Bureaucratic Rumours and Policymaking *

Shafin Shabir†

Abstract

This paper examines how bureaucrats strategically use informal communication to influence policymaking. I construct a two-period model with three agents: a Leader who plans to enact a policy, a Citizen who can block the policy, and a Bureaucrat who transmits information about the policy through rumours. I model bureaucrats as strategic agents whose preference alignment alters equilibrium outcomes, rather than passive intermediaries. I characterise Perfect Bayesian Equilibria under two bureaucratic alignment scenarios. Citizen-aligned bureaucrats use truthful information transmission when management costs are low, enabling citizens to prevent extreme policies preemptively. In contrast, leader-aligned bureaucrats strategically withhold or misrepresent information, thereby facilitating policy implementation that would otherwise face citizen opposition. I also analyse the impact of bureaucratic alignment on citizen-weighted welfare. These findings demonstrate that bureaucrats impact policy outcomes through rumours, even when bureaucrats lack formal authority.

Keywords- *Bureaucratic Rumours, Bureaucratic Alignment, Policymaking.*

*I acknowledge the assistance of Claude (Anthropic) in reviewing mathematical proofs and editorial suggestions. All core theoretical contributions and research insights are my own.

†CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 11121 Prague, Czech Republic. E-mail: shafin.shabir@cerge-ei.cz.

1 Introduction

Policymaking is a complex process characterised by informational asymmetries among key stakeholders. These asymmetries shape the strategic interaction between leaders who enact the policies, bureaucrats who implement them, and citizens who might resist and challenge them. While formal institutions like independent media traditionally mediate informational flows between the government and citizens, bureaucrats' use of informal channels to transmit information represents a critical yet understudied mechanism in this process.

This paper examines a specific communication channel: bureaucratic rumours, which are informal, often anonymous communications from bureaucrats to citizens regarding proposed policies. I develop a formal model to explore how bureaucrats strategically use rumours to influence policy implementation, particularly in countries where constraints on executive authority are weak. In particular, my research question is to understand under which conditions bureaucrats can effectively influence policy outcomes by convincing citizens to block proposed policies.

I construct a two-period model with a leader, a bureaucrat and a citizen to study the role of bureaucrats in policymaking. Information asymmetry regarding the nature of the policy exists in the first period, and in the second period, all the information is publicly available. In period 1, the leader can enact an extreme version or a moderate version of a policy, which is observable to the bureaucrat but not to the citizen. The Bureaucrat sends rumours (anonymous messages) to the citizen about the nature of the planned policy, and the citizen can choose to block the proposed policy. Citizen blocking at this stage means that a moderate version of the policy is enacted or the policy is shelved. However, if the citizen decides not to protest, the policy is enacted and becomes observable to the citizens in period 2. The citizen can still choose to block the policy, but at a higher cost.

I examine three variations of the model: a benchmark case where bureaucrats babble, a model with citizen-aligned bureaucrats, and a model with leader-aligned bureaucrats. In each case, I characterise Perfect Bayesian Equilibria and analyse how bureaucratic preferences affect information transmission and policy outcomes. This paper demonstrates that bureaucrats act as strategic agents and not just as passive intermediaries. Their preference alignment creates equilibria which would not have emerged in their absence. I find that bureaucratic alignment significantly influences equilibrium policy choices via its impact on information transmission. Citizen-aligned bureaucrats engage in truthful revelation when management costs are low, enabling precise citizen responses that effectively prevent extreme policies. Conversely, leader-aligned bureaucrats strategically misreport information to increase the likelihood of an extreme policy being implemented. I also analyse the impact of bureaucratic alignment on citizen-weighted welfare. I find that under certain conditions, citizen-aligned systems lead to higher welfare compared to leader-aligned and benchmark scenarios.

This model builds on three strands of literature. First, it extends cheap talk models (Battaglini, 2002; Crawford & Sobel, 1982; Krishna & Morgan, 2001) to a three-agent setting with strategic intermediaries. While the standard cheap talk models examine communication between an informed sender and an uninformed receiver with partially aligned interests, my model incorporates a third actor (the bureaucrat) who acts as a strategic intermediary between the leader and the uninformed citizen. The closest papers to my setting are the papers on mediated cheap talk (Ambrus et al., 2013; Blume et al., 2023; Goltsman et al., 2009), which have a three-tier structure where a mediator (usually neutral) mediates between an informed sender and an uninformed receiver. In contrast, I have a two-period setting where the mediator (bureaucrat) observes the leader's policy inclination, and the uninformed citizen actively blocks the leader's policy.

Second, it contributes to the literature on authoritarian control and power-sharing. Svolik (2009) classifies problems that an autocratic leader faces as either the problem of authoritarian control or the problem of authoritarian power sharing. The former is the political problem the leader faces in controlling the masses. This includes potential threats from the citizens to the leader's rule. Egorov and Sonin (2024) give an overview of the recent literature in economics that deals with the problem of authoritarian control. The latter issue, the problem of authoritarian power sharing, arises as a response to the former problem. To control and rule over the masses, a leader needs an alliance with whom she can entrust power to tackle the problem of authoritarian control. The literature has called this group by different names - Selectorate (De Mesquita et al., 2005), Viziers (Egorov & Sonin, 2011) or coalition (Acemoglu et al., 2008). The underlying problem with sharing power is that a leader might face threats from the people with whom they share their power. Though bureaucrats are not formal power-sharing partners, they can use their informational advantage to constrain leaders' policy choices, specifically when formal institutional checks are weak.

Third, it contributes to the broader literature on bureaucrats in political economy models (Alesina & Tabellini, 2007; Egorov & Sonin, 2011; Forand & Ujhelyi, 2021; Forand et al., 2023; Fox & Jordan, 2011; Sasso & Morelli, 2021). Unlike Egorov and Sonin (2011), who study the bureaucratic selection strategies that dictators use, or Forand and Ujhelyi (2021) who analyses the optimality of restricting political activity of bureaucrats, I focus on bureaucrats' strategic use of anonymous communication to influence policymaking. Unlike existing models where bureaucrats are either passive implementers or direct policy influencers, I model bureaucrats as strategic information intermediaries whose alignment determines equilibrium policy outcomes.

Finally, the paper is particularly significant in an increasingly less democratic world (Graham & Svolik, 2020; Guriev & Treisman, 2022; Luo & Przeworski, 2023). I show how informal mechanisms might constrain the executive overreach when formal institutions are weaker. The presence of bureaucratic intermediaries fundamentally alters equilibrium outcomes, with the direction of change depending on bureaucratic alignment- citizen-aligned bureaucrats constrain leader behaviour while leader-aligned bureaucrats can facilitate policy implementation.

This paper proceeds as follows: In Section 2, I present some examples to further motivate the importance of informal communication by bureaucrats in policy making. In Section 3, I present the formal model, and in Section 4 I characterise the Perfect Bayesian Equilibria. Section 5 deals with comparative static exercises, and Section 6 compares welfare. Finally, I conclude in Section 7.

2 Motivating Examples

This section presents two contrasting examples that illustrate the role of bureaucratic rumours in ex-ante policy revelation, which is central to my theoretical model. These examples demonstrate how bureaucrats make deliberate choices about when and how to leak policy information, and how citizens respond to such information by blocking or allowing policy implementation. The cases highlight the key elements of my model: information asymmetry between leaders and citizens, bureaucratic intermediation through anonymous communication, and differential costs of blocking policies at different stages of the policy process.

National Register of Citizens in India (2019) In 2019, the Indian government considered implementing the nationwide National Register of Citizens (NRC), after a successful but controversial implementation in the state of Assam. The purpose of NRC is to identify Indian citizens and "illegal" immigrants residing within the country. The issue with proving citizenship is that many existing citizens, especially ones from poorer economic households, might not have the required documents to prove their citizenship. In fact, in Assam, 1.9 million residents were excluded from the register (Agarwal & Salam, 2019). The list had people from different religious denominations- Hindus, Muslims, Sikhs, etc. However, NRC combined with the Citizenship (Amendment) Act 2019 (CAA,2019), which provides legal options for non-Muslims from the neighbouring countries to regain citizenship, raised concerns about potential disenfranchisement of Muslim minorities.

In the latter part of 2019, anonymous bureaucratic rumours appeared in the public regarding a possible implementation of the NRC. These rumours usually look like articles in the newspaper, citing anonymous officials within the government. For example, in this case, the following sentence from a newspaper article is a bureaucratic rumour:

"Once the NPR is completed and published, it is expected to be the basis for preparing the National Register of Indian Citizens (NRIC), a pan-India version of Assam's National Register of Citizens (NRC), an **official said**."[1] (emphasis added by author)

[1] Press Trust of India (2019)

These rumours contributed to nationwide protests across India against both the proposed NRC and the CAA,2019, ultimately forcing the government to postpone nationwide NRC implementation. This case demonstrates how bureaucratic rumours can correctly signal policy intentions and enable citizens to preemptively block the policy compared to costlier post-enactment opposition.

Demonetization in India (2016) In contrast to NRC, when the Indian government demonetised 500 and 1000 rupee notes on 8 November 2016, there were no prior bureaucratic leaks. This policy, which demonetised 86% of the currency in circulation, was implemented without any warning (Lahiri, 2020). The absence of rumours prevented preemptive citizen action, allowing the government to implement the policy without significant resistance. This case demonstrates that bureaucrats strategically choose when to transmit information, and that the absence of leaks can be as consequential as their presence.

These examples illustrate several key mechanisms in my model. First, they demonstrate information asymmetry where bureaucrats observe policy choices but citizens do not, thus creating opportunities for bureaucratic intermediation. Second, they show how bureaucrats use rumours strategically to transmit information. They do not always send rumours, but strategically decide when to reveal the policy ex-ante. Third, they highlight differential blocking costs: in the NRC case, early information enabled low-cost preemptive action (protests), while in the demonetization case, post-implementation resistance would have been much costlier and less effective. Finally, they suggest different bureaucratic motivations—the NRC leaks appear consistent with citizen-aligned bureaucrats seeking to prevent controversial policy implementation, while the absence of demonetization leaks suggests either leader-aligned bureaucrats or successful information control. I use these examples as motivation and propose a framework that captures these features in my theoretical model.

