Bureaucratic Sabotage and Policy Inefficiency*

Kun Heo[†] and Elisa M. Wirsching[‡]

June 26, 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 electoral politics 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 show that bureaucrats are most incentivized to sabotage when voters are torn between the reform and the status quo, leading them to interpret poor service provision as informative of the reform's merit. We also find that bureaucratic sabotage causes two types of policy inefficiency depending on voters' perceptions of the reform's value. 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

‡Post-Doctoral Fellow, Center for the Study of Democratic Politics, Princeton 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 electoral incentives to enact reforms. With this framework, we shed light on a unique source of political power for bureaucrats and its consequences for public policy.

We integrate bureaucratic sabotage into a model of electoral politics and policy-making. An 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 opponent. The incumbent and opponent are both office motivated as well as policy motivated. While the incumbent receives some extra benefit from having the reform, the opponent has some anti-reform bias. 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. 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 theory requires several scope conditions. First, the model assumes a context where bureaucrats have a status quo bias, i.e., a "vested interest" in avoiding reforms that affect bureaucrats' money, programs, and policy direction (Moe, 2015). Second, we operate in a partisan environment with some degree of preference asymmetry between the incumbent and the opponent regarding reform policies. One can think of policies that bear sufficient cleavages between liberals and conservatives, e.g., budgets for law enforcement, vaccine mandates for public employees, or the extent of environmental protection. Third, we abstract away from the standard and well-studied issue of political delegation, where politicians seek to control bureaucrats who shirk their duties to avoid effort costs or to influence policy (e.g., Epstein and O'Halloran (1999); Huber and Shipan (2002); Yazaki (2018); Slough (2024)).

Instead, we focus on bureaucrats with considerable discretion (e.g., street-level bureaucracies) who trade off their motivation to serve the public with their incentives to affect public service provision for political leverage. Finally, we examine contexts in which voters obtain information about the effectiveness of policies through public services but struggle to clearly attribute poor service provision to either bureaucrats' behavior or politicians' policy choices.

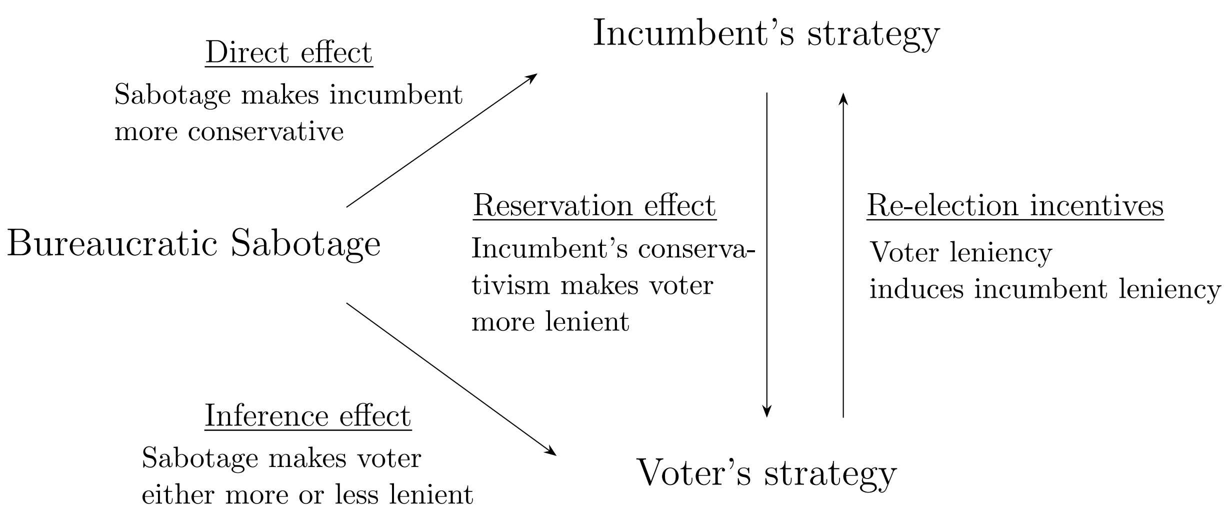
Our model produces several key insights. First, we demonstrate when and why bureaucrats sabotage public service provision for political leverage. Since voters cannot perfectly identify who is responsible for poor service quality and can only probabilistically determine whether bureaucratic sabotage has occurred, it becomes optimal for bureaucrats—provided the costs of sabotage are low enough—to engage in sabotage despite voters' awareness of this possibility. After incumbents introduce the reform, bureaucrats can exploit their intermediary role in government to affect voters' inference about the reform and undermine the incumbent's reelection chances in favor of the anti-reform opponent.

We also find that bureaucrats' incentive to sabotage is non-monotonic with respect to the voter's prior belief about the reform's value. The incentive to sabotage 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.

Second, we show how bureaucratic sabotage affects the political equilibrium between the incumbent and the voter. A naive conclusion may 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



2 Related Literature and Contributions

We make several contributions to existing scholarship on bureaucratic politics, interest group influence, and political economy.

2.1 Bureaucratic Politics and Interest Groups

First, our theory addresses a fundamental debate in bureaucratic politics between the public choice school (Tullock, 1965; Downs, 1967; Niskanen, 1971) and theories of bureaucratic control and delegation (Miller and Moe, 1983; McCubbins, 1985; McCubbins, Noll and Weingast, 1987; Banks and Weingast, 1992; Brehm and Gates, 1997). Niskanen positioned bureaucrats as the primary strategic actors and famously argued that self-interested bureaucrats use their private information to extract rents through bargaining by making take-it-or-leave-it offers to incumbents. In contrast, theories of legislative control criticized Niskanen's framework for ascribing out-sized power to bureaucrats. They framed the politician-bureaucrat relationship as a top-down principal-agent model and focused on incumbents' strategies to minimize agency loss and leverage 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 who can sway the incumbent's policy decisions by leveraging their private information, exploiting the incumbent's electoral vulnerability, and adjusting their work effort.

Additionally, we contribute to the growing literature on bureaucrats as interest groups within government. We build on Moe (2006)'s argument that bureaucrats leverage politicians' electoral vulnerability to influence who their principals are and what policies they choose in office. An extensive literature highlights bureaucrats' various means of direct political influence through their public sector unions, including 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). In contrast, we focus on a more fundamental source of bureaucratic power and explain how and when bureaucrats can exert policy influence through their role in government, i.e., *by the mere virtue of being bureaucrats.*

Third, we describe and micro-found a novel explanation for why bureaucratic agencies 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.

2.2 Formal Theory and Political Economy

Our model is closely connected to models of electoral competition where voters can predict each candidate's post-election policies, and the incumbent's actions shape voters' beliefs about the effectiveness of these policies (Bils and Izzo, 2023; Izzo, Martin and Callander, 2023; Bueno de Mesquita and Dziuda, 2023; Delgado-Vega, Dziuda and Loeper, 2023). The innovation in our model is the inclusion of strategic bureaucrats in this game.

Additionally, this paper is related to models where the incumbent and bureaucrats jointly produce government outcomes, creating difficulties for the voter to attribute responsibility between the two parties (Fox and Jordan, 2011; Ujhelyi, 2014; Yazaki, 2018; Forand and Ujhelyi, 2021; Martin and Raffler, 2021; Awad, Karekurve-Ramachandra and Rothenberg, 2023; Foarta, 2023; Slough, 2024). Yet, most of these models do not provide an explanation for why and when bureaucratic sabotage occurs.[1] One exception is Ujhelyi (2014), who also examines bureaucrats' strategic sabotage and its implications for policy-making. In both models, such sabotage distorts policy. The key difference lies in how electoral incentives operate. In Ujhelyi (2014), the incumbent has a strong intrinsic policy preference, and electoral incentives mitigate that preference. Consequently, players' equilibrium behavior largely depends on intrinsic preferences and the material costs of actions rather than voter beliefs. In contrast, in our model, electoral incentives lead to policy distortion rather than discipline (Canes-Wrone, Herron and Shotts, 2001; Gersen and Stephenson, 2014). Specifically, the incumbent in our model chooses the wrong reform *because of the election*, whereas the incumbent in Ujhelyi (2014) chooses it *in spite of the election*. Therefore, the key mechanism in our model is through voter inference rather than inherent types and thus provides new insights into the connection of voter beliefs and bureaucratic behavior.

Lastly, this paper is closely related to models of policy obstruction and sabotage (Patty, 2016; Fong and Krehbiel, 2018; Gieczewski and Li, 2022; Hirsch and Kastellec, 2022). 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 also study the implications of bureaucratic sabotage for policy-making.

[1] This is because bureaucrats are assumed to be non-strategic (e.g., their types perfectly determine their behavior) (Fox and Jordan, 2011; Martin and Raffler, 2021; Foarta, 2023), or because incumbents adjust their policy and delegation to bureaucrats based on factors influencing bureaucrats' motivation such that bureaucratic sabotage does not happen on the equilibrium path (Yazaki, 2018), or because bureaucrats and politicians are assumed to share policy preferences (Awad, Karekurve-Ramachandra and Rothenberg, 2023).

3 Model

Consider a two-period ($t = 1, 2$) electoral competition model with an incumbent (she), an opponent, a median voter (he), and the bureaucrats (they).[2] There is an election after $t = 1$ where the voter chooses between the incumbent and the opponent as a new officeholder for $t = 2$.

3.1 Politicians' Preferences and Policy-Making

We model a situation where partisan politicians decide whether to implement a reform policy that affects bureaucrats and their work. In $t = 1$, the officeholder decides whether to introduce reform. In $t = 2$, the officeholder decides whether to repeal the reform *if it was introduced in period one*. If reform is not introduced in $t = 1$, it cannot be reintroduced in $t = 2$. Thus, $t = 1$ is the window for reform (Keeler, 1993).[3]

Politicians are both office-motivated and policy-motivated. Both the incumbent and the opponent get 1 by winning the election and 0 otherwise. The incumbent is pro-reform and receives $\varepsilon \geq 0$ if and only if she chooses reform while in office, and the opponent is anti-reform and receives $\varepsilon \geq 0$ if and only if she chooses the status quo while in office. Since the voter knows the ideological leanings of the politicians, he can anticipate the second-period policy choice of each candidate. Note that we focus on cases where ε is small and dominated by re-election incentives in the first period.

Let $a \in \{0, 1\}$ indicate the incumbent's choice about whether to introduce the reform ($a = 1$) or not ($a = 0$). The incumbent's policy choice is public. The reform's period-invariant value to government outcome for each period, $\omega \in \{0, 1\}$, is unknown to the public. The common prior is $\Pr[\omega = 0] = 1/2$. If the incumbent chooses not to introduce the reform, the status quo policy of known value $q \in [0, 1]$ is implemented.[4] $1 - q$ can be interpreted as the probability that reform outperforms the status quo.

