Essays in Political Economy

Sylta Cornils

Abstract

This thesis studies the effect of rent-seeking behavior on policy, regulation and entrepreneurial activity.

The first chapter investigates the impact of corruption on risk. Corruption distorts entrepreneurial innovation and production. Both innovation and production can fail to result in a positive profit implying that they entail risk. Corruption can distort this risk in different directions because innovation and production differ in the amount of information available to entrepreneurs. In production corruption increases the risk of activities chosen leading to more output volatility. In innovation corruption discourages innovators from taking risk leading to less innovation. This results whenever the entrepreneurs' exact project choice is not known to the corrupt public official accepting the bribe. In this case the entrepreneur bases his project choice on the information about the project's success that is available when paying the bribe. If only scant information is available, as in the case of innovation, less risky project choices are made. If a lot of information is available, as in the case of production, the entrepreneur chooses riskier projects.

The second chapter sheds light on the optimal level of regulation in the presence of corruption. A higher level of regulation leads to both a better allocation of goods and a higher level of corruption. A corrupt bureaucrat in charge of distributing goods does not follow the allocation rules laid down by the government. Instead he offers agents to circumvent the official regulation if they pay a bribe. The bureaucrat either demands a low bribe, all agents pay, or a high bribe that only agents valuing the good highly pay. In terms of allocation the government prefers the second approach. Because the agents' willingness to pay higher bribes increases as the level of official regulation raises, an increase in official regulation can improve the allocation if there is corruption.

The third chapter studies how interest groups with a large membership can influence policy and which policy position these groups take. Voters form interest groups in order to influence a politician's policy choice by conditioning the group's voting behavior on the observed policy. A politician who wants to maximize the probability of reelection chooses the groups's policy if the group is sufficiently large to compensate for votes lost from unorganized voters. If groups are formed endogenously, in a symmetric equilibrium with two groups, the groups' positions are sufficiently moderate to be chosen by the politician and sufficiently extreme to benefit from a change in policy. Groups become more extreme the higher the cost of founding and the smaller the share of politicians interested solely in being reelected.

Chapter 1

Corruption and Risk: How Corruption Raises and Reduces Risk

1 Introduction

Corruption is a widespread problem in many countries and is thought to decrease both innovation and production. In the literature, the effects of corruption on innovation and production are usually studied separately. This approach neglects two important issues. Firstly, innovation and production share the common characteristic of entailing risk that results from the inherent possibility of innovation and production failing and leaving the entrepreneur without a positive profit. Secondly, a change in corruption might influence innovation and production differently such that a reduction in the level of corruption results in a potential tradeoff between encouraging production or innovation. This paper provides a unified framework for the impact of corruption on innovation and production by focusing on the level of risk, that is the chance of a project failing, tolerated by an entrepreneur. Three main findings emerge. Firstly, corruption leads to more output volatility by increasing the level of risk in production. Secondly, corruption discourages innovation by lowering the level of risk an innovator is willing to bear. Taken together, this implies that corruption influences risk in a particularly bad way. Lastly, there is a potential tradeoff between increasing production or innovation where a fall in the acceptance of corruption can increase production but decrease innovation.

One of the earliest, if not the only, statement relating corruption as a form of rent-seeking behavior to both innovation and production can be attributed to Murphy, Shleifer and Vishny (1993). They claim that "rent-seeking, particularly public rent-seeking by government officials, is likely to hurt innovative activities more than everyday production". The idea behind this claim is that the innovator bears the cost of the project alone but has to share the return. This reasoning relies on the assumption that an entrepreneur faces a binary choice of whether to innovate, that is either innovation takes place or nothing is undertaken. In this case a bribe obviously discourages innovation and production because the entrepreneur does not have to pay the bribe if no innovation or production takes place. There is, however, a second dimension to the entrepreneur's project choice and that is which project an entrepreneur chooses when deciding among several projects which vary in the level of risk. Two factors govern the entrepreneur's project choice. The first factor is whether a corrupt bureaucrat can infer the project choice such that the bribe demand depends on it. Whenever riskier projects result in a higher profit upon success, this higher profit translates into a higher bribe which discourages choosing riskier projects. If, however, the bureaucrat does not know the project choice, the bribe demand is independent of the project choice and therefore taken as given by the entrepreneur. The second factor is the amount of information the entrepreneur has about the realization of the project's risk when paying the bribe. The more information he has, the more often he can refrain from paying the bribe if the riskier project fails. This, in turn, increases the expected profit of the risky choice making it more worthwhile. The entrepreneur has more information available about a project when he had more time to gather information. Compared to innovation, the entrepreneur also has more information available in the case of production because he can draw on past experience. The different effect of corruption on risk for innovation and production therefore depends on the information the entrepreneur has about failure and success of these projects.

To illustrate the effect of the two dimensions of information, the one the bureaucrat has and the one the entrepreneur has, on the bureaucrat's bribe and the entrepreneur's project choice, we look at four different examples. A recent corruption scandal involved the Brazilian Odebrecht Organization which paid bribes in order to secure construction contracts. There are two things to note here. Firstly, the company pays the bribe before production starts, implying that only little information about the project is available at that time. Secondly, the company has a long track record of finished buildings such that the bureaucrat can tailor the bribe he accepts to the usual project and profit the company makes. The bureaucrat therefore asks for a higher bribe if the entrepreneur makes higher profits. This discourages the entrepreneur from choosing risky projects with high returns. He chooses safer projects instead resulting in a decrease in risk.