3 Model

Consider a two-period model with a Leader (L), a Bureaucrat (B), and a Citizen (C). The leader chooses whether to enact an extreme version of policy $p \in \{0, 1\}$, where $p = 1$ represents enactment of a potentially controversial or extreme version of the policy that citizens may wish to block, and $p = 0$ represents the enactment of a moderate version of the policy. The leader receives a higher payoff if the extreme version of the policy is implemented, which happens if the citizen doesn't block the policy in either period. However, if the policy is blocked in period 2 (after it has become observable to everyone), then the leader pays an additional cost in order to retract the policy. These costs represent the institutional, political, and economic penalties that the leader faces after reversing the implemented policies.

The bureaucrat observes the leader's policy choice and sends a rumour, which is a costless message $m \in \{0, 1\}$ to the citizen. I consider two different scenarios, one where the bureaucrat is aligned with the leader and hence gets a higher payoff if an extreme version of the policy is implemented, and the second where the bureaucrat is aligned with the citizen and hence prefers a moderate version of the policy. The citizen can block the policy in either period by taking action $a_i \in \{0, 1\}$ in period $i \in \{1, 2\}$, where $a_i = 1$ represents blocking the policy. In period 1, the citizen does not observe the policy and uses Bureaucratic rumours to decide whether to block or not. In period 2, the citizen observes the policy directly, as long as it was not blocked in period 1.

3.1 Payoffs

The **Leader's payoff** is given by

$$U^L(p, a_1, a_2) = p(1 - a_1)[(1 - a_2) - a_2 r] \tag{1}$$

where $p \in \{0, 1\}$ is the payoff of the policy to the leader, and $r > 0$ denotes the retraction cost of the leader.

The **Citizen's payoff** is

$$U^C(p, a_1, a_2) = -(1 - a_1)(1 - a_2)p - a_1 c_1 - (1 - a_1)a_2(c_2) \tag{2}$$

where $p \in \{0, 1\}$ is the loss the citizen faces if the policy is implemented. The parameters c_1 and c_2 denote the costs of blocking the policy in period 1 and 2, respectively, with $0 < c_1 < c_2$. The assumption $c_1 < c_2$ captures the idea that blocking a policy before its enactment (through protests or lobbying) is less costly than attempting to reverse a policy after it has been enacted. Post enactment the policy has to be challenged through expensive, time-consuming formal channels (courts, legislations, etc.) in addition to the relatively cheaper means of informal pressure (protests or lobbying). Therefore, the assumption of lower blocking costs in period 1 for the citizen does not seem excessive.

The **Bureaucrat's payoff** is

$$U^B(p, a_1, a_2) = (1 - a_1)(1 - a_2)M(p) - a_1 k_1 - (1 - a_1)a_2(k_2) \tag{3}$$

where $M(p)$ represents a bureaucrat's alignment or policy preferences. The parameters k_1 and k_2 denote the management costs the bureaucrat faces when the policy is blocked in period 1 and 2, respectively, with $0 < k_1 < k_2$. Management costs can be thought of as costs the bureaucrat bears in order to implement the policy- time and resources spent on policy preparation, consultations with stakeholders, etc. If the policy is blocked, then this effort from the bureaucrat is wasted and hence lowers the bureaucrat's payoff. If the policy is blocked in period 2, then the bureaucrat will also have to put in additional effort in order to backtrack the enacted policy. Therefore, the management costs are higher in the second period.

I consider two specifications of $M(p)$: Citizen-Aligned Bureaucrat (CAB) with $M(p) = -p$ and Leader-Aligned Bureaucrat (LAB) with $M(p) = p$. The former prefers a moderate policy and hence has similar policy preferences as the citizen, therefore aligned with the citizen. The latter prefers an extreme version of the policy and hence is aligned with the leader.

3.2 Key Assumptions

The model incorporates several key assumptions that drive the strategic interactions:

- **Information asymmetry**: The leader's policy choice is observable to the bureaucrat but not to the citizen in period 1. So the bureaucrat observes whether the leader plans to enact an extreme version or a moderate version of the policy.
- **Differential blocking costs**: $c_1 < c_2$ reflects that preemptive action against a policy is less costly than post-enactment reversal.
- **Anonymous communication**: The bureaucrat's message is costless and anonymous, preventing direct retaliation.
- **Management costs**: $k_1 < k_2$ captures that bureaucrats face higher management costs when policies are blocked after implementation.

3.3 Timing

The game proceeds as follows:

1. Period 1:
   (a) L chooses a policy $p \in \{0, 1\}$, which is observed by L and B but not C.
   (b) B sends a costless message $m \in \{0, 1\}$ to C.
   (c) C updates her belief about p conditional on the message m and decides whether to take action $a_1 \in \{0, 1\}$.
   (d) If $a_1 = 1$, i.e., C blocks the policy, the game ends and the payoffs are realised.
2. Period 2 (If $a_1 = 0$, i.e., C does not block the policy in period 1):
   (a) p is enacted and becomes observable to everyone.
   (b) C decides whether to block it, $a_2 \in \{0, 1\}$.
   (c) Payoffs are realised.

3.4 Strategies and Solution Concept

The strategies of the agents in the model are as follows:

- **Leader's strategy:** can be represented as $\sigma_L = q$, where $q \in [0, 1]$ denotes the probability with which she chooses $p = 1$.
- **Bureaucrat's strategy:** is defined by

$$\sigma_B(m, p) = \mathbb{P}(m|p) \tag{4}$$

  which specifies the probability of sending message m conditional on the observed policy p.

- **Citizen's strategy:** can be represented as:

$$\sigma_C : \mathcal{I}_C \to \{0, 1\} \tag{5}$$

  where $\mathcal{I}_C = \{(1, m) : m \in \{0, 1\}\} \cup \{(2, p) : p \in \{0, 1\}\}$ represents the citizen's information set. This mapping specifies whether the citizen blocks (action 1) or does not block (action 0) based on either the message received in period 1 or the policy observed in period 2.

The solution concept for the game is Perfect Bayesian Equilibrium (PBE), which consists of strategies $(\sigma_L, \sigma_B, \sigma_C)$ and beliefs $\mu_m(p = 1)$ such that the strategies are sequentially rational given beleifs and the citizen's beliefs are updated according to Bayes' rule:

$$\mu_m(p = 1) = \frac{\sigma_B(m, p = 1)q}{\sigma_B(m, p = 1)q + \sigma_B(m, p = 0)(1 - q)} \tag{6}$$

Perfect Bayesian Equilibrium is the appropriate solution concept as it ensures sequential rationality and consistent beliefs in this game with incomplete information. I present the model's results in the next section.

4 Results

I start by analysing a baseline case in which a bureaucrat sends uninformative messages to the citizen. This is a valid benchmark as in this case the bureaucrat babbles and therefore, the citizen relies on their prior beliefs to decide whether to block the policy in the first period. The citizen ignores the message from the bureaucrat and acts as if there is no bureaucratic messaging, and therefore acts as a plausible benchmark [2]. In the main model, I differentiate between citizen-aligned bureaucrat (CAB) and leader-aligned bureaucrat (LAB), and compare the results with the benchmark case.

[2] I also consider a benchmark without the bureaucrat, which is in Appendix A

4.1 Benchmark

A babbling equilibrium exists when the bureaucrat's messaging strategy is independent of the observed policy, $\mathbb{P}[m = 1|p = 1] = \mathbb{P}[m = 1|p = 0] = \alpha \in [0, 1]$, where α is any constant probability. Under babbling, the citizen learns nothing from messages, so: $\mu_m(p = 1) = q \; \forall m \in \{0, 1\}$, i.e., the citizen's posterior belief equals her prior regardless of the message received.

Citizen Since period 2 is a world with complete information, the citizen chooses whether to block the policy by comparing her payoffs.

$$U^C(a_2 = 0) = -p$$
$$U^C(a_2 = 1) = -c_2$$

Therefore, C's best response in period 2 is

$$a_2^* = \begin{cases} 1 & \text{if } c_2 < p \\ 0 & \text{if } c_2 \geq p \end{cases} \tag{7}$$

When $p = 0$, the best response for the citizen is to play $a = 0$. The citizen blocks a policy whenever the cost of blocking is lower than the loss from the policy. Assuming $c_2 < 1$ implies the policy will never be implemented, since the citizen always strictly prefers blocking the policy in period 2. However, $c_2 > 1$ implies that the citizen never blocks the policy in period 2. This assumption seems more plausible in a democratic government that is transitioning towards autocracy. This case represents a political system, where the independent institutions that are supposed to keep a check on the executive powers have been hijacked by the leader (or a politician) to an extent that blocking (or reversing) the policy after its enactment is extremely costly for the citizen and as a result they never block in the second period. It is precisely in this setting that the role of preemptive blocking becomes extremely important to manage and prevent the enactment of the extreme version of the policy. In the paper, I assume that if $c_2 = p$, i.e., the citizen is indifferent between blocking and allowing the policy, she breaks the indifference by allowing the policy implementation.

In period 1, since the bureaucrat babbles, the citizens' posterior beliefs are the same as their prior beliefs, which are induced by the leader's strategy. Let the leader play $p = 1$ with probability q. The expected payoff of the citizen becomes:

$$\mathbb{E}[U^C(a_1 = 1)] = -c$$
$$\mathbb{E}[U^C(a_1 = 0)] = q[-c_2 \mathbb{1}_{\{c_2<1\}} - \mathbb{1}_{\{c_2 \geq 1\}}] - (1 - q) \cdot 0 = -q \, \min(c_2, 1)$$

where $\mathbb{1}_{\{c_2<1\}}$ is an indicator function that is 1 when the condition is satisfied. If the citizen does not block the policy in period 1 ($a_1 = 0$), then the policy is enacted in period 2 ($p = 1$) with probability q. In period 2, either $c_2 < 1$ or $c_2 > 1$. Equation (19) gives us the corresponding best response for the citizen in both these cases. In period 2, $a_2 = 1$ and $a_2 = 0$ respectively. With probability $(1 - q)$, the policy is not enacted ($p = 0$) and the citizen prefers not to take any action in period 2.