Before choosing a, the incumbent *privately* observes a signal $r \in [0, 1]$ about the reform's value ω. r is drawn from a conditional density $f(r|\omega)$ that satisfies the strict monotone likelihood ratio property;

$$\rho(r) := \Pr[\omega = 1|r] = \frac{f(r|1)}{f(r|0) + f(r|1)}$$

is monotonically increasing in r. $F(\cdot|\omega)$ denotes the CDF of $f(\cdot|\omega)$.

[2] We assume that players do not discount their future payoffs, which does not affect the qualitative results.
[3] We further justify this modeling approach in Section 3.8.
[4] The result is qualitatively similar if the status quo's value is $1/2$ and $\Pr[\omega = 0] = q$.

3.2 Sabotage

If the reform is implemented ($a = 1$), the incumbent, the opponent, and the bureaucrats privately observe ω, while the voter does not. This captures an information asymmetry between elected or non-elected officeholders and voters over policy.[5] Bureaucrats intrinsically dislike the reform ($a = 1$) and get disutility of *unknown* value $-\kappa$ with common prior $\kappa \sim U[0, 1]$.[6] After observing a (and ω if $a = 1$), the bureaucrats *privately* choose whether to sabotage the policy, $s \in \{0, 1\}$, where $s = 1$ is to sabotage the policy and $s = 0$ is not to 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. Sabotage is costly for bureaucrats, i.e., they incur a known cost of $c \in [0, 1]$ if they sabotage. c captures material/reputational punishments for non-compliance[7] (Ujhelyi, 2014), bureaucrats' public service motivations and preferences for high-quality service provision (Yazaki, 2018; Forand, Ujhelyi and Ting, 2022), or coordination efforts of bureaucrats necessary to engage in sabotage.[8]

3.3 Government Outcome

The government outcome $g \in \mathbb{R}$ is produced by

$$g = \begin{cases} (1-s)\omega + \eta & \text{if } a = 1 \\ (1-s)q + \eta & \text{if } a = 0 \end{cases}$$

where η is an i.i.d. shock drawn from a log-concave density $h(\cdot)$ that has full support on $\mathbb{R}$ and is symmetric around 0.[9] Let $H(\cdot)$ denote the associated CDF of $h(\cdot)$.

[5] The main results hold when relaxing this assumption, i.e., the model allows for the case where officeholders do not have superior information compared to the voters about the policy or the canonical case where bureaucrats are better informed than elected officials (Epstein and O'Halloran, 1999; Huber and Shipan, 2002).

[6] This is a simplifying assumption. The main results do not change qualitatively when κ is drawn from a log-concave distribution with different support (see Appendix A.4).

[7] Since our focus is on the logic and consequences of bureaucratic sabotage, we treat c as exogenous and do not incorporate how politicians can adjust their monitoring of or delegation to bureaucrats to affect the bureaucrats' cost-benefit trade-offs.

[8] Readers familiar with canonical principal-agent models may question the idea that shirking (rather than working) is costly for bureaucrats. However, we are not the first to assume a mirror image where policy-motivated bureaucrats face a trade-off between the benefits of sabotaging an unwanted policy and the material, reputational, or psychological costs of doing so (Brehm and Gates, 1997; Ujhelyi, 2014; Yazaki, 2018). Instead of minimizing the costs of positive effort for government output while accounting for its benefits (e.g., higher wages, avoiding political oversight), bureaucrats in this setting maximize the benefit from negative government output while taking the costs into account.

[9] See Appendix B for a discussion of the importance of noise for the incumbent and voter in our model.

3.4 Election

After observing the chosen policy a and the realized government outcome g, the voter chooses between the incumbent and the opponent. If the voter is indifferent between the two candidates, he flips a fair coin and reelects the incumbent with probability 1/2.

3.5 Second Period

Let $\tilde{a} \in \{0, 1\}$ denote the election winner's policy choice in $t = 2$ where $\tilde{a} = 1$ is to introduce (or keep) the reform and $\tilde{a} = 0$ is to keep (or to revert back to) the status quo.

If the incumbent chose the status quo in period one, the second-period policy is fixed as the status quo: $a = 0 \Rightarrow \tilde{a} = 0$.

If the incumbent introduced the reform in period one, then the incumbent gets $\varepsilon \geq 0$ if and only if she keeps reform, $\tilde{a}\varepsilon$. In contrast, the opponent gets ε if and only if she repeals it, $(1 - \tilde{a})\varepsilon$.

The government outcome in $t = 2$, $\tilde{g}$, is given by

$$\tilde{g} = \begin{cases} (1 - \tilde{s})\omega + \tilde{\eta} & \text{if } a = \tilde{a} = 1 \\ \\ (1 - \tilde{s})q + \tilde{\eta} & \text{if otherwise} \end{cases}$$

where $\tilde{s} \in \{0, 1\}$ is the bureaucrats' decision to sabotage the policy, and $\tilde{\eta}$ is a shock drawn from $h(\cdot)$.

3.6 Payoffs

The voter gets the government outcome in each period:

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

The incumbent gets ε if she chooses reform. Also, she gets 1 from winning the election:

$$a\varepsilon + \mathbf{1}\{\text{reelection}\}(1 + a\tilde{a}\varepsilon).$$

The opponent gets ε if she chooses the status quo and gets 1 if she wins the election:

$$\mathbf{1}\{\text{reelection}\}(1 - \tilde{a})\varepsilon.$$

The bureaucrats get $-\kappa$ in each period if the reform is in place. Also, they get $-c$ if they engage in sabotage:

$$-\underbrace{a(\kappa+\tilde{a}\kappa)}_{\text{disutility from the reform}}-\underbrace{c(s+\tilde{s})}_{\text{cost of sabotage}}.$$

3.7 Timing

To recap,

0. Nature draws the reform's value ω, the incumbent's signal r, and the bureaucrats' disutility from the reform, κ.
1. The incumbent privately observes the signal r and publicly chooses whether to introduce the reform ($a = 1$) or not ($a = 0$).
2. The incumbent, the opponent, and the bureaucrats observe the reform's value ω.
3. The bureaucrats privately observe their disutility from the reform κ and choose whether to sabotage the chosen policy ($s = 1$) or not ($s = 0$).
4. The government outcome g is produced, and the voter observes it.
5. The voter chooses between the incumbent and the opponent as the new officeholder in the election.
6. The election winner chooses the policy $\tilde{a}$ and the bureaucrats chose $\tilde{s}$.
7. Payoffs are realized, and the game ends.

3.8 Comments on Window for Reform

$t = 1$ in our model is a critical "watershed" point where the reform is either implemented or abandoned (Keeler, 1993). Formally, our assumption that the incumbent cannot delay the introduction of the reform to $t = 2$ means she commits *not to introduce the reform she rejected in the previous term.* We make this assumption because it (i) allows us to focus on a non-trivial game and (ii) will arise in equilibrium in the super-game where the incumbent chooses between situations with and without the commitment power.

To see this, first assume that the incumbent cannot commit to the status quo in the second period. The voter then knows that she will choose the reform regardless of her first-period policy. Thus, the incumbent's reelection probability depends on the voter's belief about the reform's value, whether it is introduced or not.

As r increases, $\rho(r) = \mathrm{E}[\omega|r]$ increases, leading to a higher probability that the voter observes a high g. Because the voter's expectation of the reform's value depends on his observation of g, as g becomes more likely, the probability increases that the voter will believe the reform outperforms the status quo, $\mathrm{E}[\omega|g] \geq q$. This leads to a higher reelection probability as the voter expects the incumbent to continue choosing the reform in $t = 2$. In short, there is an increasing mapping from r to the reelection probability with the reform from the incumbent's perspective.

This leads to the voter's expectation that the incumbent will choose the reform if she observes a high enough r. The upside of this expectation is that the incumbent can increase the voter's confidence by choosing the reform, as it signals that r is high. The downside is that the incumbent cannot choose the status quo without damaging the voter's expectation about the reform's value since the voter believes the incumbent would only choose the reform if r is high. Specifically, a rational voter will always consider the possibility that $r = 0$ if the incumbent chooses the status quo:[10]

$$\Pr[\text{reelection}|a = 0, \text{ No commitment}, r] \leq \Pr[\text{reelection}|a = 1, r]$$

As the incumbent has no incentive to make the voter consider $r = 0$ unless r is actually 0, the incumbent's inability to commit leads to an *unraveling* result where the incumbent *must choose the reform for* $r > 0$, making the game trivial.[11]

Now, suppose that the incumbent can make the voter's expectation about the reform's value irrelevant by committing to the status quo. Then, the incumbent who observes a low r such that

$$\Pr[\text{reelection}|a = 1, r] < \Pr[\text{reelection}|a = 0, \text{ Commitment}]$$

can cut her loss by committing to the status quo. Notice that the incumbent is at least weakly better off when she can commit to the status quo than when she cannot, as the commitment provides insurance for a low r. Therefore, we can expect the incumbent to develop a commitment device to tie her hands regarding the reform once she has chosen the status quo.

[10] Strictly speaking, $r = 0$ is a zero-measure event, so the consideration of $r = 0$ per se is irrelevant. More precisely, when calculating the reform's expected value, the voter will always integrate over a non-degenerate set of r that includes $r = 0$.

[11] To prevent unraveling, there should be a case where the incumbent legitimately turns down the reform that works, and she cannot credibly communicate the true reason why she chooses the status quo. For a more detailed explanation of the logic of unraveling results, see Milgrom (1981).

4 Analysis

The solution concept is a weak Perfect Bayesian Equilibrium with pure strategies (henceforth, equilibrium).

4.1 Second-Period Behavior

It is straightforward that the incumbent who wins the election with the reform continues it, $\mathbf{1}\{a = 1\} \times \mathbf{1}\{\text{reelection}\} \Rightarrow \tilde{a} = 1$, since she gets $\epsilon \geq 0$ by doing so and 0 otherwise. If she does not introduce it or the opponent wins the election, then the status quo is chosen, $\tilde{a} = 0$.

Regardless of the election winner or the policy she chooses, the bureaucrats do not have any incentive to sabotage and take its cost $-c < 0$, i.e., $\tilde{s} = 0$.

