\dagger}))-H(g^*_B(q^{\dagger\dagger})-1)},\tag{8}$$

so $r^*_B(q^{\dagger\dagger})=r^*(q^{\dagger\dagger})$. Hold $g^*_B(q^{\dagger\dagger})=g^{\dagger\dagger}_B$ fixed. There exists a unique $g^{\dagger\dagger}(g^{\dagger\dagger}_B)<g^{\dagger\dagger}_B$ such that $\frac{H(g)-1/2}{c}\geq\frac{H(g^{\dagger\dagger}_B)-1/2}{H(g^{\dagger\dagger}_B)-H(g^{\dagger\dagger}_B-1)}$ iff $g\geq g^{\dagger\dagger}(g^{\dagger\dagger}_B)$ if $c<H(g)-H(g-1)$.

Suppose that $g^*(q^{\dagger\dagger})=g^{\dagger\dagger}(g^{\dagger\dagger}_B)\geq 1/2$. Recall $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}\geq\frac{h(g)}{h(g-1)}$ iff $g\geq 1/2$. Since

$\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ is decreasing in g,

$$\frac{h(g^*(q^{\dagger\dagger}))}{h(g^*(q^{\dagger\dagger}))+c\frac{h(g^*(q^{\dagger\dagger})-1)-h(g^*(q^{\dagger\dagger}))}{H(g^*(q^{\dagger\dagger}))-H(g^*(q^{\dagger\dagger})-1)}} \geq \frac{h(g^{\dagger\dagger})}{h(g^{\dagger\dagger})+c\frac{h(g^{\dagger\dagger}-1)-h(g^{\dagger\dagger})}{H(g^{\dagger\dagger})-H(g^{\dagger\dagger}-1)}} \geq \frac{h(g^{\dagger\dagger})}{h(g^{\dagger\dagger}-1)}.$$

Notice that this leads to a contradiction since

$$\frac{1-F(r^*(q^{\dagger\dagger})|0)}{1-F(r^*(q^{\dagger\dagger})|1)}\frac{h(g^*(q^{\dagger\dagger}))}{h(g^*(q^{\dagger\dagger}))+c\frac{h(g^*(q^{\dagger\dagger})-1)-h(g^*(q^{\dagger\dagger}))}{H(g^*(q^{\dagger\dagger}))-H(g^*(q^{\dagger\dagger})-1)}} = \frac{1-F(r^*(q^{\dagger\dagger})|0)}{1-F(r^*(q^{\dagger\dagger})|1)}\frac{h(g^{\dagger\dagger})}{h(g^{\dagger\dagger}-1)} = \frac{1-q^{\dagger\dagger}}{q^{\dagger\dagger}}$$
$$\iff \frac{h(g^*(q^{\dagger\dagger}))}{h(g^*(q^{\dagger\dagger}))+c\frac{h(g^*(q^{\dagger\dagger})-1)-h(g^*(q^{\dagger\dagger}))}{H(g^*(q^{\dagger\dagger}))-H(g^*(q^{\dagger\dagger})-1)}} = \frac{h(g^{\dagger\dagger})}{h(g^{\dagger\dagger}-1)}$$

by assumption. Therefore, $g^*(q^{\dagger\dagger}) < 1/2$. Notice that $q^{\dagger}_{SE} > 1/2$ implies that $q^{\dagger\dagger} < 1/2$.

Comparison between two different costs: Suppose $c_h > c_l > 0$ such that there exists a non-degenerate open interval of g such that $H(g) - H(g-1) > c_h$.

Define

$$L_i(g) = \frac{1-F(\hat{r}^*(g;c_i)|0)}{1-F(\hat{r}^*(g;c_i)|1)}\frac{h(g)}{h(g)+c_i\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$$

where $i = h, l$ and

$$\hat{r}^*(g;c_i) = \rho^{-1}\Big(\frac{H(g)-1/2}{c_i}\Big).$$

Notice that $\hat{r}^*(g;c_h) < \hat{r}^*(g;c_l)$ and

$$\frac{1-F(\hat{r}^*(g;c_h)|0)}{1-F(\hat{r}^*(g;c_h)|1)} > \frac{1-F(\hat{r}^*(g;c_l)|0)}{1-F(\hat{r}^*(g;c_l)|1)}.$$

If $g \leq 1/2$, then

$$\frac{h(g)}{h(g)+c_h\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} > \frac{h(g)}{h(g)+c_l\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}.$$

Thus, for g_i^* such that $L_i(g_i^*) = \frac{1-q}{q}$, $g_h^* > g_l^*$ if $g \leq 1/2$ since $L_i(g)$ is decreasing and $L_h(g) > L_l(g)$.