As a second example, consider the case where the bureaucrat has a lot of information about the entrepreneur, for example from past behavior, allowing him to tailor the bribe, and the entrepreneur, in turn, has a lot of information about the project's success. When transporting goods, the entrepreneur faces no more uncertainty regarding the transport when the goods reach the final destination. Illustrating this situation, the German logistics company Schenker paid bribes at the port of St. Petersburg in Russia to get Ford's car parts through customs more quickly (Ott 2015). Because the bureaucrat asks for higher bribes if higher profits are attained by the entrepreneur, but the entrepreneur pays the bribe conditional on a successful transport, project choices can be optimal.

A different scenario arises when the entrepreneur's past profit is not observable. We again consider the case of transportation and crossing a border. At a remote Vietnamese border, the custom official might find it difficult to exactly determine the goods supposed to cross the border and simply ask for the same bribe from everyone such that "'every relevant government office would get 1m baht [£20,000, in bribes]'" (Hodal 2013). Similarly, the person crossing the border might not know whether the goods will ever reach their final destination. Hence, in this case both the person paying and the person accepting the bribe have limited information. Because the entrepreneur does not know whether the goods will arrive, but already has to pay the bribe, he either decides to abandon the project or to choose a less risky project.

As the fourth and final example, consider the case of the bureaucrat possessing scant information about the entrepreneur, while the entrepreneur possesses a lot of information about the project. The entrepreneur is well informed about the project when the project is already underway or finished. Street vendors in Mumbai, for example, pay high bribes to policemen to keep their goods and booth (Mulye 2014). They pay after they have already begun selling such that they can infer their daily profit. Similarly, an Indian shop owner reports that he faced bribe demands when he was already running his shop and that "the municipal corporation demolished my shop several times as I did not have the money to pay bribe" (Dhingra 2012). Again, he was faced with a bribe demand after acquiring an idea about whether the shop was successful. Similarly, the last two floors of Rana Plaza in Bangladesh were illegally added and only then was a permit acquired (The Guardian 2013). In this case, investigations after the building's collapse revealed that the material was inferior and the ground swampy (The Guardian 2013). By only paying the bribe after the factory was successfully build, the entrepreneur could condition the bribe payment on the successful building process. Additionally, using adequate material would have been more expensive reducing the profit but making a successful building process more likely. Of these two options, the entrepreneur chose the riskier one.

These four examples differ in how much information the entrepreneur has about the project when paying the bribe. The second difference is how much information the bureaucrat has about the entrepreneur's project choice. In the model, the first dimension is captured by assuming that the entrepreneur observes a signal of different level of precision about the project's success. The second dimension is captured by assuming that either the bureaucrat can observe the entrepreneur's choice or vice versa.

The intuition of the model is described in the following. To enter business, an entrepreneur can choose between two projects (or strategies, investments, technologies). One project, the safe project, always results in a small profit. The other project, the risky project, can fail, but has a higher return than the safe project if successful. If there is no corruption, the entrepreneur chooses the project with the higher expected return. Now suppose that there is a corrupt bureaucrat who demands a bribe for the permit to start production and reap the project's return. If the bureaucrat can condition his bribe demand on the entrepreneur's project choice, higher project returns lead to higher bribe demands. This decreases the payoff of choosing the risky project such that project choices are inefficiently safe. If, on the other hand, the bureaucrat does not condition his bribe on the project choice, the extent of the distortion depends on the amount of information the entrepreneur has when paying the bribe. After choosing the project, but before paying the bribe, the entrepreneur observes a signal indicating whether the risky project will be successful. For production, the signal is very precise revealing a lot of information. For innovation, the signal is very imprecise revealing only a little information about the project's success. In the first case, the entrepreneur can condition payment of the bribe on the project's success and only pays the bribe if the project is very likely successful. By not paying the bribe if the project fails, the entrepreneur shifts some risk on the bureaucrat inducing the entrepreneur to take on excessive risk. In the second case, the entrepreneur pays the bribe even if the project is unsuccessful, implying that the entrepreneur has to bear the risk of the project failing. This results in the entrepreneur choosing inefficiently safe projects.

Using data from the World Bank Enterprise Survey I test the model's predictions regarding corruption and risk empirically. I find that corruption impacts the amount of risk, firms assume. As predicted by the model, the change in risk depends on the levels of information, the firm and the bureaucrat have.

By showing that the corruption-induced distortion depends on whether the entrepreneur or the bureaucrat observes the other agent's strategy, this paper adds to the literature on the organization of corruption. Other considerations in this branch of the literature include the number of bureaucrats working (Amir and Burr 2015) and the degree of centralization of corruption (Blackburn and Forgues-Puccio 2009).

Few theoretical papers relating corruption to risk or volatility exist. The most relevant to this paper include Søreide (2009), showing that higher risk-aversion may increase bribery, and Célimène, Dufrénot, Mophou and N'Guérékata (2016), showing how tax corruption can increase volatility if the evaded money is invested in private production which is more volatile than the public one. Leung, Tang and Groenewold (2006) show how rent-seeking behavior can increase growth volatility. In their model, firms choose the optimal amount of rent-seeking and, following from an assumption about the growth process, the more rent-seeking there is, the more volatile the growth process.