4.2 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. If the incumbent chooses the status quo in $t = 1$, $a = 0$, the voter is indifferent between the two candidates and reelects the incumbent with probability 1/2. If the incumbent chooses the reform, $a = 1$, the voter gets $\mathrm{E}[\omega|g]$—the conditional expectation of the reform's value given the government outcome g—if he reelects the incumbent and q if he votes for the opponent. In turn, the voter reelects the incumbent if and only if[12]

$$\mathrm{E}[\omega|g] \geq q. \tag{1}$$

To construct the voter's posterior belief, suppose the following about the incumbent's and the bureaucrats' strategies.

- The incumbent introduces the reform if she observes a high enough r, $r \geq r'$:$a^*(r) = \mathbf{1}\{r \geq r'\}$.
- The bureaucrats sabotage the reform if and only if (i) the reform works, $\omega = 1$, and their disutility from it is large enough, $\kappa \geq \kappa'$: $s^*(\kappa) = \omega \times \mathbf{1}\{\kappa \geq \kappa'\}$.

Then, the voter's conditional expectation of the reform's value is given by

$$\mathrm{E}[\omega|a = 1, g, \kappa', r'] = \Pr[\omega = 1|g, r \geq r'] = \frac{1}{1 + \lambda(g, \kappa', r')}$$

[12]Strictly speaking, the voter flips the coin if $\mathrm{E}[\omega|g] = q$, but we ignore this since it is a zero-measure event.

where

$$\lambda(g,\kappa',r') := \frac{\Pr[\omega=0]\Pr[g,r\geq r'|\omega=0]}{\Pr[\omega=1]\Pr[g,r\geq r'|\omega=1]} = \frac{1-F(r'|0)}{1-F(r'|1)}\frac{h(g)}{h(g)+\kappa'\Big(h(g-1)-h(g)\Big)}. \quad (2)$$

In turn, equation (1) can be rewritten as

$$\frac{1}{1+\lambda(g,\kappa',r')} \geq q \iff \lambda(g,\kappa',r') \geq \frac{1-q}{q}.$$

Further inspection of λ provides the following insights about the voter's inference and behavior:[13]

Lemma 1 *If $r'=0$, $E[\omega|g]=1/2$ at $g=1/2$.*

If $r'>0$ and $\kappa'>0$,

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

$$\lambda(g,\kappa',r') \geq \frac{1-q}{q} \quad (3)$$

   if and only if $g\geq\hat{g}^*$.

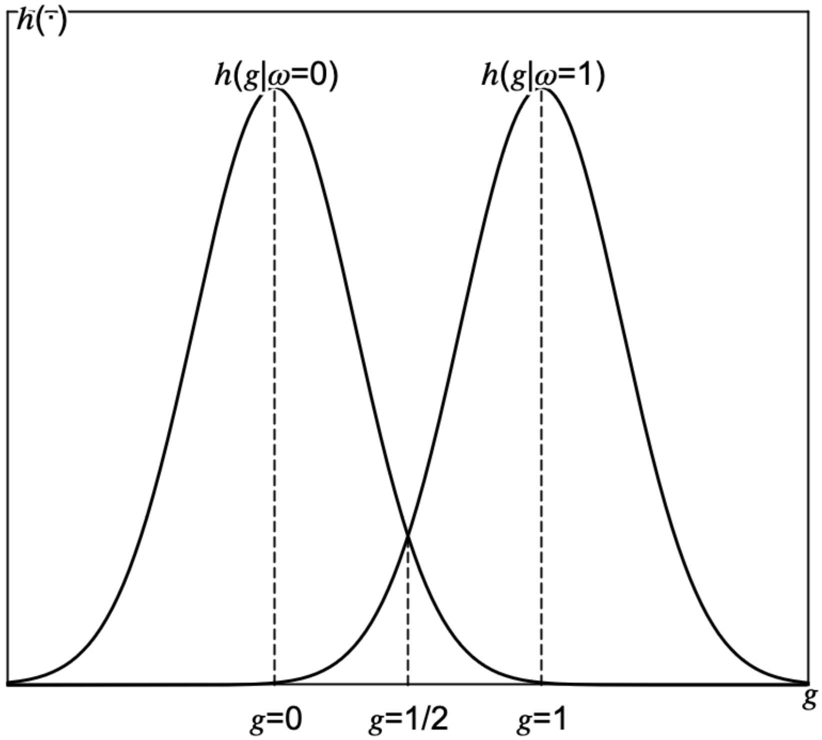
2. *$E[\omega|r\geq r',g]$ is increasing in r'; $\hat{g}^*$ that satisfies equation (3) decreases in r'.*

3. *$E[\omega|r\geq r',g]$ 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 if the incumbent introduces the reform for any $r\geq 0$, then the reform is equally likely to succeed or fail at $g=1/2$, ($\mathrm{E}[\omega|g=1/2]=1/2$) 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 ($\omega=1$) and not have been sabotaged ($s=0$) than either to fail ($\omega=0$) or have been sabotaged ($s=1$). In contrast, if $g<1/2$, then the voter's inference works in the opposite way. Figure 2 illustrates this logic, showing that $h(g)>h(g-1)$ if and only if $g<1/2$.

[13] All proofs are relegated to Appendix A.

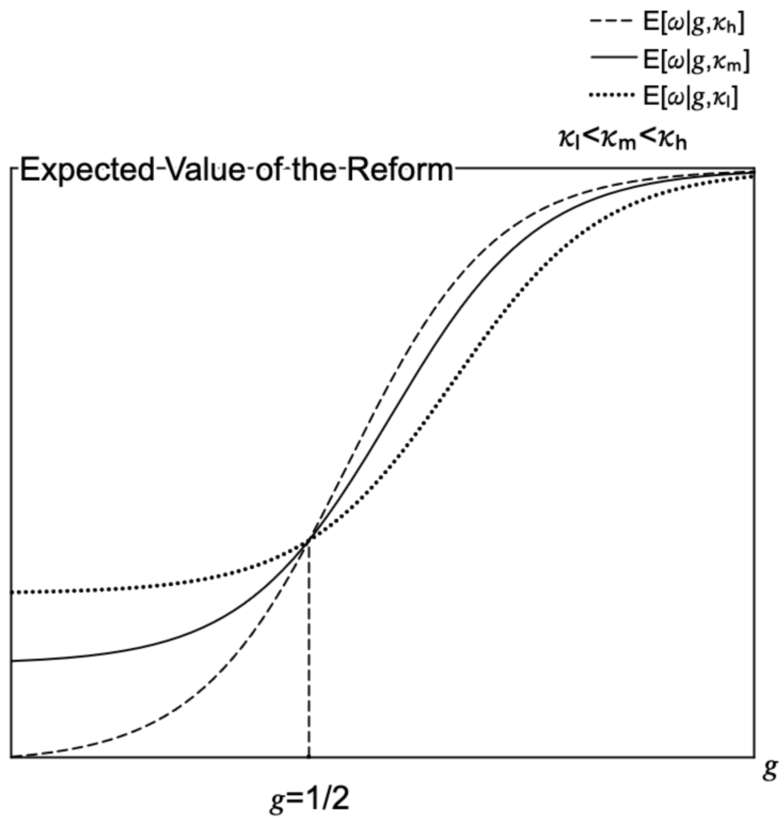
Figure 2: Likelihood Comparison



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 3 shows how this *inference effect of sabotage* 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, in the face of sabotage, he is inclined to ascribe high-quality services (high g) to mere luck rather than the reform's success, thus becoming more stringent with the incumbent when observing high-quality services. We further unpack the mechanisms for these two countervailing effects in Appendix C.

Figure 3: Sabotage's Countervailing Effects on Voter Inference



4.3 Bureaucrats' Incentives to Sabotage

We now study the bureaucrats' optimal strategy given the voter's behavior.

The bureaucrats' incentives to sabotage in $t = 1$ come from (i) their desire to avoid the reform in $t = 2$ and (ii) their ability to influence the election outcome by affecting government outcome, g. Therefore, they do not sabotage unless they can influence the election outcome by affecting the government outcome. For instance, they do not sabotage any policy in $t = 2$ because there is no election afterward. Similarly, they do not sabotage the status quo, since the voter's election decision is independent of g.

Bureaucrats experience disutility $-\kappa < 0$ in $t = 2$ if the incumbent introduces the reform and wins the election, and 0 if the opponent wins. Thus, the benefit of sabotage is given by the reform disutility κ times the gap between the reelection probabilities of the reforming incumbent with and without sabotage, and the bureaucrats sabotage if and only if this benefit is larger than its cost c. Namely, $s^*(\kappa') = 1$ if and only if

$$\kappa \times \Big(\Pr[reelection(a=1)|s=1] - \Pr[reelection(a=1)|s=0] \Big) \geq c. \tag{4}$$

Suppose the voter reelects the incumbent if and only if he observes $g \geq g'$, so

$$\Pr[reelection(a=1)|s=1] = \Pr[g \geq g'|a=1, s=1] = 1 - H(g')$$

with sabotage $s = 1$ and

$$\Pr[reelection(a=1)|s=0] = \Pr[g \geq g'|a=1, s=0] = \begin{cases} 1 - H(g') & \text{if } \omega = 0 \\ 1 - H(g'-1) & \text{if } \omega = 1 \end{cases}$$

without sabotage, $s = 0$. Then equation (4) can be rewritten as

$$\kappa \underbrace{\Big(1 - H(g') - 1 + H(g')\Big)}_{=0} \geq c \tag{5}$$

if $\omega = 0$ and

$$\kappa \Big(H(g') - H(g'-1) \Big) \geq c \tag{6}$$

if $\omega = 1$. Notice that equation (5) never holds as the reelection probability is constant with respect to sabotage, so the bureaucrats do not sabotage the reform that does not work: $s^*(\omega = 0) = 0$. In contrast, the bureaucrats can benefit from sabotaging the reform that actually works, as the reelection probability changes by $H(g') - H(g'-1)$. Therefore, the bureaucrats sabotage the reform if and only if it works ($\omega = 1$) and

$$\kappa > \hat{\kappa}^*(g') := \frac{c}{H(g') - H(g'-1)}. \tag{7}$$

4.4 Incumbent's Decision

The incumbent's decision over the reform depends on its implications for her reelection prospects. 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 given r. Recall that the probability that the incumbent wins the election with the status quo is $1/2$. Thus, the incumbent chooses the reform if and only if

$$\varepsilon + \Pr[reelection|a=1, r](1+\varepsilon) \geq \Pr[reelection|a=0] = 1/2 \tag{8}$$