The citizen's best response in period 1 depends on the relationship between $\{c_1, c_2, 1\}$. If $0 < c_1 < c_2 < 1$, then $\mathbb{E}[U^C(a_1 = 0)] = -qc_2$ and the citizen prefers to block in period 1 as long as $c_1 < qc_2$. Similarly, when $0 < c_1 < 1 < c_2$, then $\mathbb{E}[U^C(a_1 = 0)] = -q$, and the citizen prefers to block in period 1 as long as $c_1 < q$. Combining these,

$$a_1^* = \begin{cases} 1 & \text{if } c_1 < q\, min(c_2, 1) \\ 0 & \text{if } c_1 \geq q\, min(c_2, 1) \end{cases} \tag{8}$$

Leader If the leader plays $p = 0$, i.e., sets $q = 0$, then from the best response of the citizen in period 1 (20), the citizen prefers not to block the policy as $c_1 > 0$. In period 2, $p = 0$ is observed by the citizen and, following the best response (19), the citizen prefers not to block in period 2 as well. The leader's payoff is $U^L(p = 0, a_1^* = 0, a_2^* = 0) = 0$. In comparison, if the leader plays $p = 1$, her expected payoff depends on the citizen's strategy and the parameter values.

Proposition 4.1 (Benchmark)**.** *Consider the two-period game with a Leader (L), a Bureaucrat (B) and a Citizen (C), where the bureaucrat uses the babbling stategy* $\mathbb{P}[m = 1|p = 1] = \mathbb{P}[m = 1|p = 0] = \alpha \in [0, 1]$ *. If* $c_2 > 1 > c_1 > 0$*, a mixed equilibrium exists, where the leader enacts the policy* $p = 1$ *with probability* $q^* = c_1$*, and the citizen never blocks the policy:* $a_t = 0\ \forall t \in \{1, 2\}$*, with the beliefs* $\mu_m(p = 1) = q^*\ \forall m \in \{0, 1\}$*.*

Proof. When $0 < c_1 < 1 < c_2$, the leader makes the citizen indifferent between blocking or not blocking in period 1 by playing $q^* = c_1$. Using the assumption that an indifferent citizen decides not to block implies that the game goes to the second period. In period 2, the citizen does not block because $c_2 > 1$ (from equation (19)). The leader's expected payoff is $E[U^L] = q^* \cdot 1 + (1 - q^*) \cdot 0 = q^* = c_1$. Any deviation to $q > c_1$ will trigger blocking in period 1, decreasing the payoff to 0, and any deviation to $q < c_1$ will lead to the expected payoff of the leader being lower than q^*. Therefore, when $0 < c_1 < 1 < c_2$, $\{q^* = c_1, a_1^* = 0, a_2^* = 0\}$ is an equilibrium. Since the bureaucrat is payoff-irrelevant, she can play any babbling strategy $\alpha \in [0, 1]$. □

Assuming $c_2 > 1$ implies that after a policy has been enacted, blocking it is not optimal for the citizen because the blocking cost is higher than the loss from not blocking. In an autocratic setting or a democracy with an authoritarian leader, this assumption does not seem excessive. The leader creates or ensures that the institutions in a country are such that the enacted policy cannot be blocked. This can be done by hijacking the independent media, judiciary, civil service administration, etc. The higher blocking cost in the second period further strengthens the role of information sharing in the first period. Since the citizen knows that blocking in period 2 is not possible, the best course of action for her is to learn about the leader's policy in period 1. This is what the bureaucrat does in the model and is discussed in the next section.

The takeaway from the benchmark model is that the leader can sometimes strategically exploit differential blocking costs to implement extreme policies. These extreme policies might have been blocked in a world with complete information. However, in a world with asymmetric information, the leader can exploit differential blocking costs to ensure that extreme policies can be implemented. This sets up the base for a model with the bureaucrat, hinting towards how the bureaucrat might influence outcomes via information transmission in this model.

4.2 Citizen-Aligned Bureaucrat (CAB)

Now consider a model with a Bureaucrat whose preferences align with the citizen's preferences, i.e., $M(p) = p$. the bureaucrat's utility function becomes $U^B(p, a_1, a_2) = -(1-a_1)(1-a_2)p - a_1k_1 - (1-a_1)a_2k_2$. The bureaucrat's optimal messaging strategy depends on the policy and the management costs. The bureaucrat faces management costs if the citizen decides to block.

Citizen In period 2, the citizen observes the policy p and decides whether to block it. Since this is similar to the benchmark case, the citizen's best response is

$$a_2^* = \begin{cases} 1 & \text{if } c_2 \leq 1 \\ 0 & \text{if } c_2 > 1 \vee p = 0 \end{cases} \tag{9}$$

In period 1, the citizen does not observe p, but uses messages from the bureaucrat to update her prior belief. If bureacrat follows the following strategy $\sigma_B(m = 1, p = 1) = \mathbb{P}[m = 1|p = 1] = \lambda$ & $\sigma_B(m = 0, p = 0) = \mathbb{P}[m = 0|p = 0] = 1$.Then the citizens' updated beliefs will be:

$$\mu_{m=1}(p = 1) = 1 \quad \& \quad \mu_{m=0}(p = 1) = \frac{(1-\lambda)q}{(1-\lambda)q + 1 - q}$$
$$\mu_{m=1}(p = 0) = 0 \quad \& \quad \mu_{m=0}(p = 0) = \frac{1-q}{1-q+(1-\lambda)q}$$

Given the posterior beliefs, the citizen's expected payoff is

$$\mathbb{E}[U^C(a_1 = 1)|m = 1] = -c_1$$
$$\mathbb{E}[U^C(a_1 = 0)|m = 1] = -\mu_{m=1}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -\min(c_2, 1)$$
$$\mathbb{E}[U^C(a_1 = 1)|m = 0] = -c_1$$
$$\mathbb{E}[U^C(a_1 = 0)|m = 0] = \mu_{m=0}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -\mu_{m=0}(p = 1)\min(c_2, 1)$$

To solve the model, consider the two cases: $c_2 < 1$ and $c_2 > 1$. While in the former case, the citizen prefers to block the policy in period 2, whereas in the latter case citizen is better off by allowing the policy to be implemented. As already mentioned, the case $c_2 > 1$ is the more interesting case as it represents a democracy in decline. Therefore, I focus only on this case ($0 > c_1 > 1 > c_2$), while keeping the other cases in the Appendix B.

Therefore, the expected payoffs of the citizen become:

$$\begin{aligned}
&\mathbb{E}[U^C(a_1 = 1)|m = 1] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 1] = -\mu_{m=1}(p = 1)[-c_2\mathbf{1}_{\{c_2\leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -1 \\
&\mathbb{E}[U^C(a_1 = 1)|m = 0] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 0] = \mu_{m=0}(p = 1)[-c_2\mathbf{1}_{\{c_2\leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -\mu_{m=0}(p = 1)
\end{aligned}$$

When $m = 1$, the expected utility from blocking ($-c_1$), is always higher than allowing (1), therefore

$$a_1^*(m = 1) = 1 \tag{10}$$

Similarly, when $m = 0$

$$a_1^*(m = 0) = \begin{cases} 1 & \text{if } c_1 < \mu_0(p = 1) \\ 0 & \text{if } c_1 \geq \mu_0(p = 1) \end{cases} \tag{11}$$

Bureaucrat When $p = 0$, the bureaucrat always prefers message $m = 0$ and has no profitable deviation. When $p = 1$, the bureaucrat's expected utilities are:

$$\begin{aligned}
&\mathbb{E}[U^B(m_1 = 1|p = 1)] = -k_1 \\
&\mathbb{E}[U^B(m_1 = 0|p = 1)] = \begin{cases} -k_1 & \text{if } \quad c_1 < \mu_0(p = 1) \\ -1 & \text{otherwise} \end{cases}
\end{aligned}$$

The bureaucrat's best response depends on the relationship between $k_1, 1$ and c_1 as summarised below

$$\lambda^* = \begin{cases} k_1 < 1 \begin{cases} [0, 1] & \text{if } c_1 < \mu_{m=0}(p = 1) \\ 1 & \text{if } c_1 \geq \mu_{m=0}(p = 1) \end{cases} \\ \\ k_1 > 1 \begin{cases} [0, 1] & \text{if } c_1 < \mu_{m=0}(p = 1) \\ 0 & \text{if } c_1 \geq \mu_{m=0}(p = 1) \end{cases} \end{cases} \tag{12}$$

When $k_1 < 1$ and $c_1 \geq \mu_0(p = 1)$, the bureaucrat strictly prefers to send $m = 1$, i.e., set $\lambda^* = 1$, and the bureaucrat is indifferent between sending $m = 1$ and $m = 0$ when $c_1 \geq \mu_0(p = 1)$. Similarly, when $k_1 > 1$ and $c_1 \geq \mu_0(p = 1)$, the bureaucrat strictly prefers to send $m = 0$ when $p = 1$, i.e., set $\lambda^* = 0$, and the bureaucrat is indifferent between sending $m = 1$ and $m = 0$ when $c_1 < \mu_0(p = 1)$.