We can define $g_{cross-partial}(c_l) > 1/2$ such that $\frac{\partial^2}{\partial g \partial c_l}\frac{h(g)}{h(g)+c_l\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \geq 0$ if and only if $g \leq g_{cross-partial}(c_l)$. Thus, there exists $g^{\dagger}(c_l) > g_{cross-partial}(c_l)$ such that $L_h(g) \leq L_l(g)$ $g_h^* \leq g_l^*$ if and only if $g \geq g^{\dagger}(c_l)$.

Notice that the argument above implies that there exists a unique $q^{\dagger\dagger}(c_l) < 1/2$ such that $\hat{r}^*(g_h^*(q);c_h) \leq \hat{r}^*(g_l^*(q);c_l)$ iff $g \geq g^{\dagger\dagger}(c_l)$. Also, for $q_E(c_i) > 1/2$ such that $\rho(\hat{r}^*(g_i^*(q_E(c_i)))) =$

$q_E(c_i)$, $q_E(c_h) > q_E(c_l)$ (See the argument about $q^\dagger_{BE} > q^\dagger_{SE}$.).

Cost's Effect on Sabotage Incentive

Proof. By the chain rule,

$$\begin{aligned}\frac{\partial \kappa^*}{\partial c} &= \frac{H(g^*) - H(g^*-1) - c\frac{\partial H(g^*)-H(g^*-1)}{\partial c}}{[H(g^*) - H(g^*-1)]^2} \geq 0 \\ \iff & H(g^*) - H(g^*-1) \geq c\frac{\partial g^*}{\partial c}\frac{\partial H(g^*) - H(g^*-1)}{\partial g^*} \\ \iff & 1 \geq c\frac{\partial g^*}{\partial c}\frac{\frac{\partial H(g^*)-H(g^*-1)}{\partial g^*}}{H(g^*) - H(g^*-1)}\end{aligned}$$

Notice

$$\frac{\partial g^*}{\partial c}\frac{\frac{\partial H(g^*)-H(g^*-1)}{\partial c}}{H(g^*) - H(g^*-1)} = \frac{\partial}{\partial c}\log\Big(H(g^*) - H(g^*-1)\Big).$$

So, $\frac{\partial \kappa^*}{\partial c} \geq 0$ if and only if $\frac{1}{c} \geq \frac{\partial}{\partial c}\log\Big(H(g^*) - H(g^*-1)\Big)$.

Recall that g^* is decreasing in c if and only if $g^* < g^\dagger < 1/2$. Suppose that $g^* < 1/2$, so g^* is decreasing in c. As g^* decreases at $g^* < 1/2$, $H(g^*) - H(g^*-1)$ is decreasing, so $\frac{\partial}{\partial c}\log\Big(H(g^*) - H(g^*-1)\Big) < 0$.

If $g^* > 1/2$, $H(g^*) - H(g^*-1)$ is decreasing in g^*, so $\frac{\partial}{\partial c}\log\Big(H(g^*) - H(g^*-1)\Big) < 0$.

Thus, κ^* is increasing in c. ■ ■

A.4 Extension: Type

Proof for Lemma 2 and Proposition 5. Notice that L^*_τ is increasing in τ if and only if $1-F(\hat{r}^*(g)|0) > 1-F(\hat{r}^*(g)|1)$. However, $1-F(r'|0) < 1-F(r'|1) \iff F(r'|0) > F(r'|1)$ for $r' > 0$, because MLRP implies the first-order stochastic dominance (Bagnoli and Bergstrom, 2006). ■

A.5 Cost from a non-uniform distribution

Suppose that κ is drawn from a log-concave distribution $P(\cdot)$ with support $[0,1]$ and associated pdf $p(\cdot)$. Then the equilibrium probability of sabotage is $1-P(\kappa^*) = 1-P\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big)$. Thus, equation (3) is now

$$\rho(r)P(\kappa^*)\Big(H(g^*) - H(g^*-1)\Big) + 1 - H(g^*) \geq 1/2$$

and

$$r^* = \rho^{-1}\left(\frac{H(g^*) - 1/2}{P\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big)\Big(H(g^*) - H(g^* - 1)\Big)}\right).$$

$P\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big)\Big(H(g^*) - H(g^* - 1)\Big) < \Big(H(g^*) - H(g^* - 1)\Big)$, and can be larger or smaller than c depending on that P is concave or convex, but does not qualitatively affect r^*'s property. To see this, notice that because P's log-concave, there exists a unique $g_p^* \in (0, 1)$ such that $p'\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big) < 0$ if and only if g^* is larger than that g_p^*. For instance if $p'\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big) > 0$, then $P\Big(\frac{c}{H(g^*)-H(g^*-1)}\Big)\Big(H(g^*) - H(g^* - 1)\Big) < c$. Then q is small enough to make $g^* < g_p^*$ in equilibrium, the sabotage's effect on the incumbent's action is smaller than when κ is drawn from a uniform distribution. In contrast, if q is larger than that value, so $g^* > g_p^*$ in equilibrium, the effect of sabotage on the incumbent's action is larger than when κ is drawn from a uniform distribution. But since the effect is marginal and mean-preserving spread across g_p^*, it does not affect the uniqueness of g^*, κ^*, and r^*.