To focus solely on the incumbent's reelection incentives, we assume that $\varepsilon \to^{+} 0$ in the discussion from this point on, emphasizing the electoral incentives behind the incumbent's policy-making.[14] Then, inequality (8) is

$$\Pr[reelection|a = 1, r] \geq 1/2. \tag{9}$$

Suppose that the voter reelects the incumbent if and only if he observes $g \geq g'$ and the bureaucrats sabotage the reform if and only if $\omega = 1$ and $\kappa > \hat{\kappa}^*(g')$. Then, $\Pr[reelection|a = 1, r]$ is given by

$$\begin{aligned}&\rho(r)\left[\hat{\kappa}^*(g')\Big(1 - H(g'-1)\Big) + [1 - \hat{\kappa}^*(g')]\Big(1 - H(g')\Big)\right] + [1 - \rho(r)]\Big(1 - H(g')\Big)\\ &= \rho(r)\hat{\kappa}^*(g')\Big(H(g') - H(g'-1)\Big) + 1 - H(g').\end{aligned}$$

Then, given $\hat{\kappa}^*(g')$ in equation (7), we can rewrite inequality (9) to yield the incumbent's decision rule:

$$a^*(r; g') := \mathbf{1}\left\{\rho(r) \geq \frac{H(g') - 1/2}{c}\right\}. \tag{10}$$

Given the properties of $\rho(r) = \Pr[\omega = 1|r]$, we can define

$$\hat{r}^*(g') := \rho^{-1}\left(\frac{H(g') - 1/2}{c}\right). \tag{11}$$

4.5 Equilibrium Construction

By bringing together all three actors, we can characterize the equilibrium. Namely, the voter's equilibrium threshold g^* defines the equilibrium threshold value of the bureaucrats,

$$\kappa^* := \hat{\kappa}^*(g^*) = \frac{c}{H(g^*) - H(g^* - 1)}$$

and, that of the incumbent,

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

[14] As we only consider the pure-strategy equilibrium with a unique threshold for each player, the probability measure for each player's best response in behavioral strategy (i.e., the unique threshold value) is a continuous function of the other players, unlike in a completely mixed strategy equilibrium (Echenique and Edlin, 2003). Consequently, a slight perturbation in ε creates a proportionate, slight change in threshold values (Milgrom and Roberts, 1994). Namely, $\varepsilon > 0$ does not affect the equilibrium qualitatively other than slightly increasing the incumbent's incentives to choose the reform.

In turn, we can attain g^* by plugging these two back into λ in equality (2). Specifically, for $\kappa^* = \hat{\kappa}^*(g^*)$ and $r^* = \hat{r}^*(g^*)$, g^* satsifies

$$\begin{aligned}\lambda(g^*, \hat{\kappa}^*(g^*), \hat{r}^*(g^*)) &= \frac{1 - F(\hat{r}^*(g^*)|0)}{1 - F(\hat{r}^*(g^*)|1)} \frac{h(g^*)}{h(g^*) + \hat{\kappa}^*(g^*)\Big(h(g^* - 1) - h(g^*)\Big)} \\ &= \frac{1 - F(\hat{r}^*(g^*)|0)}{1 - F(\hat{r}^*(g^*)|1)} \frac{h(g^*)}{h(g^*) + c\frac{h(g^*-1)-h(g)}{H(g^*)-H(g^*-1)}} := \Lambda(g^*) = \frac{1-q}{q}.\end{aligned} \tag{12}$$

After endogenizing other players' strategies, the voter's posterior belief that the reform works, $\dfrac{1}{1+\lambda(g)}$, is still increasing in g:

Lemma 2 *$\Lambda(g)$ is monotonically decreasing in g.*

By monotonicity of $\Lambda(\cdot)$, we can define its inverse, $\Lambda^{-1}(\cdot)$, as a decreasing real function and define g^* as

$$g^* := \Lambda^{-1}\left(\frac{1-q}{q}\right) \tag{13}$$

Since g^* is uniquely defined by the monotonicity of Λ^{-1}, κ^* and r^* are also uniquely defined.

Proposition 1 *There exists a unique pure strategy equilibrium with a unique set of threshold values $\{g^*, \kappa^*, r^*\}$ such that*

- *The voter reelects the incumbent if and only if he observes a high enough government outcome:*

$$\mathbf{1}\{reelection^*|g\} = \mathbf{1}\{g \geq g^*\}.$$

- *The bureaucrats sabotage if and only if*

  (i) The incumbent introduces the reform, $a = 1$, ***and***

  (ii) The introduced reform works, $\omega = 1$, ***and***

  (iii) The disutility from it is high enough, $\kappa > \kappa^*$ in $t = 1$

  and do not sabotage in $t = 2$:

$$\begin{aligned}&s^*(a, \omega, \kappa) = a \times \omega \times \mathbf{1}\{\kappa > \kappa^*\} \\ &\tilde{s}^* = 0\end{aligned}$$

- *The incumbent introduces the reform if and only if she observes a high enough signal in $t = 1$ and continues her policy in $t = 2$:*

$$a^*(r) = \mathbf{1}\{r \geq r^*\}$$
$$\tilde{a}^* = a$$

- *The opponent chooses the status quo if she wins the election.*

5 Comparative Statics

We now consider how the exogenous parameters q and c affect equilibrium outcomes. Importantly, by comparing equilibrium outcomes across various levels of q and c, we can describe in detail when bureaucrats sabotage reforms and how sabotage leads to different types of policy inefficiencies.[15]

5.1 Bureaucrats' Equilibrium Behavior

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 $c < \bar{c}$.*[16]

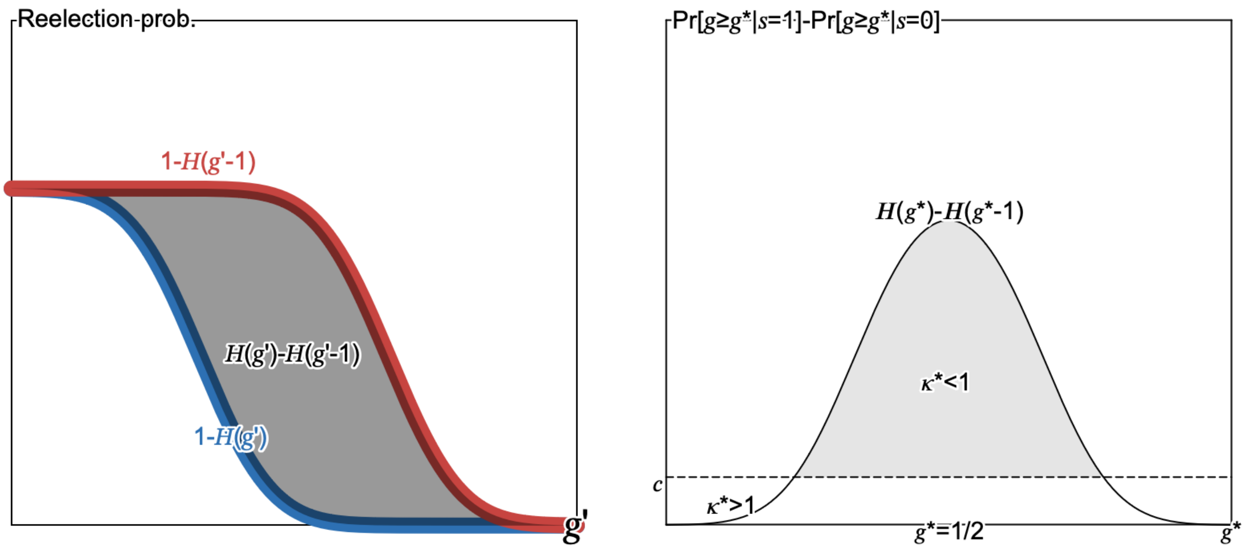
Figure 4 provides intuition for why sabotage incentives are non-monotonic with respect to q. Recall that sabotage reduces the probability that the voter observes a positive enough signal to reelect a reforming incumbent. Without sabotage, the probability that the voter observes $g \geq g^*$ is $1-H(g^*-1)$ (the red line). Sabotage decreases this probability to $1-H(g^*)$ (the blue line). 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.

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 $g^* < 1/2$, then the voter may

[15] Note that the proofs largely compare cases with and without sabotage instead of marginal effects of c. See Appendix A.3.1 for a detailed discussion of the benchmark case without sabotage.

[16] As we show in Appendix A.3.2, the effect of c on k^* is only unambiguous for $g > g^{\dagger\dagger}$, where k^* increases with c.

Figure 4: 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.

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 $g^* = 1/2$.

As panel (b) of Figure 4 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^*.

5.2 Voter's Equilibrium Behavior

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. *There exist* $c^\dagger(0,1) \in$ *and* $q^\dagger \in (1/2, 1)$ *such that*

   - $g^*(q,c)$ *is weakly increasing in* c *if* $c < c^\dagger$ *or* $q < q^\dagger$.
   - $g^*(q,c)$ *is weakly decreasing in* c *if* $c > c^\dagger$ *and* $q > q^\dagger$.

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 $\Lambda(g)$ into two components:

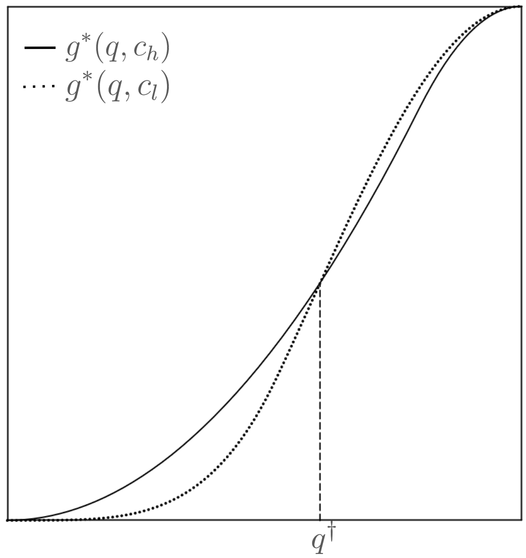
$$\Lambda(g^*) = \underbrace{\frac{1 - F(\hat{r}^*(g^*)|0)}{1 - F(\hat{r}^*(g^*)|1)}}_{\text{Reservation Effect, } \downarrow g^*} \times \underbrace{\frac{h(g^*)}{h(g^*) + c\frac{h(g^*-1)-h(g^*)}{H(g^*)-H(g^*-1)}}}_{\text{Inference Effect, } \uparrow g^* \text{ iff } g^*>1/2}.$$

The second component captures the inference effect of sabotage. As discussed in section 4.2, 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 $\Lambda(g)$.