The idea that underlying parameters drive both corruption and economic outcomes was pioneered by Bliss and Di Tella (1997). They show that the same parameters influence the level of both corruption and competition. This paper shows that underlying economic conditions can lead to a high level of both bribery and output, with no causal relation between the latter two.

Lastly, this paper extends the theoretical literature modeling how corruption results in inefficient outcomes by adding inefficient levels of risk to the list of inefficiencies. It has been shown that bribery leads to misallocation of permits (Cadot 1987; Ahlin and Bose 2007), delayed issuance of licenses (Ahlin and Bose 2007), entrepreneurs deciding not to apply for permits (Yoo 2008) as well as firms delaying entry (Choi and Thum 2003). Harstad and Svensson (2011) show that bribes decrease investments by firms. The most closely related, in that it also studies technology choices, is the paper by Choi and Thum (2004). They show that firms faced with uncertain future bribe demands choose technologies with inefficiently high operating costs and inefficiently low fixed costs. This paper extends this literature by adding excessively high or low levels of risk to the list of inefficiencies resulting from corruption in the context of licensing.

The next two sections introduce and analyze the model for static and repeated games. Then, the effect of corruption on output and its volatility is investigated under different economic conditions. Thereafter, I test the model's predictions empirically. The last section concludes.

2 The model

The economy consists of one bureaucrat, B, and one firm, F. The firm can choose between two production technologies: safe and risky. The safe technology gives a certain output of r while the risky technology results in an output of $R > r$ with probability p and an output of 0 with probability $1 - p$.

The bureaucrat demands a bribe b to maximize his income taking into account the probability of being caught and punished for corruption. The detection probability depends on the size of the bribe relative to the firm's return and the government's detection efforts, $\pi > \frac{1}{2}$. This specification captures that it is easier to observe a bribe if almost all the firm's profit is taken away than if only a small fraction is used for bribing. If caught, the bureaucrat loses all his income attained by accepting bribes.

After choosing the risky project, the firm observes a signal s about the risky project's success. The signal can take on two values, failure and success of the project, $s \in \{0, 1\}$. The signal is correct with probability $1 - \varepsilon$, where $\varepsilon \in \left[0, \frac{1}{2}\right]$, such that ε measures the imprecision of the signal.[1] A very imprecise signal, a high value of ε, corresponds to an innovative project, while a very precise signal, a low value of ε, corresponds to everyday production. After observing the signal, the firm can either continue with the project and pay the bribe or discontinue the project and not pay the bribe. The crucial assumption is that the firm cannot switch from the risky to the safe project after observing the signal.

We assume that the probability of success p is private information of the firm. The bureaucrat only knows that it follows a certain distribution, $p \sim f(p)$ on $[0, 1]$. Because of this asymmetric information, the timing of the bureaucrat's bribe demand and the firm's technology choice matters. While the firm can reveal information, the bureaucrat generally cannot.

3 Solution without corruption

In this section the firm's project choice in the absence of corruption is derived. We will see later that corruption distorts the firm's optimal behavior derived here. A firm chooses the risky project if its success probability of the risky project is so high that the expected value of the risky project is higher than that of the safe project. Figure 1 depicts the firm's choice without corruption. If the firm observes signal $s = 1$ ($s = 0$), indicating the project's success (failure), the firm reaches information set I_1 (I_0). At both information sets the firm chooses to continue with the project. The signal's precision ε is not important because production is costless. Therefore, there is no distortion from firms not continuing production. Firms choose the safe project if the expected profit of the safe project, $\mathbb{E}(safe) = r$, is higher than that of choosing the risky project, $\mathbb{E}(risky) = p\varepsilon R + p(1 - \varepsilon)R = pR$. The firm

[1] Then, $prob(s = 1|success) = 1 - \varepsilon$, $prob(s = 0|success) = \varepsilon$, $prob(s = 1|failure) = \varepsilon$ and $prob(s = 0|failure) = 1 - \varepsilon$. The probability of observing a signal of success is $p(1 - \varepsilon) + (1 - p)\varepsilon$. A very precise signal, $\varepsilon = 0$, corresponds to the case of ex post bribing (the firm pays the bribe after the project's risk realized), while a very imprecise signal, $\varepsilon = 1/2$, corresponds to the case of ex ante bribing (the firm pays the bribe before the project's risk realized).