Leader Following from our earlier analysis, the leader's expected utility is

$$\begin{aligned}\mathbb{E}[U^L(p = 1)] &= \lambda \cdot 0 + (1 - \lambda)[\mathbb{1}_{\{c_1 < \mu_{m=0}(p=1)\}} \cdot 0 + \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)\}}] \\ &= (1 - \lambda)\mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)\}}\end{aligned}$$

Therefore, the leader's best response is given by

$$q^* = \begin{cases} 1 & \text{if } (1 - \lambda)\, \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)\}} > 0 \\ [0, 1] & \text{if } (1 - \lambda)\, \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)\}} = 0 \end{cases} \tag{13}$$

Proposition 4.2 (Citizen-Aligned Bureaucrat)**.** *Consider a two-period game with a Leader (L), a Citizen-Aligned Bureaucrat (B), and a Citizen (C). When $0 < c_1 < 1 < c_2$, and $k_1 < 1$ there exists an equilibrium where the bureaucrat follows the strategy $\lambda^* = 1$, the leader mixes between enacting and not enacting a policy $q^* \in [0, 1]$ and the citizen blocks in period 1 upon receiving message $m = 1$, and does not block upon receiving message $m = 0$: $a_1^*(m = 0) = 0, a_1^*(m = 1) = 1$. In period 2, if the policy is enacted, the citizen does not blocks it $a_2^*(p) = 0\ \forall p$, and the beliefs of the citizen are $\mu_{m=1}(p = 1) = 1$, $\mu_{m=0}(p = 0) = 1$.*

Proof. **Truthtelling Equilibrium**: When $k < 1$, then $\lambda^* = 1$, i.e., the bureaucrat reports truthfully, the posterior belief of the citizen becomes $\mu_0(p = 1) = 0$. From Citizen's best response in period 1 in equation (10) and equation (11), the citizen plays $a_1^*(m = 0) = 0, a_1^*(m = 1) = 1$. Since $c_2 > 1$, this implies in period 2, the citizen does not block the policy $a_2^*(p = 1) = 0$, $a_2^*(p = 0) = 0$. From the bureaucrat's best response, since $k_1 < 1$ and $c_1 > \mu_0(p = 1) = 0$, the bureaucrat gets a strictly higher payoff when she sends message $m = 1$ (payoff $-k_1$) given that $p = 1$. From the leader's best response (13), the leader is indifferent between setting $p = 0$ or $p = 1$, i.e., $q^* \in [0, 1]$ because both choices give her the same payoff. □

Since the leader is indifferent between $q \in [0, 1]$, it is not clear what really happens in equilibrium. In order to pin down the leader's choice, I refine the equilibrium by selecting the equilibrium that pareto dominates the other equilibria. An equilibrium is pareto dominant iff a pareto improvement is not possible, i.e., no one can be made better-off without making someone else worse off. When $q = 0$, $U^L = 0$, $U^C = 0$, $U^B = 0$, whereas, when $q > 0$, $U^L = 0$, $U^C = -q \cdot c_1 < 0$, $U^B = -q \cdot k_1 < 0$. Since $q = 0$ pareto dominates $q > 0$, the refined equilibrium choice of the leader is $q^* = 0$. A way to interpret this is that the threat of revelation from the bureaucrat acts as a deterrence to the leader, and eventually, the leader does not enact extreme policies in equilibrium.

Therefore, in regimes where the cost of blocking is very high ($c_2 > 1$), ex-ante information on policies becomes extremely useful in preventing the implementation of extreme policies. In contrast to the benchmark case (Proposition A.1), where the leader might successfully implement a policy, in this case (Proposition 4.2), the extreme policy is blocked by the citizen in period 1, given the bureaucrat's truthful revelation. This illustrates that in a society where independent checks (media, judiciary, legislatures) have been hijacked by the leader, informal communication from citizen-aligned bureaucrats is extremely important in convincing the citizen to preemptively block the implementation of extreme policies.

4.3 Leader-Aligned Bureaucrat (LAB)

In contrast to the previous case, consider a model with a Bureaucrat whose preferences align with the leader, i.e., $M(p) = -p$. the bureaucrat's utility function becomes $U^B(p, a_1, a_2) = (1-a_1)(1-a_2)p - a_1 k_1 - (1-a_1)a_2 k_2$. The bureaucrat's optimal messaging strategy depends on the policy and the management costs. The bureaucrat faces management costs if the citizen decides to block.

Citizen In period 2, the citizen observes the policy p and decides whether to block it by comparing her utilities. Using a similar methodology as in the Citizen-Aligned Bureaucrat case, the best Citizen's best response as

$$a_2^* = \begin{cases} 1 & \text{if } c_2 \leq 1 \\ 0 & \text{if } c_2 > 1 \vee p = 0 \end{cases} \tag{14}$$

In period 1, the expected payoffs of the citizen given the message are

$$\begin{aligned}
&\mathbb{E}[U^C(a_1 = 1)|m = 1] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 1] = -\mu_{m=1}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2 > 1\}}] = -\min(c_2, 1) \\
&\mathbb{E}[U^C(a_1 = 1)|m = 0] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 0] = \mu_{m=0}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2 > 1\}}] = -\mu_{m=0}(p = 1)\min(c_2, 1)
\end{aligned}$$

Once again, consider the two cases separately: $c_2 < 1$ and $c_2 > 1$. While in the former case, the citizen prefers to block the policy in period 2, whereas in the latter case citizen is better off by allowing the policy to be implemented. As already mentioned, the case $c_2 > 1$ is the more interesting case as it represents a democracy in decline. Therefore, I focus only on this case ($0 > c_1 > 1 > c_2$), while keeping the other cases in the Appendix B.

Therefore, the expected payoffs of the citizen become:

$$\begin{aligned}
&\mathbb{E}[U^C(a_1 = 1)|m = 1] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 1] = -\mu_{m=1}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2 > 1\}}] = -1 \\
&\mathbb{E}[U^C(a_1 = 1)|m = 0] = -c_1 \\
&\mathbb{E}[U^C(a_1 = 0)|m = 0] = \mu_{m=0}(p = 1)[-c_2\mathbf{1}_{\{c_2 \leq 1\}} - \mathbf{1}_{\{c_2 > 1\}}] = -\mu_{m=0}(p = 1)
\end{aligned}$$

Using similar arguments as in the CAB case, the best response of the citizen in period 1 will be

$$a_1^*(m = 1) = 1 \tag{15}$$

Similarly, when $m = 0$

$$a_1^*(m = 0) = \begin{cases} 1 & \text{if } c_1 < \mu_0(p = 1) \\ 0 & \text{if } c_1 \geq \mu_0(p = 1) \end{cases} \tag{16}$$

Bureaucrat When $p = 0$, the bureaucrat always prefers message $m = 0$ and has no profitable deviation as in the previous case. When $p = 1$, the bureaucrat's expected utilities are:

$$\begin{aligned}
&\mathbb{E}[U^B(m_1 = 1|p = 1)] = -k_1 \\
&\mathbb{E}[U^B(m_1 = 0|p = 1)] = \begin{cases} -k_1 & \text{if } \quad c_1 < \mu_0(p = 1) \\ 1 & \text{otherwise} \end{cases}
\end{aligned}$$

Therefore, the bureaucrat's best response is

$$\lambda^* = \begin{cases} [0, 1] & \text{if } c_1 < \mu_{m=0}(p = 1) \\ 0 & \text{if } c_1 \geq \mu_{m=0}(p = 1) \end{cases} \tag{17}$$

When $c_1 \geq \mu_0(p = 1)$, the bureaucrat strictly prefers to send $m = 0$ when $p = 1$, i.e., set $\lambda^* = 0$, and the bureaucrat is indifferent between sending $m = 1$ and $m = 0$ when $c_1 < \mu_0(p = 1)$.

Leader Following from our earlier analysis, the leader's expected utility when $p = 1$ is

$$\begin{aligned}
\mathbb{E}[U^L(p = 1)] &= \lambda \cdot 0 + (1 - \lambda)[\mathbb{1}_{\{c_1 < \mu_{m=0}(p=1)\}} \cdot 0 + \mathbb{1}_{\{c_1 \geq \mu_{m=0}(p=1)\}}] \\
&= (1 - \lambda)\mathbb{1}_{\{c_1 \geq \mu_{m=0}(p=1)\}}
\end{aligned}$$

Therefore, the leader's expected utility is

$$\begin{aligned}\mathbb{E}[U^L] &= q \cdot \mathbb{E}[U^L(p=1)] + (1-q) \cdot 0 \\ &= q \cdot (1-\lambda)\mathbb{1}_{\{c_1 \geq \mu_{m=0}(p=1)\}}\end{aligned}$$

Proposition 4.3 (Leader-Aligned Bureaucrat)**.** *Consider the two-period game with a Leader (L), a Leader-Aligned Bureaucrat (B), and a Citizen (C). When $0 < c_1 < 1 < c_2$ the following constitutes perfect Bayesian Equilibria :*

1. ***Status-quo Equilibrium**: There exists an equilibrium where the leader never enacts the policy $q^* = 0$, λ^* is arbitrary (off-path), and the citizen does not block in period 1: $a_1^*(m = 0) = 0$, $a_1^*(m = 1) = 1$(off-equilibrium path), and in period 2 $a_2^*(p = 0) = 0$, $a_2^*(p = 1) = 1$, with the beliefs $\mu_{m=0}(p = 0) = 1$, $\mu_{m=1}(p = 1) = 1$ (off-equilibrium path).*

2. ***Pooling Equilibrium**: There exists an equilibrium where $\lambda^* = 0$, and $q^* = c_1$, and the citizen would block in period 1 upon receiving message $m = 1$ (off-equilibrium) but does not block upon receiving $m = 0$: $a_1^*(m = 1) = 1$ (off-equilibrium), $a_1^*(m = 0) = 0$). In period 2, the citizen does not blocks it: $a_2^*(p) = 0 \quad \forall p$, and the beliefs of the citizen are $\mu_{m=1}(p = 1) = 1$(off-equilibrium path), $\mu_{m=0}(p = 1) = q$.*

Proof. **Status-quo Equilibirum**: If the leader sets $q^* = 0$, then the posterior belief of the citizen becomes $\mu_0(p = 0) = 1$. Given this, the citizen strictly prefers not to block when $m = 0$ in period 1 ((30)). Since the leader never chooses $p = 1$, the bureaucrat's strategy for $p = 1$ is off the equilibrium path. Recall that the bureaucrat strictly prefers to send $\lambda^* = 1$ whenever $c_1 \geq \mu_0(p = 1)$. Since in this case $c_1 \geq 0$, which is true by assumption, it is plausible to believe the bureaucrat's best response is to play $\lambda^* = 1$, i.e., truthfully reveal the policy. And hence, the citizen chooses to block the policy in period 1 ($a_1^*(m = 1) = 1$) and the posterior belief of the citizen will be $\mu_{m=1}(p = 1) = 1$. Since no one has an incentive to deviate and the beliefs are consistent, this strategy constitutes a PBE.

Pooling Equilibrium: If the bureaucrat follows the strategy $\lambda^* = 0$, the posterior belief of the citizen becomes $\mu_{m=0}(p = 1) = q$. From Citizen's best response in period 1 equation (15) and equation (16), the citizen plays $a_1^*(m = 0) = 0$ when $c_1 \geq q$, and $a_1^*(m = 1) = 1$ (off-equilibrium). Since $c_2 > 1$, this implies in period 2, the citizen does not block the policy $a_2^*(p = 1) = 0$, $a_2^*(p = 0) = 0$.