B Example for Sabotage's Countervailing Effects on Voter Inference

Here we provide a specific example for the results discussed in Section 4.3.1, fixing the values of g' to those shown in Figure 6. The area of each cell represents the probability of each event and adds up to one. In both panels, the ex-ante total probability of successful reform $\Pr[x = 1] = \Pr[TP] + \Pr[FN] = 1/2$. Without sabotage,

$$\Pr[x = 1|g \geq g'] = \Pr[x = 1|g < g'] = \frac{1}{2}.$$

If bureaucrats sabotage, they do so with probability $1/2$, and TP (blue shaded area in broken lines, "Sabotaged TP") becomes FN (red shaded area in solid lines, "Added FN").

In Panel (a), the voter's cutoff is high ($g' = 0.7$), so observing a high signal is rare ($\Pr[g \geq g'] = 0.3$). As sabotage decreases $\Pr[TP]$ by 50%,

$$\Pr[x = 1|g \geq g'] = \frac{\Pr[TP]}{\Pr[TP] + \Pr[FP]} = \frac{0.3 * 0.5 * 0.5}{0.3 * 0.5 * 0.5 + 0.3 * 0.5} = \frac{1}{3} < \frac{1}{2},$$

and

$$\Pr[x = 1|g < g'] = \frac{\Pr[FN]}{\Pr[FN] + \Pr[TN]} = \frac{0.7 * 0.5 + 0.3 * 0.5 * 0.5}{0.7 * 0.5 + 0.3 * 0.5 * 0.5 + 0.7 * 0.5} = \frac{0.85}{1.55} \approx 0.548 > \frac{1}{2}.$$

Evidently, the PPV effect is larger than the FOR effect.

In Panel (b), the voter's cutoff is low ($g' = 0.3$), so a positive signal is relatively more prevalent ($\Pr[g \geq g'] = 0.7$). Without sabotage,

$$\Pr[x = 1|g \geq g'] = \Pr[x = 1|g < g'] = \frac{1}{2}.$$

As sabotage decreases $\Pr[TP]$ by 50%,

$$\Pr[x = 1|g \geq g'] = \frac{\Pr[TP]}{\Pr[TP] + \Pr[FP]} = \frac{0.7 * 0.5 * 0.5}{0.7 * 0.5 * 0.5 + 0.7 * 0.5} = \frac{1}{3} < \frac{1}{2},$$

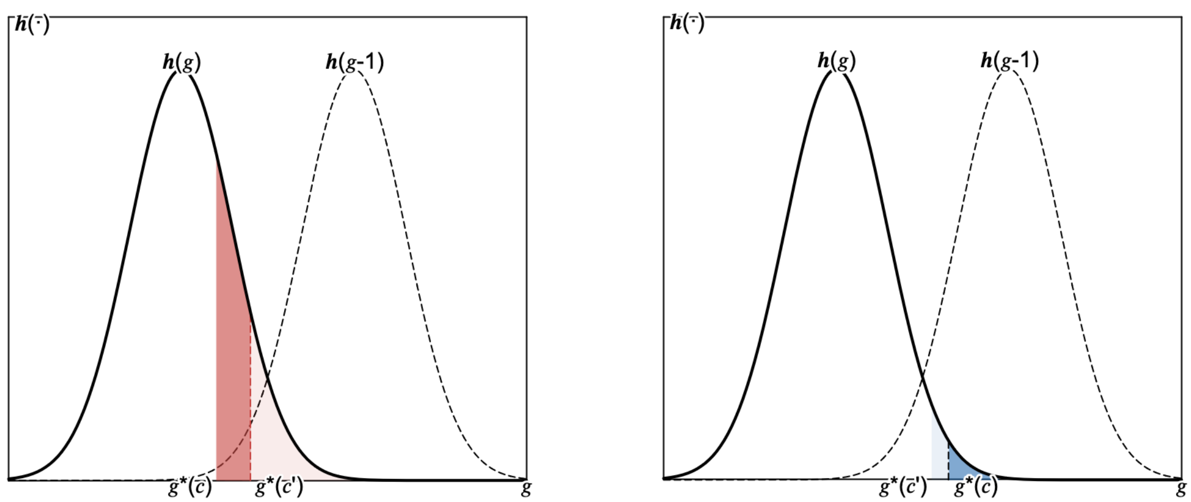
and

$$\Pr[x = 1|g < g'] = \frac{\Pr[FN]}{\Pr[FN] + \Pr[TN]} = \frac{0.3 * 0.5 + 0.7 * 0.5 * 0.5}{0.3 * 0.5 + 0.7 * 0.5 * 0.5 + 0.3 * 0.5} = \frac{0.65}{0.95} \approx 0.684 > \frac{1}{2}.$$

Here, the FOR effect is larger and dominates the PPV effect. For the general result, see the Appendix of Heo and Landa (2024).