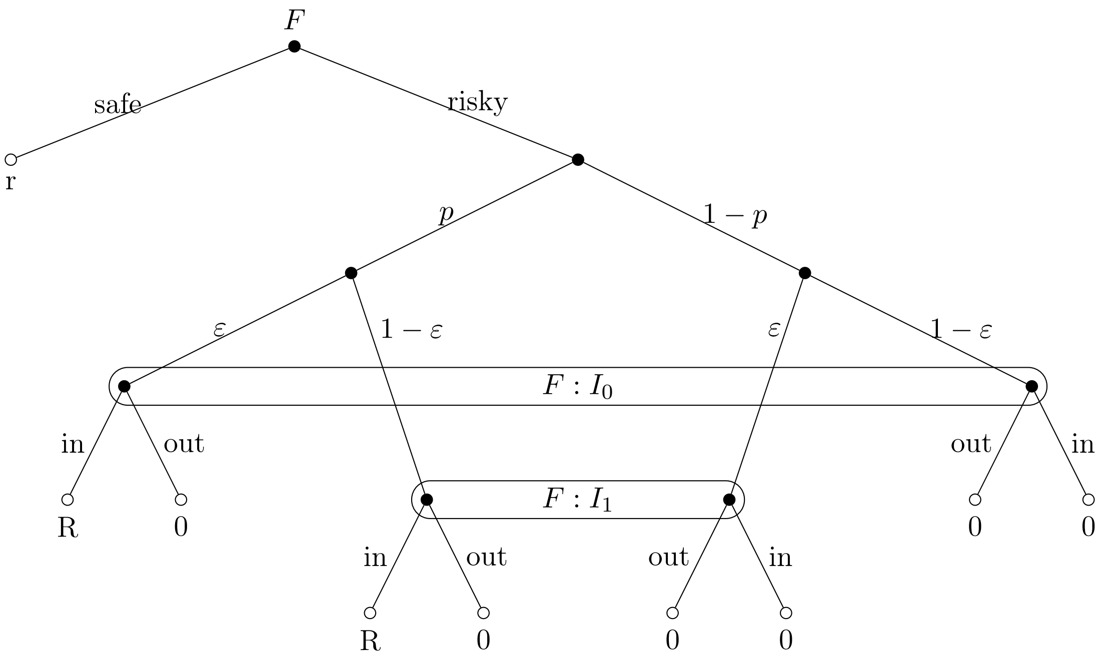


Figure 1: No corruption

compares expected output of both projects and decides to choose the risky project if indifferent.

Lemma 1. *If there is no corruption, firms choose the safe project if* $\frac{r}{R} > p$*.*

Whenever firms with a success probability of $p \geq \frac{r}{R}$ choose the risky project, project choices are efficient.

4 The model: Static game

In this section, we derive the intuition for why it matters whether the firm's or the bureaucrat's strategy is observable to the other agent. In order to do this, we assume two different timelines, where either the firm or the bureaucrat chooses a strategy first which is then observed by the other agent. First, we consider the case of the bureaucrat observing the firm's project choice before choosing the bribe. Second, we look at the case of the firm observing the bribe before choosing its project.

1. (a) the firm chooses the technology
   (b) after observing the signal about the project's success the firm decides whether to pay the bribe
   (c) the bureaucrat, observing the firm's choice, sets the bribe

   (d) the firm pays the bribe and production takes place

2. (a) the bureaucrat sets the bribe
   (b) the firm, observing the bribe demand, decides on the project
   (c) after observing the signal about the project's success the firm decides whether to pay the bribe
   (d) if the firm pays the bribe, production takes place

Definition 1. *In equilibrium, the firm derives beliefs using Bayes' rule wherever possible, the bureaucrat's strategy is optimal given the firm's strategy and the firm's strategy is optimal given its beliefs and the bureaucrat's strategy.*

4.1 Technology chosen before bribe

This section shows that in most of the cases firms make inefficiently safe project choices if the bureaucrat can condition the bribe on the project choice. The only exception to this observation is the case of a perfectly precise signal which leads to the firm choosing the efficient project. Because the bribe depends on the firm's choice, the firm faces a higher bribe demand if it chooses the risky project. If the signal is perfectly precise, the firm only has to pay the higher bribe demand if the project is successful leading to the efficient project choice.

In order to solve for the equilibrium of the game, we first determine the bureaucrat's optimal bribe demand if the firm's project choice is observable. Let $\rho = r, R$ denote the return such that the detection probability is given by $\pi \frac{b}{\rho}$. If the firm can pay the bribe, the expected bribe payment is given by

$$B(b) = b - \pi \frac{b}{\rho} b. \tag{1}$$

Because a bribe cannot be higher than the firm's return, bribes can be expressed as a share α of the return. The bureaucrat chooses a bribe $b = \alpha\rho$ where α maximizes the bureaucrat's income. If the return ρ is observable, the share α does not depend on the return.

Lemma 2. *If the firm's profit is observable, the bureaucrat asks for a small bribe αr if the safe project is chosen and for a high bribe αR if the risky project is chosen where*

$$\alpha = \frac{1}{2\pi}.$$

Proof. Equation (1) can be rewritten as

$$B = \alpha\rho - \pi\frac{\alpha\rho}{\rho}\alpha\rho = \alpha\rho - \pi\rho\alpha^2$$

From the first-order condition, we get α. The high bribe is higher than the safe return if $\pi < \frac{R}{2r}$. □

The game in extensive form with the firm as first mover is summarized in figure 2. If the firm observes a signal indicating success, $s = 1$, it reaches information set I_1, but does not know whether the project will be successful. If, on the other hand, the firm observes a signal indicating failure, $s = 0$, it reaches information set I_0.

The firm makes two decisions: first, whether to choose the safe or risky project and second, whether to pay the bribe after observing the signal corresponding to the risky project. We first derive which firms pay the bribe when choosing the risky project and then determine the firm's optimal project choice.