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. We call this the *direct effect of sabotage*. 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*. 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. For small costs $c < c^\dagger$, the reservation effect always dominates the inference effect, because incumbents are extremely sensitive to changes in the costs of sabotage. For high enough costs $c > c^\dagger$, the reservation effect amplifies the inference effect for low q but weakens the inference effect for high q. As Figure 5 illustrates, lower costs of sabotage, therefore, only induce the voter to be more strict with the incumbent for sufficiently high $q > q^\dagger$ and sufficiently high $c > c^\dagger$.

[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.

Figure 5: Sabotage's Equilibrium Effect on Voting Decision, $c > c^\dagger$



The voter's equilibrium cutoff g^* with high $c_h > c^\dagger$ and low $c_l > c^\dagger$.

5.3 Incumbent's Equilibrium Behavior: Sabotage & Policy-Making

To facilitate the discussion of policy inefficiencies induced by sabotage, consider the following normative benchmark illustrated in Figure 6. 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.*

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

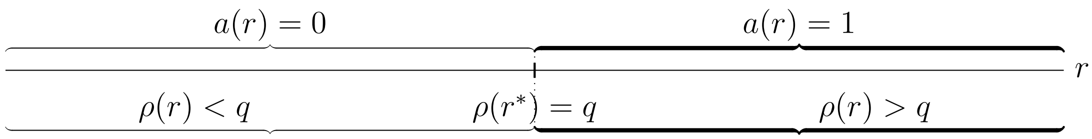
1. *As the status quo's value increases, the incumbent requires a higher signal to introduce reform; $r^*(q, c)$ is weakly increasing in q.*

2. *If $c > c^\dagger$, $r^*(q, c)$ is weakly increasing in c if and only if $q < q^{\dagger\dagger} < 1/2$.*[18]

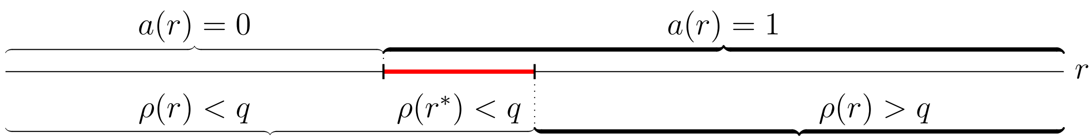
Given the voter's tendency to become more stringent with large q, the probability that a reforming incumbent gets reelected decreases as q increases. In response, the incumbent becomes more conservative and requires a higher signal r to introduce reform.

[18]For brevity, we focus on the more interesting case where $c > c^\dagger$, since it nests the mechanisms for cases where $c < c^\dagger$.

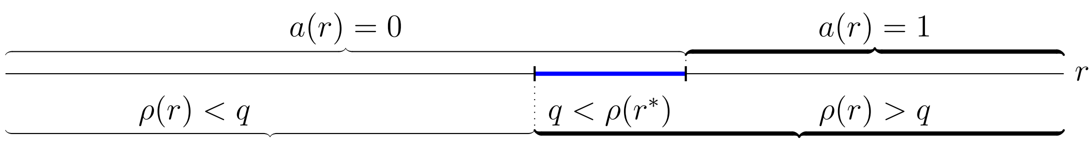
Figure 6: Normative Criterion and Policy Inefficiencies



(a) Efficient policy-making



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



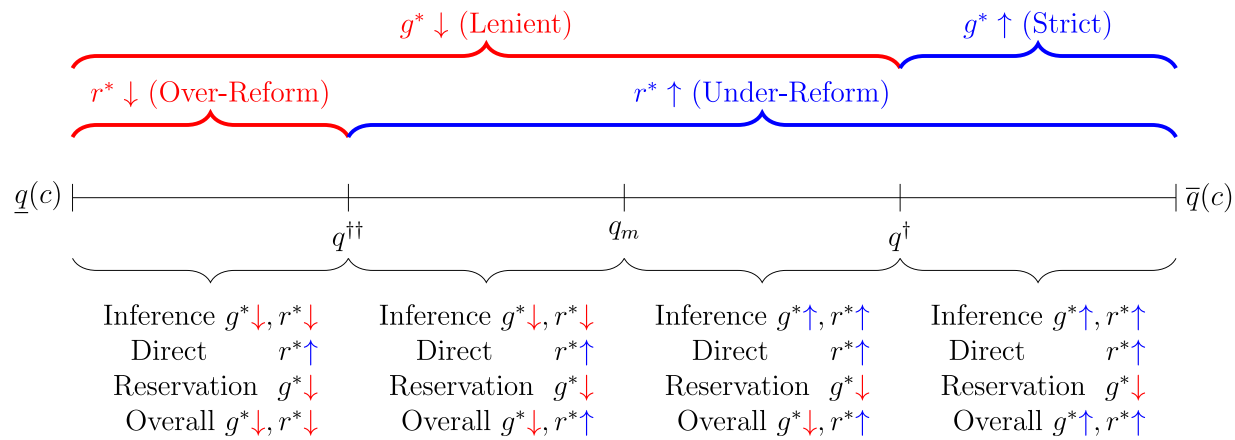
(c) Inefficiency due to **under**-reforming

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.

What are the mechanisms for these differential effects of sabotage in equilibrium? Figure 7 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 ($\kappa \to r^* \to g^*$).

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 and size of the inference effect depend on q. Consider the extreme cases first. For the highest levels of q ($q^\dagger < q < \overline{q}(c)$), where voters have little trust in the success of the reform, the inference effect makes the voter more strict by a large margin and dominates the reservation effect. Consequently, the incumbent is overly conservative with sabotage. The opposite holds if voters have high ex-ante trust in the reform. For the lowest levels of q ($\underline{q}(c) < q < q^{\dagger\dagger}$), sabotage makes the voter lenient through a large inference

Figure 7: Mechanisms of Equilibrium Effects of Sabotage



The figure illustrates the mechanisms for the effect of easier sabotage (decreases in c) on voter inference and incumbent behavior.

effect, which feeds into the incumbent's electoral incentives and dominates the direct and reservation effects of sabotage. In turn, sabotage leads to over-reform.

For more moderate levels of q, however, the direct and reservation effects are strong because voters who are torn between the status quo and the reform are more susceptible to changes in government services. For intermediary low levels of q ($q^{\dagger\dagger} < q < q_m$), 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, even though the voter is more forgiving. For intermediary high levels of q ($q_m < q < q^{\dagger}$), the inference effect makes the voter more strict, but it is not large enough to cancel out the reservation effect. Again, sabotage leads to under-reform.

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]

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.

[19] Figure A3 shows the equilibrium predictions given these dynamics of over- and under-reforming.

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 opponent 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 opponent 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

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

threatening a change in the quality of law 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

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

federal bureaucracy was working to undermine his administration, Trump weaponized fears 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. Our model illustrates 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 politics 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.

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 sabotage ($1-\kappa^*$), the probability of reform ($1-r^*$), and the probability of reelection ($1-g^*$). Similarly, scholars could use surveys to empirically evaluate the impact of bureaucratic sabotage (i.e., variation in c) on voters' perceptions of reform merit ($\mathrm{E}[\omega|g,c]$), conditional on the realized government quality (g). Our results suggest that sabotage dampens voters' preference for reform for low government quality, while it strengthens their perceptions of reform merits for high government quality (see Lemma A1.3 and Figure 3).

When selecting empirical cases for such analyses, scholars want to pay close attention to two 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 government outcomes. 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/.

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

Awad, Emiel, Varun Karekurve-Ramachandra and Lawrence Rothenberg. 2023. "Politicians, Bureaucrats, and the Battle for Credit." Working Paper.
URL: *https://osf.io/preprints/socarxiv/ajrey/*

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.

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.

Bils, Peter and Federica Izzo. 2023. "Ideological Infection." *Working Paper* .

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 Wioletta Dziuda. 2023. "Partisan Traps." NBER Working Paper No. w31827. https://www.nber.org/system/files/working_papers/w31827/w31827.pdf.

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.

Delgado-Vega, Alvaro, Wioletta Dziuda and Antoine Loeper. 2023. "Politics of Repeals." Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4632868.

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.

Echenique, Federico and Aaron Edlin. 2003. "Mixed equilibria are unstable in games of strategic complements." *Journal of Economic Theory* .

Echenique, Federico and Aaron Edlin. 2004. "Mixed equilibria are unstable in games of strategic complements." *Journal of Economic Theory* 118(1):61–79.

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/.

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 and Gergely Ujhelyi. 2021. "Don't hatch the messenger? On the desirability of restricting the political activity of bureaucrats." *Journal of Theoretical Politics* 33(1):95–139.

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.

Fox, Justin and Stuart V. Jordan. 2011. "Delegation and Accountability." *https://doi.org/10.1017/S0022381611000491* 73(3):831–844.
URL: *https://www.journals.uchicago.edu/doi/10.1017/S0022381611000491*

Gersen, Jacob E. and Matthew C. Stephenson. 2014. "Over-Accountability." *Journal of Legal Analysis* .

Gieczewski, Germán and Christopher Li. 2022. "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.

Izzo, Federica, Gregory J. Martin and Steven Callander. 2023. "Ideological Competition." *American Journal of Political Science* .

Keeler, John. 1993. "Opening the Window for Reform: Mandates, Crises, and Extraordinary Policy-Making." *Comparative Political Studies* .

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*

Martin, Lucy and Pia J. Raffler. 2021. "Fault Lines: The Effects of Bureaucratic Power on Electoral Accountability." *American Journal of Political Science* 65(1):210–224.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12530*

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.

Milgrom, Paul. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics* .

Milgrom, Paul and John Roberts. 1994. "Comparing Equilibria." *American Economic Review* .

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. 2015. "Vested Interests and Political Institutions." *Political Science Quarterly* 130(2):277–318.
URL: *http://www.jstor.org/stable/43828574*

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.

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.

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 Voter Inference

Proof for Lemma 1. Lemma 1-1.

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

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

Since $h(g-1) = h(g)$ if and only if $g = 1/2$, $\lambda = 0$ if and only if $g = 1/2$. ■

Lemma 1-2.

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

Observe that this is not positive if and only if

$$\frac{f(r'|0)}{1-F(r'|0)} \geq \frac{f(r'|1)}{1-F(r'|1)}. \tag{14}$$

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

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

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 (14). ■

Lemma 1-3. Observe that

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

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 $g < 1/2$. Therefore, $\lambda(g, \kappa', r')$ is increasing in κ if and only if $q < 1/2$. ■ ■

A.2 Endogenizing Sabotage and Inference

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 $g < 1/2$. ■

Proof for Proposition 1. From equation (7), 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 λ to get

$$\lambda(g,r') := \lambda(g,r',\kappa^*) = \frac{1-F(r'|0)}{1-F(r'|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}. \tag{15}$$

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

Lemma A1 *For $\lambda(g, r') := \lambda(g, r', \kappa^*)$*

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

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

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 $sign\frac{\partial L}{\partial r'} = sign\frac{\partial \lambda}{\partial r'}$.