Since $c_1 > 0$, the bureaucrat gets a strictly higher payoff when she sends message $m = 0$ given that $p = 1$, as can be seen from the bureaucrat's best response (17). The leader wants to maximize her expected utility: $\mathbb{E}[U^L] = q \cdot \mathbb{1}_{\{c_1 \geq q}$. The leader's expected utility is $\mathbb{E}[U^L] = q \cdot \mathbf{1}_{\{c_1 \geq q\}}$, which equals q when $q \leq c_1$ and 0 when $q > c_1$. Therefore, the leader maximises utility by choosing $q^* = c_1$. Therefore, this strategy consitutes a PBE with beliefs: $\mu_{m=0}(p = 1) = q^*$, $\mu_{m=1}(p = 1) = 1$. □

In regimes where the cost of blocking becomes higher ($c_2 > 1$), ex-ante information on policies becomes extremely useful in preventing the implementation of extreme policies, as noted in the previous case.

However, the leader aligned Bureaucrat wants the policy to be implemented and hence tries to conceal or hide the information about the policy. In fact, the bureaucrat gets a higher payoff from misreporting and sending an incorrect message, i.e., $m = 0$ when $p = 1$. The leader optimally chooses $q^* = c_1$, which makes the citizen exactly indifferent between blocking and not blocking in period 1. Given the tie-breaking assumption, the citizen does not block, allowing the policy to succeed with probability c_1. The bureaucrat's behaviour is intuitive. Since they want the policy to be implemented, the best they can do is not to leak the policy ex-ante and therefore mimic the strategy that bureaucrats in the benchmark follow, i.e., send uninformative messages or babble. The citizen is aware of this and therefore uses her prior beliefs to decide whether to block a policy preemptively.

The presence of the bureaucrat fundamentally alters equilibrium outcomes. Citizen-aligned bureaucrats can constrain leaders through information revelation, while leader-aligned bureaucrats facilitate policy implementation through strategic information pooling. Comparing across cases: the benchmark allows leaders to exploit information asymmetries when $c_2 > 1$, citizen-aligned bureaucrats eliminate this advantage through truthful revelation, while leader-aligned bureaucrats restore and enhance the leader's strategic position through successful information manipulation.

5 Comparative Statics

This section demonstrates how marginal changes in key parameters affect the equilibrium strategies and outcomes in the model.

5.1 CAB

Case 1: $c_2 < 1$ In the truthtelling equilibrium with $\lambda^* = 1$, the citizen blocks after $m = 1$ and does not block after $m = 0$. Recall, the Bureaucrat's expected utility when $p = 1$ is

$$\mathbb{E}[U^B(m = 1|p = 1)] = -k_1$$
$$\mathbb{E}[U^B(m = 0|p = 1)] = -k_2$$

The bureaucrat strictly prefers to send $m = 1$ when $k_1 < k_2$. When k_1 increases, we have

$$\frac{\partial(\mathbb{E}[U^B(m = 1|p = 1)] - \mathbb{E}[U^B(m = 0|p = 1)])}{\partial k_1} = -1 < 0$$

This means that as k_1 increases, the utility advantage of the bureaucrat of sending $m = 1$ decreases. The truthtelling equilibrium persists as long as $k_1 < k_2$. However, as $k_1 \to k_2$ truthtelling weakens. In terms of policy implications, this suggests that the leader can reduce the role of information transmission by increasing the bureaucratic burden.

Similarly, when k_2 increases $\frac{\partial(\mathbb{E}[U^B(m=1|p=1)]-\mathbb{E}[U^B(m=0|p=1)])}{\partial k_2} = 1 > 0$. That is the utility advantage of the bureaucrat of sending $m = 1$ increases, further strengthening the truthtelling equilibrium.

Effect of c_1: As c_1 increases but remains below 1, there is no effect on bureaucrats' incentives. However, when $c_1 = 1$, the threshold condition changes the Citizen's strategy, potentially changing the equilibrium structure.

Case 2: $c_2 > 1 > c_1 > 0$ Similarly, the Bureaucrat's expected utility when $p = 1$ is

$$\mathbb{E}[U^B(m = 1|p = 1)] = -k_1$$

$$\mathbb{E}[U^B(m = 0|p = 1)] = \begin{cases} -k_1 & \text{if citizen blocks after m=0} \\ -1 & \text{if citizen does not block after m=0} \end{cases}$$

From Proposition (4.2), truthtelling equilibrium exists as long as $k_1 < 1$. In this case, the bureaucrat strictly prefers to send message $m = 1$ when $p = 1$ as it gives her higher utility. As k_1 increases above 1, the equilibrium breaks down. Since in this case, the leader will have a strong incentive to set $q = 1$, which will cause the citizen to block preemptively even after receiving $m = 0$. Therefore, as k_1 increases above 1, the bureaucrat's optimal strategy changes, and as a result, truthtelling equilibrium breaks down.

This suggests that even when the Bureaucrat is aligned with the citizen, they will want to truthfully reveal policy as long as the management costs are low. Therefore, if the leader wants to ensure that the policy is not revealed preemptively to the citizen, she has to increase the citizen's management cost. Again, management costs can be increased by increasing the bureaucratic burden, weakening whistleblower protections, etc.

5.2 LAB

Case 1: $c_2 > 1 > c_1 > 0$ In the pooling equilibrium, $q^* = c_1$, $\lambda^* = 0$, and the citizen does not block in period 1 if the message is $m = 0$ and blocks if the message is $m = 1$. Note that as c_1 increases, the equilibrium probability $q^* = c_1$ increases, allowing the leader to implement extreme policy more frequently while maintaining citizen indifference. However, if $c_1 \to 1$, the citizen's incentive to block in period 1 decreases, potentially destabilising the equilibrium. In contrast, as k_1 increases, revealing truthfully becomes more costly for the bureaucrat, thereby strengthening the bureaucrat's pooling strategy of $\lambda^* = 0$.

This suggests that when the bureaucrat is aligned with the leader, they will want to mislead the citizen. However, as management costs in period 1 decrease, the bureaucrat's tendency or likelihood to mislead the citizen decreases as well. Therefore, in terms of policy considerations, keeping the bureaucrat's management cost lower might lead to better information transmission and prevent extreme policies from being implemented.

6 Welfare Analysis

In this section, I evaluate the desirability of different bureaucratic systems using a citizen-weighted social welfare function. I define social welfare as:

$$W = \mathbb{E}[U^C] + \theta(\mathbb{E}[U^L] + \mathbb{E}[U^B]) \tag{18}$$

where $\theta \in (0, 1)$ represents the relative weight on leader and bureaucrat welfare compared to citizen welfare. This specification reflects a social welfare function which puts more weight on citizen welfare, while considering leader and bureaucratic welfare as well. The result below compares the welfare rankings for the different cases.

Proposition 6.1 (Welfare)**.** *When $0 < c_1 < 1 < c_2$, the welfare under different cases is:*

1. $W^{bench} = c_1(\theta - 1) < 0$
2. $W^{CAB} = -q(c_1 + \theta k_1) = 0$
3. $W^{LAB} = c_1(2\theta - 1)$

Proof. **Benchmark**: From Proposition (A.1), the mixed equilibrium has $q^* = c_1$. Therefore, expected payoffs are: $\mathbb{E}[U^L] = q^* = c_1$, $\mathbb{E}[U^C] = -q^* = -c_1$. Therefore, $W^{bench} = c_1(\theta - 1) < 0$, since $\theta < 1$.

CAB: From Proposition (4.2), in truthtelling equilibrium the expected payoffs are: $\mathbb{E}[U^L] = q^* \cdot 0 = 0$, $\mathbb{E}[U^C] = \mu_{m=0}(p = 1)(-c_1) + (1 - \mu_{m=0}(p = 1))(-1) = -q^*c_1 = 0$, and $\mathbb{E}[U^B] = -q^*k_1 = 0$. Therefore, $W^{CAB} = -q^*c_1 - \theta(q^*k_1) = -q^*(c_1 + \theta k_1) = 0$.

LAB: From Proposition (4.3), in pooling equilibrium $q^* = c_1$, $\lambda^* = 0$. The expected payoffs are: $\mathbb{E}[U^L] = q^* \cdot 1 = q^* = c_1$, $\mathbb{E}[U^C] = q^* \cdot (-1) + (1 - q^*) \cdot 0 = -q^* = -c_1$, and $\mathbb{E}[U^B] = q^* \cdot (1) + (1 - q^*)(0) = q^* = c_1$. Therefore, $W^{LAB} = -c_1 + \theta(c_1 + c_1) = c_1(2\theta - 1)$.

□

Comparing the Welfare under benchmark, CAB and LAB, we get the following results:

Proposition 6.2 (Welfare under CAB higher)**.** *The CAB case has higher welfare compared to the benchmark as $\theta < \frac{1}{2}$.*

Proof. Identify when $W^{CAB} > W^{bench}$, i.e.,

$$0 > c_1(\theta - 1)$$

which is true if $\theta < \frac{1}{2}$. □

To summarise, it seems that the presence of a bureaucrat is not always welfare-improving. The leader-aligned bureaucratic systems can be welfare-optimal when the weight on leader and bureaucrat welfare relative to citizen welfare is sufficiently high ($\theta > \frac{1}{2}$). In contrast, when the weight on the leader and citizen is low, then a system with CAB dominates other systems. This happens as CAB enables precise blocking at relatively lower cost for the citizen. However, when k is high, the cost of bureaucratic intermediation exceeds the benefits of information transmission, therefore weakening the truthtelling equilibrium. An important takeaway is that higher management costs can deter citizen-aligned bureaucrats from truthfully revealing the policy. These costs can be increased by overburdening the bureaucrats by forcing them to spend a lot more time on policy preparation than they usually would.

7 Conclusion

This paper develops a model to examine the role of bureaucratic rumours in policymaking. Unlike existing models where bureaucrats are either passive implementers or direct policy influencers, I demonstrate how bureaucrats act as strategic intermediaries whose alignment determines equilibrium outcomes. I characterise three different Perfect Bayesian Equilibria: the benchmark equilibrium without bureaucrats, the citizen-aligned equilibrium in which the bureaucrat reports truthfully and triggers precise blocking from citizens when blocking costs are low, and the leader-aligned equilibrium in which the bureaucrat pools information about policy strategically, facilitating policy implementation. The results show that the citizen-aligned bureaucrats prevent extreme policies when $c_1 < 1$, while the leader-aligned bureaucrats successfully facilitate policy implementation when $0 < c_1 < 1 < c_2$. Additionally, the welfare comparison between the systems depends critically on parameter values. When the weight on the leader and bureaucrat is high ($\theta > \frac{1}{2}$), the leader-aligned bureaucrat system achieves higher welfare. When this weight is low, citizen-aligned systems dominate.