Taking as given a risky project choice and a high bribe demand, we look at the firm's behavior after observing the signal. At information set I_0, after observing a signal $s = 0$, the firm can decide between paying or not paying the bribe. Expected payoffs are given by

$$\mathbb{E}(pay|I_0) = \frac{p\varepsilon}{p\varepsilon + (1-p)(1-\varepsilon)}(R - \alpha R) + \frac{(1-p)(1-\varepsilon)}{p\varepsilon + (1-p)(1-\varepsilon)}(-\alpha R) \tag{2}$$

$$\mathbb{E}(not|I_0) = 0 \tag{3}$$

The firm chooses to pay the bribe if the expected payoff of doing so is higher than not paying the bribe, $\mathbb{E}(pay|I_0) \geq \mathbb{E}(not|I_0)$. Firms choose to pay the bribe if

$$p \geq \frac{(1-\varepsilon)\alpha}{\varepsilon(1-\alpha) + (1-\varepsilon)\alpha} \equiv p_0. \tag{4}$$

Similarly, at information set I_1, after observing signal $s = 1$, the firm can decide between paying and not paying the bribe. Expected payoffs are given by

$$\mathbb{E}(pay|I_1) = \frac{p(1-\varepsilon)}{p(1-\varepsilon) + (1-p)\varepsilon}(R - \alpha R) + \frac{(1-p)\varepsilon}{p(1-\varepsilon) + (1-p)\varepsilon}(-\alpha R) \tag{5}$$

$$\mathbb{E}(not|I_1) = 0 \tag{6}$$

The firm pays the bribe if $\mathbb{E}(pay|s = 1) \geq \mathbb{E}(not|s = 1)$ which holds if its success

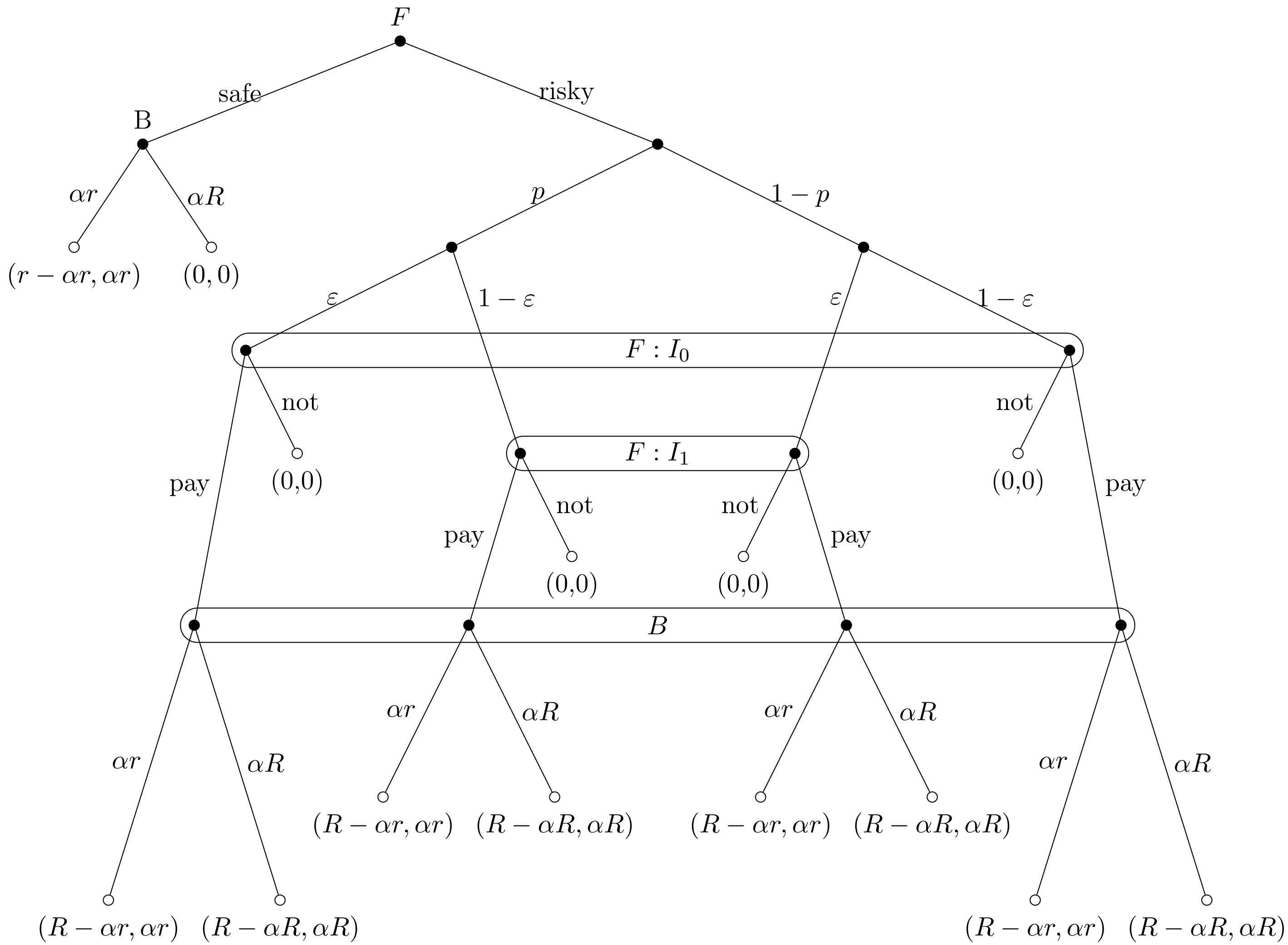


Figure 2: Technology chosen before bribe

probability is sufficiently high.