To see $sign\frac{\partial L}{\partial g} = sign\frac{\partial \lambda}{\partial g}$, 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

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

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 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.

$\lambda(g, r')$ depends on c as well. $\lambda(g, r')$ 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, $\frac{1-q}{q}$ decreases. Since $\lambda(g^*, r')$ is monotonically decreasing in g^*, g^* is increasing in q.

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

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

Proof for Lemma A2. $\frac{\partial\lambda(g,\hat{r}^*(g))}{\partial g} = \frac{\partial}{\partial g}\Big(\frac{1-F(\hat{r}^*(g)|0)}{1-F(\hat{r}^*(g)|1)}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}\Big)$ 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, $\frac{\partial\lambda(g,\hat{r}^*(g))}{\partial g}\leq 0$. ■ By the property of the monotone real mapping (the intermediate value theorem), there exists unique $g^*(q)$ such that $\lambda(g^*(q),\hat{r}^*(g^*(q)))=\frac{1-q}{q}$. ■

A.3 Comparative Statics

A.3.1 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

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

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*

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

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

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

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

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

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

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 $\lambda_B(g, \hat{r}_B^*(g))$ is decreasing in g, $g_B^*(q)$ such that $\lambda_B(g_B^*(q), \hat{r}_B^*(g_B^*(q))) = \frac{1-q}{q}$ 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 $q_{BE}^\dagger \in (1/2, 1)$ such that $\rho(r_B^*(q_{BE}^\dagger)) = q_{BE}^\dagger$ and $\rho(r_B^*(q)) < q$ if $q \in (0, q_{BE}^\dagger)$ and $\rho(r_B^*(q)) > q$ if $q \in (q_{BE}^\dagger, 1)$.*

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

Proof for Proposition A3. Observe that

$$\frac{\partial^2}{\partial g^2}\frac{H(g)-1/2}{H(g)-H(g-1)}$$
$$=\frac{1}{2[H(g)-H(g-1)]^2}\left(\begin{array}{l}[2H(g)-1]\Big([H(g)-H(g-1)]h''(g-1)-2[h(g)-h(g-1)]h(g-1)\Big)\\ +[1-2H(g-1)]\Big([H(g)-H(g-1)]h''(g)-2[h(g)-h(g-1)]h(g)\Big)\end{array}\right).$$

which is 0 at $g=1/2$ by the symmetry at the point. Notice $[2H(g)-1]\Big([H(g)-H(g-1)]h''(g-1)-2[h(g)-h(g-1)]h(g-1)\Big)$ is a point reflection of $[1-2H(g-1)]\Big([H(g)-H(g-1)]h''(g)-2[h(g)-h(g-1)]h(g)\Big)$. First, $[2H(g)-1]>0$ and $[1-2H(g-1)]>0$ are symmetric around $g=1/2$. Second, $[H(g)-H(g-1)]h''(g-1)$ is a point reflect of $[H(g)-H(g-1)]h''(g)$ around $(1/2,0)$ since $[H(g)-H(g-1)]$ is symmetric around $1/2$ and $h''(g-1)$ is a point reflection of $h''(g)$ (derivatives of symmetric functions are point reflections of each other). Furthermore, since the multiplication of an even function[2] and an odd function[3] is an odd function and two even functions are an even function. Thus, $[2H(g)-1]\Big([H(g)-H(g-1)]h''(g-1)-2[h(g)-h(g-1)]h(g-1)\Big)$ is an odd function around $1/2$. By a similar logic, $[1-2H(g-1)]\Big([H(g)-H(g-1)]h''(g)-2[h(g)-h(g-1)]h(g)\Big)$ is also an odd function around $1/2$. Let $\zeta(g):=[2H(g)-1]\Big([H(g)-H(g-1)]h''(g-1)-2[h(g)-h(g-1)]h(g-1)\Big)$. Notice that $-\zeta(1-g)=[1-2H(g-1)]\Big([H(g)-H(g-1)]h''(g)-2[h(g)-h(g-1)]h(g)\Big)$. Thus,

$$\frac{\partial^2}{\partial g^2}\frac{H(g)-1/2}{H(g)-H(g-1)}\iff\zeta(g)\geq\zeta(1-g)\iff g\geq 1/2.$$

This implies that $\frac{H(g)-1/2}{H(g)-H(g-1)}$ is a reverse S-shaped function (concave if $g<1/2$ and convex if $g>1/2$) 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 $\rho(r)\geq q$.

[2] A real function $\psi(z):\mathbb{R}\to\mathbb{R}$ is an even function if $\psi(-z)=\psi(z)$.
[3] A real function $\psi(z):\mathbb{R}\to\mathbb{R}$ is an odd function if $\psi(-z)=-\psi(z)$.

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

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

Notice that F and ρ^{-1} are increasing, so $\frac{\partial}{\partial q}F(\rho^{-1}(q)|\omega) \geq 0$. 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 a 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.[4] 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

[4]For a real function φ, the derivative of its inverse is $\frac{1}{\varphi'}$. A derivative of a reversed S-shaped function is U-shaped with a unique minimum, whose reciprocal is single-peaked with a unique maximum. If a derivative of a real function is single-peaked, then the function is S-shaped.

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 $\overline{P} := 1 - \frac{H(g^*_B(1/2))+H(g^*_B(1/2)-1)}{2} > 1/2$. 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.3.2 Effect of Status Quo and Sabotage in Equilibrium

Proof for Propositions 3, 2, and 4.

Effects of the Status Quo Value

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

Recall that $\hat{\kappa}^*(g')$ is a U-shaped function of g' that attains its minimum at $g^* = 1/2$ and g^* is increasing in q: $\frac{\partial}{\partial q}\kappa^* = \underbrace{\frac{\partial g^*}{\partial q}}_{\text{increasing}} \underbrace{\frac{\partial}{\partial g^*}\hat{\kappa}^*(g^*)}_{\text{U-shaped}}$.

Sabotage's Effect on Voter Inference

We want to show that there exists $q^\dagger$ such that $\hat{r}^*(g) = \hat{r}^*_B(g)$ if $q = q^\dagger$. This implies that

$$\begin{aligned}
&\Lambda_B(g^*) = \Lambda(g^*)\\
&\iff \log \Lambda_B(g^*) = \log \Lambda(g^*)\\
&\iff \log \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} - \log \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)} = \log \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)}
\end{aligned}$$

for $q = q^\dagger$. We limit our focus for $g \in (1/2, \overline{g}(c))$ where $c = H(\overline{g}(c)) - H(\overline{g}(c) - 1)$ as the inference effects have the direction as the reservation effects if $g^* < 1/2$ and the bureaucrats do not sabotage if $g \geq \overline{g}(c)$.

Lemma A3
$$\log \frac{1 - F(\hat{r}^*_B(g)|0)}{1 - F(\hat{r}^*_B(g)|1)} - \log \frac{1 - F(\hat{r}^*(g)|0)}{1 - F(\hat{r}^*(g)|1)}$$

is positive and single-peaked in $g \in [1/2, \overline{g})$.

Recall $\rho(\hat{r}^*(g)) = \frac{H(g)-1/2}{c}$ is concave for $g \geq 0$ and $\rho(\hat{r}^*_B(g)) = \frac{H(g)-1/2}{H(g)-H(g-1)} \geq 0$ is convex if and only if $g \geq 1/2$. Therefore, $\rho(\hat{r}^*(g)) - \rho(\hat{r}^*_B(g))$ is concave in $g \geq 1/2$.

Concavity of $\rho(\hat{r}^*(g)) - \rho(\hat{r}^*_B(g))$ implies that there exist $\delta_l(\nu)$ and $\delta_h(\nu)$ such that $1/2 \leq \delta_l(\nu) \leq \delta_h(\nu) \leq \overline{g}(c)$ and

$$\rho(\hat{r}^*(g)) - \rho(\hat{r}^*_B(g)) \leq \nu \iff \rho(\hat{r}^*(g)) \leq \rho(\hat{r}^*_B(g)) + \nu_1 \tag{16}$$

if $g \in (1/2, \delta_l(\nu)] \cup [\delta_h(\nu), \overline{g}(c))$ for any $\nu > 0$. Let $\underline{\delta}(\nu) := \max \delta_l(\nu)$ and $\overline{\delta}(\nu) := \min \delta_h(\nu)$. By concavity, $\underline{\delta}(\nu)$ is increasing and $\overline{\delta}(\nu)$ is decreasing in ν.

Let $LR(\cdot) := \log \frac{1-F(\rho^{-1}(\cdot)|0)}{1-F(\rho^{-1}(\cdot)|1)}$. Since ρ (increasing), $\frac{1-F(\cdot|0)}{1-F(\cdot|1)}$ (decreasing), and log (increasing) are differentiable, LR is decreasing function. Then we can rewrite inequality (16) as

$$LR(\rho(\hat{r}^*(g)) \geq LR(\rho(\hat{r}^*_B(g)) + \nu_1). \tag{17}$$

Let $\nu_2 := LR(\rho(\hat{r}^*_B(g))) - LR(\rho(\hat{r}^*_B(g)) + \nu_1) > 0$. Then inequality (17) can be rewritten as

$$LR(\rho(\hat{r}^*(g))) \geq LR(\rho(\hat{r}^*_B(g))) - \nu_2 \iff LR(\rho(\hat{r}^*_B(g))) - LR(\rho(\hat{r}^*(g))) \leq \nu_2. \tag{18}$$

Thus, inequality (18) holds if $g \in (1/2, \delta_l(\nu)] \cup [\delta_h(\nu), \overline{g}(c))$ for any $\nu_2 > 0$ ($LR(0) = \log 1 = 0$).