Overall, this paper demonstrates that bureaucrats act as strategic agents whose preference alignment fundamentally alters policy equilibria. The key policy implication is nuanced: while citizen-aligned bureaucrats provide valuable checks on executive power through informal information transmission, leader-aligned bureaucrats can improve coordination and welfare under certain parameter configurations. Additionally, for leaders seeking policy implementation, managing bureaucratic management costs becomes crucial, while those prioritising citizen oversight should focus on protecting bureaucratic independence.

References

Acemoglu, D., Egorov, G., & Sonin, K. (2008). Coalition Formation in Non-Democracies. *The Review of Economic Studies*, *75*(4), 987–1009.

Agarwal, R., & Salam, K. (2019). *India is betraying its founding fathers*. Retrieved December 17, 2019, from https://foreignpolicy.com/2019/12/17/india-citizenship-law-protests-narendra-modi-amit-shah-founding-fathers/#cookie_message_anchor (accessed: 09.11.2023).

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

Ambrus, A., Azevedo, E. M., & Kamada, Y. (2013). Hierarchical Cheap Talk. *Theoretical Economics*, *8*(1), 233–261.

Battaglini, M. (2002). Multiple Referrals and Multidimensional Cheap Talk. *Econometrica*, *70*(4), 1379–1401.

Blume, A., Lai, E. K., & Lim, W. (2023). Mediated Talk: An Experiment. *Journal of Economic Theory*, *208*, 105593.

Crawford, V. P., & Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, 1431–1451.

De Mesquita, B. B., Smith, A., Siverson, R. M., & Morrow, J. D. (2005). *The Logic of Political Survival*. MIT press.

Egorov, G., & Sonin, K. (2011). Dictators and their Viziers: Endogenizing the Loyalty-Competence Trade-off. *Journal of the European Economic Association*, *9*(5), 903–930.

Egorov, G., & Sonin, K. (2024). The Political Economics of Non-democracy [Publisher: American Economic Association]. *Journal of Economic Literature*. https://doi.org/10.1257/jel.20221494

Forand, J. G., & Ujhelyi, G. (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, J. G., Ujhelyi, G., & Ting, M. M. (2023). Bureaucrats and Policies in Equilibrium Administrations. *Journal of the European Economic Association*, *21*(3), 815–863.

Fox, J., & Jordan, S. V. (2011). Delegation and Accountability. *The Journal of Politics*, *73*(3), 831–844.

Goltsman, M., Hörner, J., Pavlov, G., & Squintani, F. (2009). Mediation, Arbitration and Negotiation. *Journal of Economic Theory*, *144*(4), 1397–1420.

Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States. *American Political Science Review*, *114*(2), 392–409.

Guriev, S., & Treisman, D. (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st century*. Princeton University Press.

Krishna, V., & Morgan, J. (2001). A Model of Expertise. *The Quarterly Journal of Economics*, *116*(2), 747–775.

Lahiri, A. (2020). The Great Indian Demonetization. *Journal of Economic Perspectives*, *34*(1), 55–74. https://doi.org/10.1257/jep.34.1.55

Luo, Z., & Przeworski, A. (2023). Democracy and its Vulnerabilities: Dynamics of Democratic Backsliding. *Quarterly Journal of Political Science*, *18*(1), 105–130. https://doi.org/10.1561/100.00021112

Press Trust of India, P. (2019). *Government to prepare npr to lay foundation for pan-india nrc*. Retrieved August 3, 2019, from https://www.indiatoday.in/india/story/government-npr-foundation-pan-india-nrc-1576860-2019-08-03 (accessed: 09.11.2023).

Sasso, G., & Morelli, M. (2021). Bureaucrats Under Populism. *Journal of Public Economics*, *202*, 104497.

Svolik, M. W. (2009). Power Sharing and Leadership Dynamics in Authoritarian Regimes. *American Journal of Political Science*, *53*(2). Retrieved September 22, 2023, from https://www.jstor.org/stable/25548130

A Benchmark- No Bureaucrat

The first benchmark is a case without a bureaucrat. The motivation for keeping this as a benchmark stems from the fact that ex-ante, it is not clear whether we need bureaucrats in such a model. Analysing a model without bureaucrats and comparing the results with the CAB and LAB case, I argue that the bureaucrats act as important strategic players and can potentially improve equilibrium outcomes for the citizen.

Citizen: For the benchmark case, I look at a model without bureaucrats. Since period 2 is a world with complete information, the citizen chooses whether to block the policy by comparing her payoffs.

$$U^C(a_2 = 0) = -p$$
$$U^C(a_2 = 1) = -c_2$$

Therefore, C's best response in period 2 is

$$a_2^* = \begin{cases} 1 & \text{if } c_2 < p \\ 0 & \text{if } c_2 \geq p \end{cases} \tag{19}$$

When $p = 0$, the best response for the citizen is to play $a = 0$. The citizen blocks a policy whenever the cost of blocking is lower than the loss from the policy. Assuming $c_2 < 1$ implies the policy will never be implemented, since the citizen always strictly prefers blocking the policy in period 2. However, $c_2 > 1$ implies that the citizen never blocks the policy in period 2.

In period 1, the citizens' beliefs are induced by the leader's strategy. Let the leader play $p = 1$ with probability q. The expected payoff of the citizen becomes:

$$\mathbb{E}[U^C(a_1 = 1)] = -c$$
$$\mathbb{E}[U^C(a_1 = 0)] = q[-c_2 \mathbb{1}_{\{c_2<1\}} - \mathbb{1}_{\{c_2 \geq 1\}}] - (1-q) \cdot 0 = -q \min(c_2, 1)$$

where $\mathbb{1}_{\{c_2<1\}}$ is an indicator function that is 1 when the condition is satisfied. If the citizen does not block the policy in period 1 ($a_1 = 0$), then the policy is enacted in period 2 ($p = 1$) with probability q. In period 2, either $c_2 < 1$ or $c_2 > 1$. Equation (19) gives us the corresponding best response for the citizen in both these cases. In period 2, $a_2 = 1$ and $a_2 = 0$ respectively. With probability $(1 - q)$, the policy is not enacted ($p = 0$) and the citizen prefers not to take any action in period 2.

The citizen's best response in period 1 depends on the relationship between $\{c_1, c_2, 1\}$. If $0 < c_1 < c_2 < 1$, then $\mathbb{E}[U^C(a_1 = 0)] = -qc_2$ and the citizen prefers to block in period 1 as long as $c_1 < qc_2$. Similarly, when $0 < c_1 < 1 < c_2$, then $\mathbb{E}[U^C(a_1 = 0)] = -q$, and the citizen prefers to block in period 1 as long as $c_1 < q$. Combining these, we have

$$a_1^* = \begin{cases} 1 & \text{if } c_1 < q\, min(c_2, 1) \\ 0 & \text{if } c_1 \geq q\, min(c_2, 1) \end{cases} \tag{20}$$

Leader: If the leader plays $p = 0$, i.e., sets $q = 0$, then from the best response of the citizen in period 1 (20), the citizen prefers not to block the policy as $c_1 > 0$. In period 2, $p = 0$ is observed by the citizen and, following the best response (19), the citizen prefers not to block in period 2 as well. The leader's payoff is $U^L(p = 0, a_1^* = 0, a_2^* = 0) = 0$. In comparison, if the leader plays $p = 1$, her expected payoff depends on the citizen's strategy and the parameter values.

Proposition A.1 (Benchmark)**.** *Consider the two-period game with a Leader (L), a Bureaucrat (B) and a Citizen (C), where the bureaucrat babbles. The equilibria of this game are:*

1. ***Status-Quo Equilibrium*** *There exists an equilibrium where the leader chooses $q^* = 0$, and the citizen never blocks: $a_1^* = 0, a_2^*(p = 0) = 0$.*
2. *If $c_1 < c_2 < 1$, an equilibrium exists where the leader chooses to enact the surprise policy $q^* = 1$ and the citizen blocks in period 1.*
3. *If $c_2 > 1 > c_1 > 0$, a mixed equilibrium exists, where the leader enacts the policy $p = 1$ with probability $q^* = c_1$, and the citizen never blocks the policy.*
4. *If $1 < c_1 < c_2$, an equilibrium exists where the leader always play $q^* = 1$ and the citizen never blocks.*

Proof. The proof of proposition A.1 for each case is as follows:

1. "Status-Quo Equilibrium": If $q^* = 0$, Citizen's best response in period 1 is not to block since $c_1 > 0$ (from equation (20)). In period 2, the best response to $p = 0$ from equation (19) is to play $a_2 = 0$. The leader also has no reason to deviate to $p = 1$, as it would not give her a higher payoff. Therefore, $\{q^* = 0, a_1^* = 0, a_2^* = 0\}$ is an equilibrium.
2. When $q^* = 1$ and $c_1 < 1$, the citizen's expected payoff from not blocking in period 1 is $-q \min(c_2, 1) = -c_2$, which is less than the payoff from blocking, $-c_1$. Therefore, the citizen blocks in period 1, and the leader receives a payoff of 0. The leader also doesn't have any profitable deviation, and hence $\{q^* = 1, a_1^* = 0, a_2^* = 0\}$ is an equilibrium.
3. When $0 < c_1 < 1 < c_2$, the leader makes the citizen indifferent between blocking or not blocking in period 1 by playing $q^* = c_1$. Using the assumption that an indifferent citizen decides not to block implies that the game goes to the second period. In period 2, the citizen does not block because $c_2 > 1$ (from equation (19)). The leader's expected payoff is $E[U^L] = q^* \cdot 1 + (1 - q^*) \cdot 0 = q^* = c_1$. Any deviation to $q > c_1$ will trigger blocking in period 1, decreasing the payoff to 0, and any

deviation to $q < c_1$ will lead to the expected payoff of the leader being lower than q^*. Therefore, when $0 < c_1 < 1 < c_2$, $\{q^* = c_1, a_1^* = 0, a_2^* = 0$ is an equilibrium.