$$p \geq \frac{\varepsilon\alpha}{(1-\varepsilon)(1-\alpha)+\varepsilon\alpha} \equiv p_1 \tag{7}$$

There are three groups of firms. Firms either pay the bribe after any signal, after no signal or only after a positive signal. The profit of choosing the safe project is the same for all firms and given by $\mathbb{E}(safe) = (1-\alpha)r$. The expected profit of choosing the risky project depends on the firm's bribe payment decision and is given by

$$\mathbb{E}(risky) = \begin{cases} 0 \text{ if } p < p_1 \\ p(1-\varepsilon)R - \alpha R((1-p)\varepsilon + p(1-\varepsilon)) \text{ if } p_1 \leq p < p_0 \\ (p-\alpha)R \text{ if } p_0 \leq p \end{cases} \tag{8}$$

Firms with $p < p_1$ always choose the safe project because $\mathbb{E}(risky) = 0 < \mathbb{E}(safe)$. Firms with $p_1 \leq p < p_0$ choose the risky project if $\mathbb{E}(risky) \geq \mathbb{E}(safe)$ which holds if

$$p \geq \frac{r(1-\alpha)+\alpha R\varepsilon}{R((1-\varepsilon)(1-\alpha)+\varepsilon\alpha)} \equiv p_m. \tag{9}$$

Firms with $p_0 \leq p$ pay the bribe both at I_0 and I_1 and choose the risky project if $\mathbb{E}(risky) \geq \mathbb{E}(safe)$ which holds if their success probability is sufficiently high.[2]

$$p \geq \frac{(1-\alpha)r + \alpha R}{R} \equiv p_h \tag{10}$$

There are a number of different cutoff-values. It can be shown that some are always larger than others.

Lemma 3. $p_0 \geq p_1$ *and* $\max\{p_m, p_1\} = p_m$.

Proof. Inserting the values for p_0 and p_1 gives $(1-2\varepsilon)(R-\alpha R) \geq 0$ which always holds because $\varepsilon \leq 1/2$. $p_m \geq p_1$ can be rewritten as $(1-\varepsilon)(1-\alpha)+\alpha\varepsilon \geq 0$ which always holds. □

Firms with a sufficiently high success probability choose the risky project and to pay the bribe, where the cutoff levels for these two choices differ.

Proposition 1. *In equilibrium, if the firm chooses the safe project, the bureaucrat chooses the small bribe. If the firm chooses the risky project, the bureaucrat chooses*

[2] Firms with $p \geq p_0$ only make positive profits if $\mathbb{E}(risky) = (p-\alpha)R \geq 0$ or $p \geq \alpha$. Because it can be shown that $p_0 \geq \alpha$, as this can be simplified to $1/2 \geq \varepsilon$, these firms indeed make positive profits.

the high bribe. A firm chooses the risky project if either $p \geq \max\{p_0, p_h\}$ *or* $p_0 > p \geq p_m$*. If*

$$\frac{\alpha(R-r)}{2\alpha(R-r)+r} \equiv \bar{\varepsilon} \geq \varepsilon$$

all firms with $p \geq p_m$ *choose the risky project. If* $\bar{\varepsilon} \leq \varepsilon$*, all firms with* $p \geq p_h$ *choose the risky project.*

Proof. (i) $p \geq \max\{p_0, p_h\}$: $p_0 \geq p_h$ if $\bar{\varepsilon} \geq \varepsilon$. If $\bar{\varepsilon} \geq \frac{1}{2}$, this holds $\forall\varepsilon$. Because this inequality can be rewritten as $0 > r$, both $p_0 > p_h$ and $p_h > p_0$ are possible. $1 \geq \max\{p_0, p_h\}$ because $1 > p_h$ and $1 \geq p_0$. (ii) $p_0 > p \geq p_m$: $p_0 \geq p_m$ if $\bar{\varepsilon} \geq \varepsilon$. $p_h \geq p_m$ if $\frac{\alpha(R-r)}{2\alpha(R-r)+r} \geq \varepsilon$.

□

If the bureaucrat can condition his bribe demand on the firm's choice, the higher return of the risky project leads to a higher bribe demand. The risky project is chosen by firms with a sufficiently high success probability. The exact cutoff level depends on ε, the amount of information the firm has when paying the bribe.

We continue by investigating under which conditions firms make the efficient project choice. If the firm's project choice is observed by the bureaucrat, the risky project leads to higher bribe payments making it less attractive. The only exception is the case of perfect information, $\varepsilon = 0$. In this case the firm only pays the bribe if the project is successful reducing the firm's choice to comparing $r(1-\alpha)$ and $pR(1-\alpha)$ and resulting in the efficient project choice.

Proposition 2. *Firms make inefficiently safe project choices unless* $\varepsilon = 0$.