Notice that we can still use $\underline{\delta}(\nu)$ and $\overline{\delta}(\nu)$ for inequality (18). Since ν_2 is increasing in ν, $\underline{\delta}(\nu_2)$ is increasing and $\overline{\delta}(\nu_2)$ is decreasing in ν_2. Therefore, for an arbitrary $\nu_2 > 0$,

$LR(\rho(\hat{r}_B^*(g))) - LR(\rho(\hat{r}^*(g))) = \nu_2$ at most twice, at $g = \underline{\delta}(\nu_2)$ and $g = \overline{\delta}(\nu_2) > \underline{\delta}(\nu_2)$. This implies single-peakedness. To see this, pick arbitrary ν_3, ν_4, ν_5 such that $0 < \nu_3 < \nu_4 < \nu_5$, there exists at least one g such that $LR(\rho(\hat{r}_B^*(g))) - LR(\rho(\hat{r}^*(g))) = \nu_5$, and $\nu_3 \geq LR(\rho(\hat{r}_B^*(1/2))) - LR(\rho(\hat{r}^*(1/2)))$. Then,

$$\underline{\delta}(\nu_3) < \underline{\delta}(\nu_4) < \underline{\delta}(\nu_5) \leq \overline{\delta}(\nu_5) \leq \overline{\delta}(\nu_4) < \overline{\delta}(\nu_3).$$

Recall that LR is a differentiable function. Let $g_{peak,LR}$ denote the peak of $LR(\rho(\hat{r}_B^*(g))) - LR(\rho(\hat{r}^*(g)))$.

Lemma A4

$$\log \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)}$$

is positive and single-peaked in $g \in (1/2, \overline{g})$.

Observe

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

Recall that

$$\frac{\partial}{\partial g}\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)} > 0.$$

Thus,

$$\frac{\partial^2}{\partial g \partial c} \frac{h(g)}{h(g) + c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \leq 0 \iff \frac{1}{c} \leq \frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}.$$

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 $g_{cp}(c) > 1/2$ such that

$$\frac{1}{c} = \frac{[h(g_{cp}(c)-1)/h(g_{cp}(c))]-1}{H(g_{cp}(c)) - H(g_{cp}(c)-1)} \tag{19}$$

and $\frac{\partial}{\partial g}\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ is increasing in c if and only if $g < g_{cp}(c)$.

Therefore, $\frac{h(g)}{h(g-1)}$ increases faster than $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ as g increases if $g < g_{cp}(c)$ and this makes $\log\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log\frac{h(g)}{h(g-1)}$ increase in g. Conversely, if $g < g_{cp}(c)$, $\frac{h(g)}{h(g-1)}$ increases slower than $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ as g increases if $g > g_{cp}(c)$ and this makes $\log\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} -$

$\log \frac{h(g)}{h(g-1)}$ increase slow or decrease in g. We know that $\log \frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} - \log \frac{h(g)}{h(g-1)} = 0$ at $g = \overline{g}$, so there exists its unique peak $\geq g_{cp}$.

Notice that at $g = g_{cp}$, $\log \frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} = \log \frac{1}{1+c\frac{[h(g-1)/h(g)]-1}{H(g)-H(g-1)}} = \log \frac{1}{2}$.

Lemma A5 *If*

$$\log \frac{1 - F(\hat{r}_B^*(g_{cp})|0)}{1 - F(\hat{r}_B^*(g_{cp})|1)} - \log \frac{1 - F(\hat{r}^*(g_{cp})|0)}{1 - F(\hat{r}^*(g_{cp})|1)} < \log \frac{1}{2}\frac{h(g_{cp}-1)}{h(g_{cp})} \tag{20}$$

then there exists a $g^\dagger \in (1/2, g_{cp})$ *such that* $g_B^* \leq g^*$ *if and only if* $g_B^* \geq g^\dagger$.

The argument holds trivially by the intermediate value theorem (Recall the LHS is strictly positive and the RHS is 0 at $g = 1/2$).

Lemma A6 *There exists* $\underline{c}$ *such that inequality* (20) *does not hold if* $c < \underline{c}$.

Notice that g_{cp} is decreasing in c (see equation (19).). So as c decreases, $\log \frac{1}{2}\frac{h(g_{cp}-1)}{h(g_{cp})}$ increases by the MLRP. However, its increase is bounded as there still exists $c_{cp}(c) < \overline{g}$ for $c > 0$.

In contrast, that does not need to be case for $\log \frac{1-F(\hat{r}_B^*(g_{cp})|0)}{1-F(\hat{r}_B^*(g_{cp})|1)} - \log \frac{1-F(\hat{r}^*(g_{cp})|0)}{1-F(\hat{r}^*(g_{cp})|1)}$. Holding $g = g_{cp}$, $\log \frac{1-F(\hat{r}_B^*(g_{cp})|0)}{1-F(\hat{r}_B^*(g_{cp})|1)} - \log \frac{1-F(\hat{r}^*(g_{cp})|0)}{1-F(\hat{r}^*(g_{cp})|1)}$ monotonically decreasing in c as the gap between $\rho(\hat{r}_B^*(g_{cp})) = \frac{H(g_{cp})-1/2}{H(g_{cp})-H(g_{cp}-1/2)}$ and $\rho(\hat{r}^*(g_{cp})) = \frac{H(g_{cp})-1/2}{c}$ increases. Then, notice that if $c \leq H(g_{cp}) - 1/2$, then $\rho(\hat{r}^*(g_{cp})) = 1$, which means that $\log \frac{1-F(\hat{r}^*(g_{cp})|0)}{1-F(\hat{r}^*(g_{cp})|1)} \to -\infty$. Thus, there exists $\underline{c} > H(g_{cp}) - 1/2$ such that inequality (20) does not hold if $c < \underline{c}$.

Lemma A7 *There exists* $c^\dagger < 1$ *such that the inequality* (20) *holds.*

Consider $\overline{c^\dagger} < 1$ such that $c > H(g_{cp}) - H(g_{cp} - 1)$ for $c > \overline{c^\dagger}$, so the LHS is zero, but the RHS is positive for $c \in (\overline{c^\dagger}, 1)$. Then for $c \geq \overline{c^\dagger}$, the inequality (20) holds trivially. Also, there exists an open neighborhood around $\overline{c^\dagger}$ such that the inequality holds. Notice that $H(g) - H(g-1) < 1$ by the property of a CDF, so $\overline{c^\dagger}$ exists.

Thus, there exists $c^\dagger \in (\underline{c}, \overline{c^\dagger})$ such that the inequality (20) holds if $c > c^\dagger$.

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 if positive. 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 $q^\dagger_{SE} := q^*_{NS}(g^*_{SE})$.

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

Notice that, at $q = q^\dagger$, so $g^* = g^*_B = g^\dagger$,

$$\rho(\hat{r}^*(g^\dagger)) = \frac{H(g^\dagger) - 1/2}{c} < \rho(\hat{r}^*_B(g^\dagger)) = \frac{H(g^\dagger) - 1/2}{H(g^\dagger) - H(g^\dagger - 1)}.$$

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

$$\frac{H\left(g^*(q^{\dagger\dagger})\right) - 1/2}{c} = \frac{H\left(g^*_B(q^{\dagger\dagger})\right) - 1/2}{H\left(g^*_B(q^{\dagger\dagger})\right) - H\left(g^*_B(q^{\dagger\dagger}) - 1\right)} \tag{21}$$

so $r^*_B(q^{\dagger\dagger}) = r^*(q^{\dagger\dagger})$. Notice that $g^*(q^{\dagger\dagger}) > g^*_B(q^{\dagger\dagger})$.

Holding $g^*_B(q^{\dagger\dagger}) = g^{\dagger\dagger}_B$ 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 $c < H(g) - H(g-1)$. Thus, $g^*(q^{\dagger\dagger}) > g^{\dagger\dagger}_B$.

Suppose that $g^*(q^{\dagger\dagger}) > g^{\dagger\dagger}(g_B^{\dagger\dagger}) > 1/2$. Recall $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} > \frac{h(g)}{h(g-1)}$ iff $g > 1/2$. Since $\frac{h(g)}{h(g)+c\frac{h(g-1)-h(g)}{H(g)-H(g-1)}}$ is decreasing in g,

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

Notice that this leads to a contradiction since

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

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

$r^*(g^{\dagger\dagger}) = r_B^*(g^{\dagger\dagger})$.

Notice that $q_{SE}^{\dagger} > 1/2$ 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

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

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 $g \le 1/2$, 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 $\Lambda_i(g_i^*) = \frac{1-q}{q}$, $g_h^* > g_l^*$ if $g \le 1/2$ since $\Lambda_i(g)$ is decreasing and $\Lambda_h(g) > \Lambda_l(g)$.

We can define $g_{cp}(c_l) > 1/2$ such that $\frac{\partial^2}{\partial g \partial c_l}\frac{h(g)}{h(g)+c_l\frac{h(g-1)-h(g)}{H(g)-H(g-1)}} \ge 0$ if and only if $g \le g_{cp}(c_l)$. Thus, there exists $g^\dagger(c_l) > g_{cp}(c_l)$ such that $\Lambda_h(g) \le \Lambda_l(g)$ $g_h^* \le g_l^*$ if and only if $g \ge g^\dagger(c_l)$.

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,

$$\frac{\partial \log \kappa^*}{\partial c} = \frac{1}{c} - \frac{\partial}{\partial c} \log H(g^*) - H(g^* - 1) \geq 0$$
$$\iff \frac{1}{c} \geq \frac{\partial g^*}{\partial c} \frac{\partial}{\partial g^*} \log \Big(H(g^*) - H(g^* - 1) \Big).$$

Suppose $\frac{\partial r^*}{\partial c} \leq 0 \iff \frac{\partial}{\partial c} \rho^{-1} \left(\frac{H(g^*) - 1/2}{c} \right) \leq 0 \iff \frac{\partial}{\partial c} \log(H(g^*) - 1/2) \leq \frac{d}{dc} \log c = 1/c$.

Notice that $\frac{\partial}{\partial g^*} \log \Big(H(g^*) - H(g^* - 1) \Big) < \frac{\partial}{\partial g^*} \log \Big(H(g^*) - 1/2 \Big)$:

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

Suppose that $H(g) > 1/2$. Then

$$\iff \frac{h(g-1)}{h(g)} > \frac{H(g-1) - 1/2}{H(g) - 1/2},$$

and this holds by the MLRP. Suppose $H(g) < 1/2$. Then

$$\iff \frac{h(g-1)}{h(g)} < \frac{1/2 - H(g-1)}{1/2 - H(g)},$$

and this also holds by the MLRP.

If $g^* \geq g^{\dagger\dagger}$, so $\frac{\partial r^*}{\partial c} \leq 0 \iff \frac{\partial}{\partial c}(H(g^*) - 1/2) \leq 1/c$, then $\frac{\partial}{\partial g^*} \log \Big(H(g^*) - H(g^* - 1) \Big) < 1/c \iff \frac{\partial \log \kappa^*}{\partial c} \geq 0$.