4. When $1 < c_1 < c_2$, and q=1, the citizen's expected payoff in period 1 is $-min(c_2, 1) = -1$, which is higher than the payoff from blocking, $-c_1$. In period 2, the citizen will not block if $p = 1$ because $c_2 > 1$ (from equation (19)). The leader gets a payoff of 1 and has no incentive to deviate to any q¡1, which would give her a lower payoff. Therefore, $\{q^* = 1, a_1^* = 0, a_2^* = 0\}$ is an equilibrium when $1 < c_1 < c_2$.

□

B Alternate Cases

B.1 Benchmark

Proposition B.1 (Benchmark). *Consider the two-period game with a Leader (L), a Bureaucrat (B) and a Citizen (C), where the bureaucrat uses the babbling stategy* $\mathbb{P}[m = 1|p = 1] = \mathbb{P}[m = 1|p = 0] = \alpha \in [0, 1]$. *The equilibria of this game are:*

1. ***Status-Quo Equilibrium*** *There exists an equilibrium where the leader chooses* $q^* = 0$, *and the citizen never blocks:* $a_1^* = 0, a_2^*(p = 0) = 0$, *with the beliefs* $\mu_m(p = 1) = q^*\ \forall m \in \{0, 1\}$.

2. *If* $c_1 < c_2 < 1$, *an equilibrium exists where the leader chooses to enact the surprise policy* $q^* = 1$ *and the citizen blocks in period 1, with the beliefs* $\mu_m(p = 1) = q^*\ \forall m \in \{0, 1\}$.

3. *If* $1 < c_1 < c_2$, *an equilibrium exists where the leader always play* $q^* = 1$ *and the citizen never blocks, with the beliefs* $\mu_m(p = 1) = q^*\ \forall m \in \{0, 1\}$.

Proof. Since the bureaucrat is payoff-irrelevant, she can play any babbling strategy $\alpha \in [0, 1]$. The proof of proposition B.1 for each case is as follows:

1. "Status-Quo Equilibrium": If $q^* = 0$, Citizen's best response in period 1 is not to block since $c_1 > 0$ (from equation (20)). In period 2, the best response to $p = 0$ from equation (19) is to play $a_2 = 0$. The leader also has no reason to deviate to $p = 1$, as it would not give her a higher payoff. Therefore, $\{q^* = 0, a_1^* = 0, a_2^* = 0\}$ is an equilibrium.

2. When $q^* = 1$ and $c_1 < 1$, the citizen's expected payoff from not blocking in period 1 is $-q \min(c_2, 1) = -c_2$, which is less than the payoff from blocking, $-c_1$. Therefore, the citizen blocks in period 1, and the leader receives a payoff of 0. The leader also doesn't have any profitable deviation, and hence $\{q^* = 1, a_1^* = 0, a_2^* = 0\}$ is an equilibrium.

3. When $1 < c_1 < c_2$, and q=1, the citizen's expected payoff in period 1 is $-min(c_2, 1) = -1$, which is higher than the payoff from blocking, $-c_1$. In period 2, the citizen will not block if $p = 1$ because

$c_2 > 1$ (from equation (19)). The leader gets a payoff of 1 and has no incentive to deviate to any q¡1, which would give her a lower payoff. Therefore, $\{q^* = 1, a_1^* = 0, a_2^* = 0\}$ is an equilibrium when $1 < c_1 < c_2$.

□

B.2 CAB

Proposition B.2 (Citizen-Aligned Bureaucrat)**.** *Consider the two-period game with a Leader (L), a Citizen-Aligned Bureaucrat (B), and a Citizen (C). The Perfect Bayesian Equilibria of this game are:*

1. *When $c_2 < 1$, the following constitutes Perfect Bayesian Equilibria :*
   (a) ***Truthtelling Equilibrium**: There exists an equilibrium where the bureaucrat follows the strategy $\lambda^* = 1$, the leader mixes between enacting and not enacting a policy $q \in [0, 1]$ and the citizen blocks in period 1 upon receiving message $m = 1$, and does not block upon receiving message $m = 0$: $a_1^*(m = 0) = 0, a_1^*(m = 1) = 1$. In period 2, if the policy is enacted, the citizen blocks it: $a_2^*(p) = p \ \forall p$, and the beliefs of the citizen are $\mu_{m=1}(p = 1) = 1$, $\mu_{m=0}(p = 0) = 1$.*
   (b) ***Status-quo Equilibrium**: There exists an equilibrium where the leader never enacts the policy $q^* = 0$, λ^* is arbitrary (off-equilibrium path) but would equal 1 under reasonable beliefs, and the citizen blocks in period 1 after receiving message $m = 1$ and does not block after receiving $m = 0$: $a_1^*(m = 1) = 1$ (off-equilibrium path), $a_1^*(m = 0) = 0$, $a_2^*(p) = p$, with the beliefs $\mu_{m=1}(p = 1) = 1$ (off-equilibrium path), and $\mu_{m=0}(p = 0) = 1$.*
2. *When $1 < c_1 < c_2$, there exists an equilibrium where the bureaucrat follows the strategy $\lambda^* \in [0, 1]$, the leader strictly prefers to enact the policy $q^* = 1$, and the citizen does not block in either period: $a_t^* = 0 \ \forall t$, and the beliefs of the citizen are $\mu_{m=1}(p = 1) = 1$, $\mu_{m=0}(p = 1) = 1$.*

Case 1:$c_2 < 1$ When $m = 1$, the expected utility from blocking is always higher, as $c_1 > c_2$ by assumption. Therefore,

$$a_1^*(m = 1) = 1 \tag{21}$$

Similarly, when $m = 0$

$$a_1^*(m = 0) = \begin{cases} 1 & \text{if } c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q} c_2 \\ 0 & \text{if } c_1 \geq \frac{(1-\lambda)q}{(1-\lambda)q+1-q} c_2 \end{cases} \tag{22}$$

Note that as $q \to 0$. i.e., the leader chooses $p = 0$ with probability 1, the citizen never blocks in either period. When $\lambda \to 1$, i.e., the bureaucrat reveals the true p with probability 1, and the citizen believes the bureaucrat and decides not to block in period 1. Given the citizen's best response, the bureaucrat should not have a profitable deviation. When $p = 0$, the bureaucrat cannot increase her payoff by deviating to $m = 1$, which gives her a payoff of $-k_1$ in comparison to 0 when she sends $m = 0$. Therefore, $\mathbb{P}(m = 0|p = 0) = 1$ is optimal for the bureaucrat. When $p = 1$, the bureaucrat's expected utilities are as follows:

$$\mathbb{E}[U^B(m_1 = 1|p = 1)] = -k_1 \tag{23}$$

$$\mathbb{E}[U^B(m_1 = 0|p = 1)] = \begin{cases} -k_1 & \text{if} \quad c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2 \\ -k_2 & \text{otherwise} \end{cases} \tag{24}$$

Since $k_1 < k_2$, the bureaucrat is indifferent between messages if $c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2$, and strictly prefers to send $m = 1$ given $p = 1$, i.e., $\lambda = 1$ when $c_1 > \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2$. Now, the leader's expected payoff from choosing $p = 0$ is 0. If the leader chooses $p = 1$ and the citizen blocks in period 1, the leader gets 0. Whereas, if the citizen blocks in period 2, the leader gets $-r$. Since $c_2 < 1$, if B sends $m = 1$ then C plays $a_1^*(m = 1) = 1$ and L gets a payoff of 0. If B sends $m = 0$, then if $c_1 < \mu_{m=0}(p = 1)c_2$, C plays $a_1^*(m = 0) = 1$. Therefore, the leader's expected utility is

$$\begin{aligned}\mathbb{E}[U^L(p = 1)] &= \lambda \cdot 0 + (1 - \lambda)[\mathbb{1}_{\{c_1 < \mu_{m=0}(p=1)c_2\}} \cdot 0 + \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)c_2\}}(-r)] \\ &= -(1 - \lambda)r\mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)c_2\}}\end{aligned}$$

Therefore, the leader's best response is given by

$$q^*(m = 0) = \begin{cases} 0 & \text{if } -(1 - \lambda)\, r\, \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)c_2\}} < 0 \\ [0, 1] & \text{if } -(1 - \lambda)\, r\, \mathbb{1}_{\{c_1 > \mu_{m=0}(p=1)c_2\}} = 0 \end{cases} \tag{25}$$

Proof.

- **Truthtelling Equilibrium**: When $\lambda^* = 1$, i.e., the bureaucrat reports truthfully, the posterior belief of the citizen becomes $\mu_0(p = 1) = 0$. From Citizen's best response in period 1 (21) and (22), the citizen plays $a_1^*(m = 0) = 0, a_1^*(m = 1) = 1$. Since $c_2 < 1$, this implies in period 2, citizen blocks the policy $a_2^*(p = 1) = 1$, $a_2^*(p = 0) = 0$. Given this, the bureaucrat has no incentive to deviate, i.e., she strictly prefers to send $m = 1$ given $p = 1$ (from (33)). From the leader's best response (25), the leader is indifferent between setting $p = 0$ or $p = 1$, i.e., $q^* \in [0, 1]$ because both choices give her the same payoff.

- **Status-quo Equilibrium**: If the leader sets $q^* = 0$, then the posterior belief of the citizen becomes $\mu_0(p = 1) = 0$. Given this, the citizen strictly prefers not to block when $m = 0$ and in period 1 ((22)). Since the leader never chooses $p = 1$, the bureaucrat's strategy for $p = 1$ is off the equilibrium path. Recall that the bureaucrat strictly prefers to send $\lambda^* = 1$ whenever $c_1 \geq \mu_0(p = 1)c_2$. Since in this case $c_1 \geq 0$, which is true by assumption, it is plausible to believe the bureaucrat's best response is to play $\lambda^* = 1$, i.e., truthfully reveal the policy. And hence, the citizen chooses to block the policy in period 1 ($a_1^*(m = 1) = 1$) and the posterior belief of the citizen will be $\mu_{m=1}(p = 1) = 1$. Since no one has an incentive to deviate and the beliefs are consistent, this strategy constitutes a PBE.