Proof. See Appendix. □

A perfectly precise signal, $\varepsilon = 0$, corresponds to income taxation where the bribe is paid after the project's risk realized. Projects with a high level of ε can be interpreted as innovative projects such that well-established firms should innovate less the more corrupt the country. The following example illustrates this result for the two extreme values of ε.

Example 2. *For* $\varepsilon = 0$, $p_0 = 1$ *and* $p_m = \frac{r}{R}$*. Hence, we observe the efficient project choice if* $\varepsilon = 0$*. All firms with* $p \geq p_m = \frac{r}{R}$ *choose the risky project (and no firm pays the bribe if* $s = 0$ *is observed). For* $\varepsilon = \frac{1}{2}$, $p_0 = \alpha$, $p_1 = \alpha$, $p_m = \frac{2(1-\alpha)r+\alpha R}{R}$ *and* $p_h = \alpha + (1-\alpha)\frac{r}{R}$*. Because* $p_0 = p_1$, p_m *is irrelevant and choices are inefficiently safe because only firm with* $p_h > \frac{r}{R}$ *choose the risky project.*

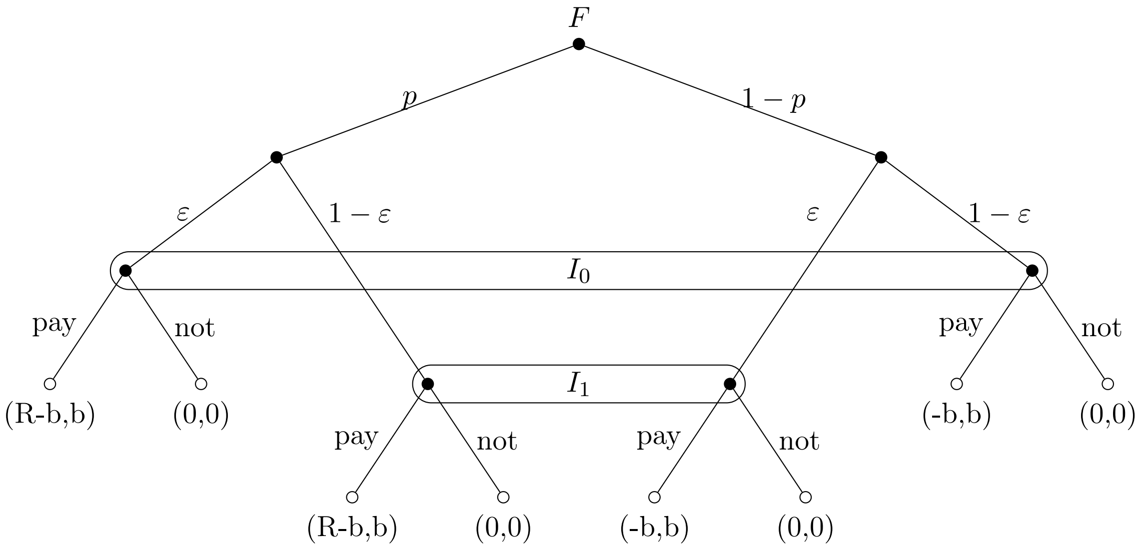


Figure 3: Firm's choice given bribe demand b

4.2 Bribe chosen before technology

This section shows the influence of corruption on the firm's technology choice if the bribe demand is chosen before the technology. In this case, corruption induces some firms to make inefficiently risky technology choices if the signal is sufficiently precise.

The bureaucrat knows the distribution $f(p)$ and the precision of the signal, ε. He knows that his bribe demand b influences the firm's project choice. Given b, the firm's behavior is depicted in figure 3. At I_0, firms with $p \geq p_0$, and, at I_1, firms with $p \geq p_1$, pay the bribe. Now the bureaucrat's bribe demand also influences which firms pay the bribe. At information set I_0, firms pay the bribe if

$$p \geq \frac{(1-\varepsilon)b}{\varepsilon(R-b)+(1-\varepsilon)b} \equiv p_0(b)$$

and at information set I_1, firms pay the bribe if

$$p \geq \frac{\varepsilon b}{(1-\varepsilon)(R-b)+\varepsilon b} \equiv p_1(b).$$

Lemma 4. *A higher bribe b makes payment by the firms less likely.*

Proof. A higher bribe increases the cutoff levels because $\frac{\partial p_0}{\partial b} = \frac{(1-\varepsilon)\varepsilon R}{(\varepsilon(R-b)+(1-\varepsilon)b)^2} \geq 0$, $\frac{\partial p_1}{\partial b} = \frac{(1-\varepsilon)\varepsilon R}{((1-\varepsilon)(R-b)+\varepsilon b)^2} \geq 0$ and $\frac{\partial p_m}{\partial b} = \frac{\varepsilon(1-\varepsilon)R+(1-\alpha)r(1-2\varepsilon)}{((1-\varepsilon)(R-b)+b\varepsilon)^2} > 0$ because $\varepsilon \leq 1/2$. □

The bribe demand can be either higher or lower than the safe return. Therefore, any bribe $b \leq r$ can be expressed as βr with $\beta \leq 1$ and any bribe $b \leq R$ can be