C Additional Illustrations

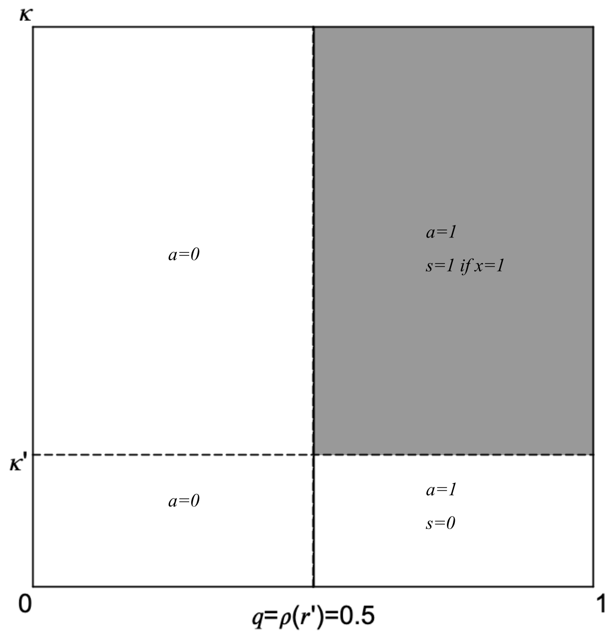
Figure A1: Direct and Inference Effect of Sabotage on Policy-Making



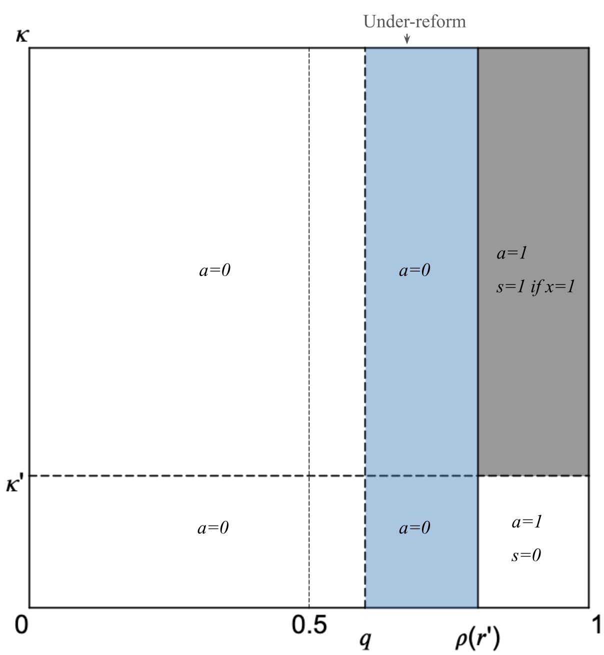
(a) Low q: Sabotage *increases* re-election chances.

(b) High q: Sabotage *decreases* re-election chances.

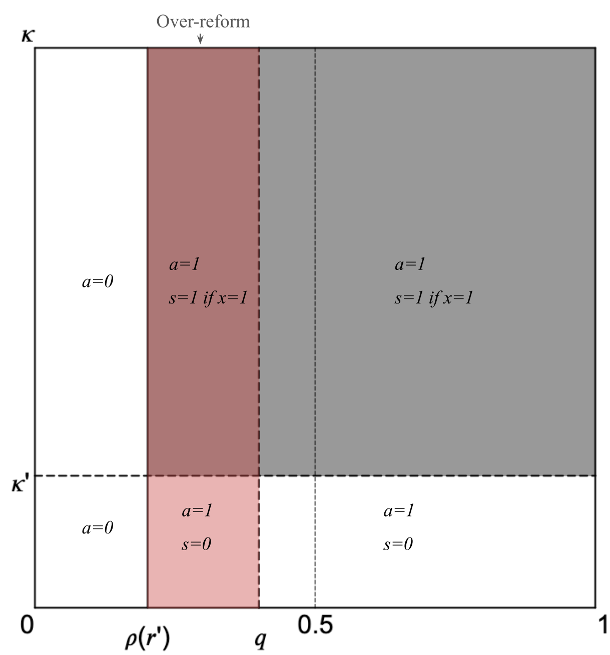
Figure A2: Equilibrium Outcomes



(a) Efficient Reform, $q = 0.5$



(b) Under-Reform, $q > 0.5$



(c) Over-Reform, $q < 0.5$