Case 2:$1 < c_1 < c_2$

Since $c_2 > 1$, using similar arguments as above, the best response of the citizen in period 1 will be

$$a_1^*(m = 1) = 0 \tag{26}$$

Since $c_1 > 1$ and $\mu_0(p = 1) \in [0, 1]$, when $m = 0$

$$a_1^*(m = 0) = 0 \tag{27}$$

Step 3.2: Bureaucrat's Decision when $1 < c_1 < c_2$

When $p = 0$, the bureaucrat always prefers message $m = 0$ and has no profitable deviation as in the previous case. When $p = 1$, the bureaucrat's expected utilities are as follows.

$$\mathbb{E}[U^B(m = 1|p = 1)] = -1$$
$$\mathbb{E}[U^B(m = 0|p = 1)] = -1$$

So the bureaucrat is indifferent between sending $m = 1$ and $m = 0$ when $c_1 > 1$.

Step 4.2: Leader's Decision when $1 < c_1 < c_2$

Since in this case, the citizen has a dominant strategy to not block the policy, she ignores the messages from the bureaucrat. The leader, therefore, in equilibrium wants to enact the policy, given that implementing it gives her a payoff of 1, which is higher than what she gets when she does not enact the policy.

$$\mathbb{E}[U^L(p = 1)] = 1$$

Therefore, the leader's best response is given by

$$q^* = 1 \tag{28}$$

Step 5.2: Equilibrium when $1 < c_1 < c_2$

- **Implementation Equilibrium**: From Citizen's best response in period 1 (26), (27), the citizen plays $a_t^* = 0 \ \forall t$. Given this, the bureaucrat is indifferent between sending message $m = 1$ or $m = 0$, therefore $\lambda^* \in [0, 1]$. From the leader's best response (28), the leader strictly prefers setting $p = 1$, i.e., $q^* = 1$.

□

B.3 LAB

Proposition B.3 (Leader-Aligned Bureaucrat)**.** *Consider the two-period game with a Leader (L), a Leader-Aligned Bureaucrat (B), and a Citizen (C). The Perfect Bayesian Equilibria of this game are:*

1. *When $c_2 < 1$, there exists an equilibrium where the leader never enacts the policy $q^* = 0$, λ^* is arbitrary (off-path), and the citizen does not block in period 1: $a_1^*(m = 0) = 0$, $a_1^*(m = 1) = 1$ and in period 2 $a_2^*(p = 0) = 0$, $a_2^*(p = 1) = 1$, with the beliefs $\mu_{m=0}(p = 0) = 1$, $\mu_{m=1}(p = 1) = 1$ (off-path).*

2. *When $1 < c_1 < c_2$, there exists an equilibrium where the leader always enacts the policy $q^* = 1$, $\lambda^* \in [0, 1]$, and the citizen does not block in period 1: $a_1^*(m) = 0 \forall m$ and in period 2 $a_2^*(p) = 0$, with the beliefs $\mu_m(p = 1) = 1 \forall m$.*

Proof. Recall that the citizen uses Bayes Rule to update her beliefs after receiving the bureaucrat's message, i.e., $\mu_m(p) \equiv \mathbb{P}(p|m) = \frac{\sigma_B(m,p=1)\mathbb{P}(p)}{\sum_{p' \in \{0,1\}} \sigma_B(m,p')\mathbb{P}(p')}$. Suppose the bureaucrat follows the following strategy $\sigma_B(m = 1, p = 1) = \mathbb{P}[m = 1|p = 1] = \lambda \ \ \& \ \ \sigma_B(m = 0, p = 0) = \mathbb{P}[m = 0|p = 0] = 1$.Then the citizens' updated beliefs will be:

$$\mu_{m=1}(p = 1) = 1 \quad \& \quad \mu_{m=0}(p = 1) = \frac{(1 - \lambda)q}{(1 - \lambda)q + 1 - q}$$
$$\mu_{m=1}(p = 0) = 0 \quad \& \quad \mu_{m=0}(p = 0) = \frac{1 - q}{1 - q + (1 - \lambda)q}$$

Now, the proof proceeds in the following steps:

Step 1: Citizen's Period 2 Decision

In period 2, the citizen observes the policy p and decides whether to block it by comparing her utilities. Using a similar methodology as in the Citizen-Aligned Bureaucrat case, the best Citizen's best response as

$$a_2^* = \begin{cases} 1 & \text{if } c_2 \leq 1 \\ 0 & \text{if } c_2 > 1 \lor p = 0 \end{cases} \tag{29}$$

Step 2: Citizen's Period 1 Decision

In period 1, the expected payoffs of the citizen given the message are

$$\mathbb{E}[U^C(a_1 = 1)|m = 1] = -c_1$$
$$\mathbb{E}[U^C(a_1 = 0)|m = 1] = -\mu_{m=1}(p = 1)[-c_2\mathbf{1}_{\{c_2\leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -\min(c_2, 1)$$
$$\mathbb{E}[U^C(a_1 = 1)|m = 0] = -c_1$$
$$\mathbb{E}[U^C(a_1 = 0)|m = 0] = \mu_{m=0}(p = 1)[-c_2\mathbf{1}_{\{c_2\leq 1\}} - \mathbf{1}_{\{c_2>1\}}] = -\mu_{m=0}(p = 1)\min(c_2, 1)$$

Once again, consider the two cases separately: $c_2 < 1$ and $c_2 > 1$. While in the former case, the citizen prefers to block the policy in period 2, whereas in the latter case citizen is better off by allowing the policy to be implemented.

Case 1:$c_2 < 1$

When $m = 1$, the expected utility from blocking is always higher, as $c_1 > c_2$ by assumption. Therefore,

$$a_1^*(m = 1) = 1 \tag{30}$$

Similarly, when $m = 0$

$$a_1^*(m = 0) = \begin{cases} 1 & \text{if } c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2 \\ 0 & \text{if } c_1 \geq \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2 \end{cases} \tag{31}$$

Note that as $q \to 0$. i.e., the leader chooses $p = 0$ with probability 1, the citizen never blocks in either period. When $\lambda \to 1$, i.e., the bureaucrat reveals the true p with probability 1, and the citizen believes the bureaucrat and decides not to block in period 1 if the message is $m = 0$.

Step 3.1: Bureaucrat's Decision when $c_2 < 1$

In this case, the bureaucrat wants the policy to be implemented. However, maintenance costs (k_1, k_2) ensure that the bureaucrat doesn't always try to deceive the citizen. When $p = 0$, the bureaucrat cannot increase her payoff by deviating to $m = 1$, which gives her a payoff of $-k_1$ in comparison to 0 when she sends $m = 0$. Therefore, $\mathbb{P}(m = 0|p = 0) = 1$ is optimal for the bureaucrat. When $p = 1$, the bureaucrat's expected utilities are as follows:

$$\mathbb{E}[U^B(m_1 = 1|p = 1)] = -k_1 \tag{32}$$

$$\mathbb{E}[U^B(m_1 = 0|p = 1)] = \begin{cases} -k_1 & \text{if} \quad c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2 \\ -k_2 & \text{otherwise} \end{cases} \tag{33}$$

Since $k_1 < k_2$, the bureaucrat is indifferent between messages if $c_1 < \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2$, and strictly prefers to send $m = 1$ if $p = 1$, i.e., $\lambda = 1$ when $c_1 \geq \frac{(1-\lambda)q}{(1-\lambda)q+1-q}c_2$.

Step 4.1: Leader's Decision when $c_2 < 1$

Now, the leader's expected payoff from choosing $p = 0$ is 0. If the leader chooses $p = 1$ and the citizen blocks in period 1, the leader gets 0. Whereas, if the citizen blocks in period 2, the leader gets $-r$. Since $c_2 < 1$, if B sends $m = 1$ then C plays $a_1^*(m = 1) = 1$ and L gets a payoff of 0. If B sends $m = 0$, then if $c_1 < \mu_{m=0}(p = 1)c_2$, C plays $a_1^*(m = 0) = 1$. Therefore, the leader's expected utility is

$$\begin{aligned}\mathbb{E}[U^L(p = 1)] &= \lambda \cdot 0 + (1 - \lambda)[\mathbb{1}_{\{c_1<\mu_{m=0}(p=1)c_2\}} \cdot 0 + \mathbb{1}_{\{c_1>\mu_{m=0}(p=1)c_2\}}(-r)] \\ &= -(1 - \lambda)r\mathbb{1}_{\{c_1>\mu_{m=0}(p=1)c_2\}}\end{aligned}$$

Therefore, the leader's best response is given by

$$q^*(m = 0) = \begin{cases} 0 \quad \text{if } -(1-\lambda)\, r\, \mathbb{1}_{\{c_1>\mu_{m=0}(p=1)c_2\}} < 0 \\ [0, 1] \quad \text{if } -(1-\lambda)\, r\, \mathbb{1}_{\{c_1>\mu_{m=0}(p=1)c_2\}} = 0 \end{cases} \tag{34}$$

Step 5.1: Equilibrium when $c_2 < 1$

- **Status-quo Equilibrium**: If the leader sets $q^* = 0$, then the posterior belief of the citizen becomes $\mu_0(p = 0) = 1$. Given this, the citizen strictly prefers not to block when $m = 0$ in period 1 ((31)). Since the leader never chooses $p = 1$, the bureaucrat's strategy for $p = 1$ is off the equilibrium path. Recall that the bureaucrat strictly prefers to send $\lambda^* = 1$ whenever $c_1 \geq \mu_0(p = 1)c_2$. Since in this case $c_1 \geq 0$, which is true by assumption, it is plausible to believe the bureaucrat's best response is to play $\lambda^* = 1$, i.e., truthfully reveal the policy. And hence, the citizen chooses to block the policy in period 1 ($a_1^*(m = 1) = 1$) and the posterior belief of the citizen will be $\mu_{m=1}(p = 1) = 1$. Since no one has an incentive to deviate and the beliefs are consistent, this strategy constitutes a PBE.

The proof for the status quo equilibrium is the same for other cases as well. Therefore, I refrain from writing it down explicitly for the other cases.

Case 2:$1 < c_1 < c_2$

- Since $c_1 > 1$, from the citizen's best response in period 1 and 2, she never blocks: $a_t^* = 0 \ \forall t$. Given this, the bureaucrat is indifferent between sending message $m = 1$ or $m = 0$, therefore $\lambda^* \in [0, 1]$. From the leader's best response, the leader strictly prefers setting $p = 1$, i.e., $q^* = 1$. Therefore, this is an equilibrium.

□