If $g < g^{\dagger\dagger}$, then the sign of $\frac{\partial \kappa^*}{\partial c}$ is unclear and depends on parameters. ∎

Recall

$$s^*(\kappa) = \mathbf{1} \Big\{ \kappa \Big(H(g^*) - H(g^* - 1) \Big) > c \Big\}.$$

Let $\kappa = 1$. Then as c increases there exists two g^*s such that $H(g^*) - H(g^* - 1) = c$ as $H(g^*) - H(g^* - 1)$ is single-peaked with respect to g^* (and symmetric around 1/2.).

Let $\bar{c} := H(1/2) - H(-1/2) = 2H(1/2) - 1$. If $c > \bar{c}$, then $\{g : H(g^*) - H(g^* - 1) = c\} = \varnothing$.

If $\frac{\partial}{\partial c}\Big(H(g^*) - H(g^*-1)\Big) = \frac{\partial g^*}{\partial c}\frac{\partial}{\partial g^*}\Big(H(g^*) - H(g^*-1)\Big) < 1$, then

$$\underline{g}^*(c) := \min_{g^*}\{g : H(g^*) - H(g^*-1) = c\} \text{ increases}$$
$$\overline{g}^*(c) := \max_{g^*}\{g : H(g^*) - H(g^*-1) = c\} \text{ decreases.}$$

■

A.4 Log-concave distribution of κ

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 (9) 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 Implications of Noise in Player's Observations

One may ask why we need noise in players' observation. In the following section, we indicate that perverse comparative statics results can arise if players' observation is noiseless. We choose our model over alternative models without the various sources of noise for two reasons.

First, the equilibrium in our model is *continuous* as it changes marginally for the marginal possibility of mistakes. This is because the best responses of players are *continuous* functions of other players' strategies. Thus, the comparative statics results are robust under any degree of noise other than 0.

In contrast, comparative statics from the equilibrium in the alternative games we discuss now are knife-edge cases where noise or the possibility of a small change from the equilibrium is zero. This is because players' best responses have a *discontinuous* point where a small change in one player's strategy can lead to a drastic change in the opponent player's BR. So, the results from the following models are not robust if we consider an η chance of deviation from the equilibrium. See Echenique and Edlin (2004) for more on this topic.

Second, setting aside robustness concerns, the requirement that each player *must* completely conjecture others' strategy without any mistake or noise is unrealistic given the complex strategic environment of our context.

B.1 Noise in Voter's Observation

Suppose that the voter's observation has no noise, i.e. g is a deterministic function of ω, a, and s. Then r^* is *decreasing* in q.

B.2 Noise in Incumbent's Observation

Suppose that the incumbent observes ω without any noise. One can show that g^* is then constant to q.

C Sabotage's Countervailing Effects on Voter Inference

C.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 an *obfuscated* signal of the reform's true value of ω. 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 $g = x + \eta$.

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.

Table 1: Confusion Matrix for Voter Inference

		Prediction	
		$g < g'$	$g > g'$
Actual condition	$\omega = 1$	FN	TP
	$\omega = 0$	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 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.

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

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

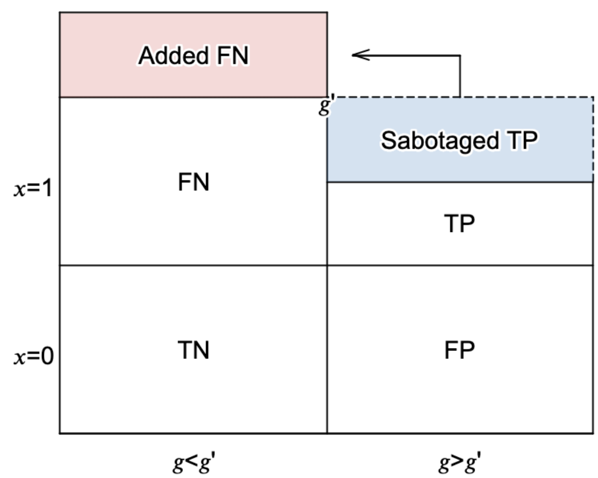
On the other hand, if the voter is too stringent so that g' is too high, he risks not choosing the reform when it is better than the status quo. So, he wants to pick a low enough g' such that the false omission rate (FOR), i.e.

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

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 ω. 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 $Pr[\omega = 1|g]$. Thus, to maintain indifference at the cutoff, the voter lowers g' as the incumbent filters more failed reform by increasing r'.

Figure A1: 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.

Finally, consider the additional obfuscation through bureaucratic sabotage. Assume bureaucrats sabotage reform that would otherwise be successful and supported by voters (i.e., $\omega = 1$ 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 A1 provides the intuition for this result. Firstly, it *decreases* $\Pr[\omega = 1|g \geq g']$ 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 $\omega = 1$), the voter is inclined 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 $1 - \kappa'$,

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

We call this the *PPV effect*.

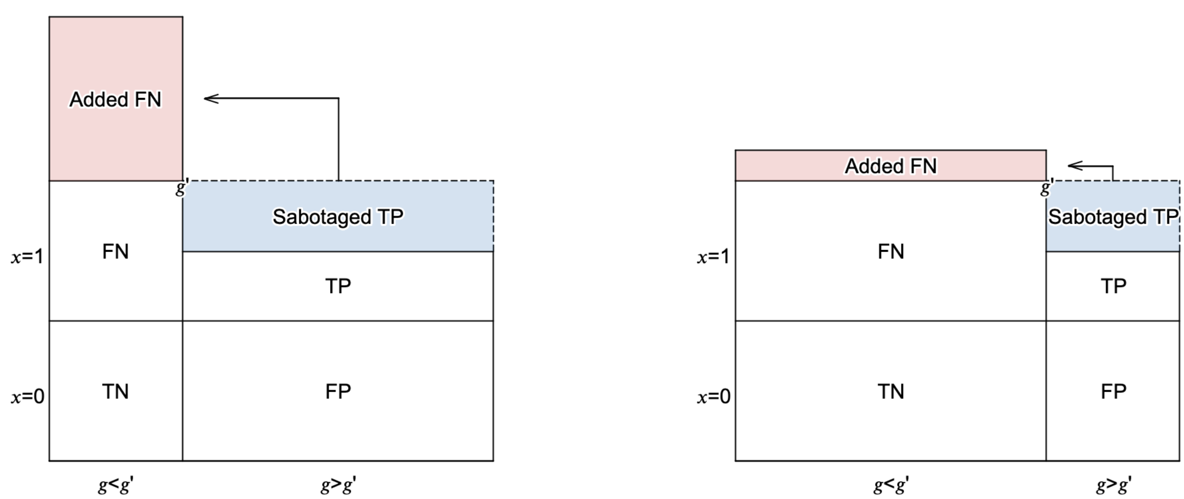
Secondly, the change from TP to FN *increases* $\Pr[\omega = 1|g < g']$ by increasing $\Pr[FN]$ (the red shaded area "Added FN"). Namely, when the voter takes into account the fact that some of the negative signals that he observes are due to sabotage, his evaluation of the reform given a negative signal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

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

We call this the *FOR effect*.

Which effect dominates depends on the initial level of g'. See Figure A2 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.[5] 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.[6] Taken together, the effect of sabotage on voter behavior depends on what 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 A2: 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 ω when they make their decision on sabotage, as long as sabotage can affect g's distribution

[5] We provide calculations of these quantities based on Figure A2 in the next section.

[6] 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).

only when the reform actually works, the logic above holds.

C.2 Examples of Inference Effects

Here, we provide a specific example for the results discussed in Section C.1, fixing the values of g' to those shown in Figure A2. 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 $\Pr[\omega = 1] = \Pr[TP] + \Pr[FN] = 1/2$. Without sabotage,

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

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%,

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

and

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

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,

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

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

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

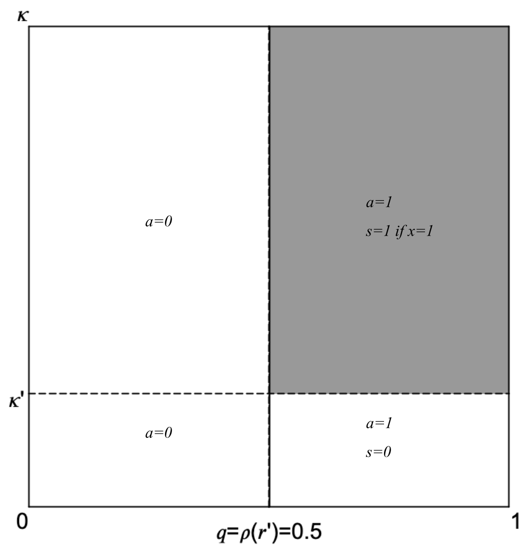
and

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

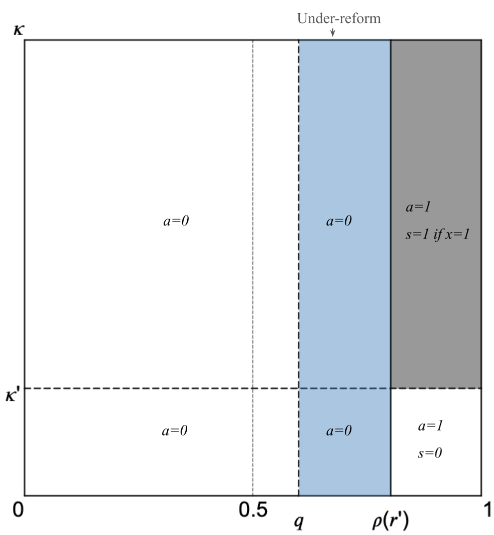
Here, the FOR effect is larger and dominates the PPV effect. For the general result, see the Appendix of Heo and Landa (2024).

D Additional Figures

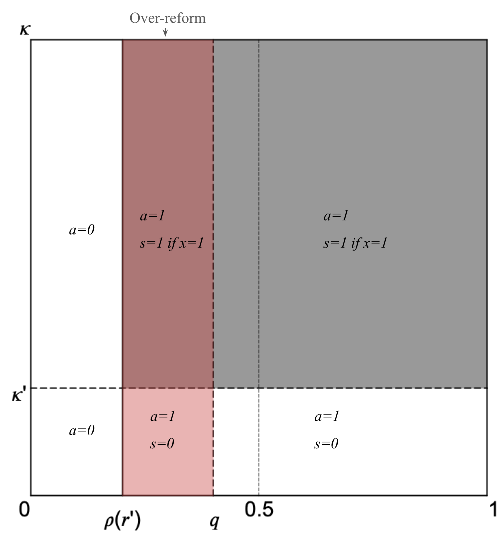
Figure A3: Equilibrium Outcomes



(a) Efficient Reform, $q = 0.5$



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



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