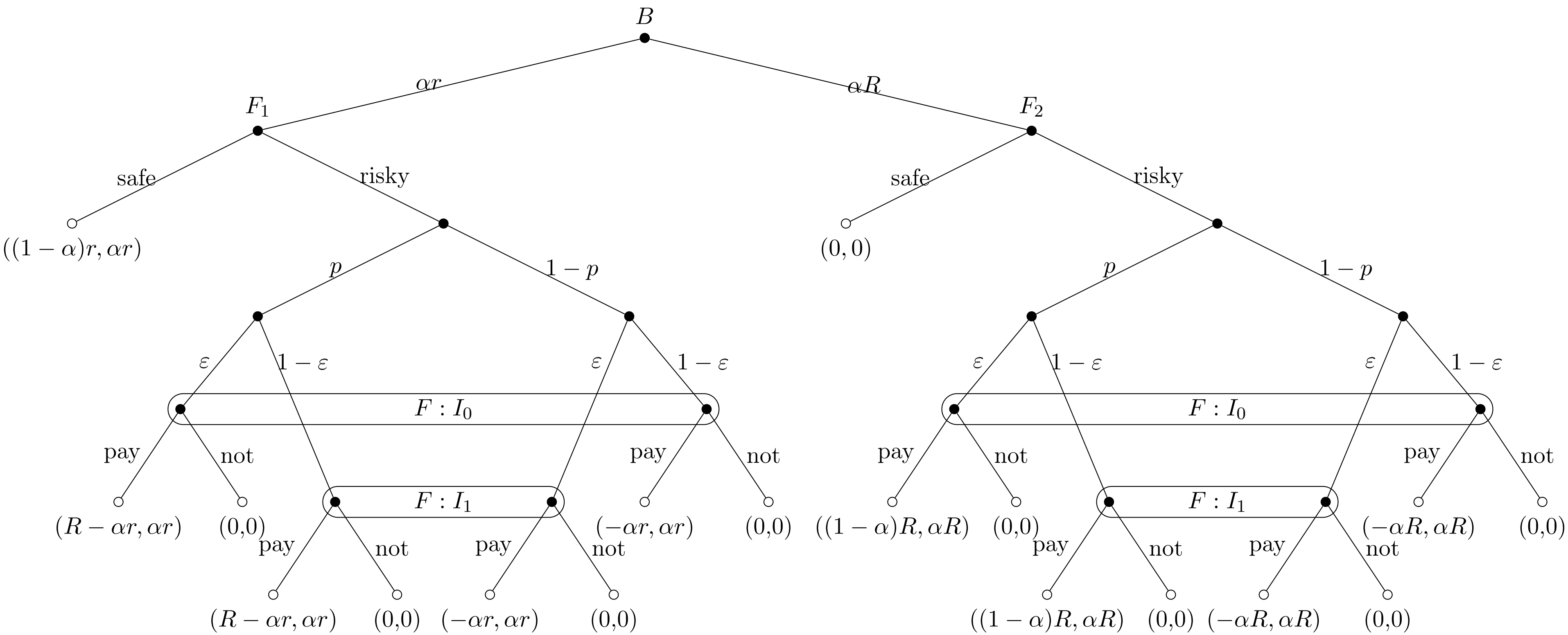


Figure 4: Bribe chosen before technology

expressed as γR with $\gamma \leq 1$. In the following two sections, let αr denote the bribe that maximizes the bureaucrat's income if firms choose both the safe and the risky project and let $\alpha R > r$ denote the bribe that maximizes the bureaucrat's income if all firms choose the risky project.[3] Because firms can pay αr with both the safe and risky project returns, we will refer to αr as the low bribe demand. A high bribe demand αR can only be paid with the risky project return. Given these two bribe demands, we first look at firm choices and then at the bureaucrat's bribe demand.

4.2.1 Low bribe demand

If the bureaucrat chooses the low bribe, firms choose inefficiently risky projects if they have a lot of information about the risky project when paying the bribe. If, on the other hand, the firms have only limited information about the risky project's success when paying the bribe, the bribe payment becomes a fixed cost such that project choices are efficient. We first determine the relevant cutoff levels by inserting $b = \alpha r$ in $p_0(b)$ and $p_1(b)$.

$$p_0(\alpha r) = \frac{(1-\varepsilon)\alpha r}{\varepsilon(R-\alpha r) + (1-\varepsilon)\alpha r} \text{ and } p_1(\alpha r) = \frac{\varepsilon \alpha r}{(1-\varepsilon)(R-\alpha r) + \varepsilon \alpha r} \tag{11}$$

There are three groups of firms: firms that always pay the bribe ($p \geq p_0(\alpha r)$), firms that never pay ($p_1(\alpha r) > p$) and firms that pay if they observe a signal of success ($p_0(\alpha r) > p \geq p_1(\alpha r)$). The expected payoff of choosing the safe project is given by $\mathbb{E}(safe) = (1-\alpha)r$. The expected payoff of choosing the risky project depends on the firm's success probability.

$$\mathbb{E}(risky) = \begin{cases} 0 \text{ if } p < p_1(\alpha r) \\ p(1-\varepsilon)R - \alpha r(p(1-\varepsilon) + (1-p)\varepsilon) \text{ if } p_1(\alpha r) \leq p < p_0(\alpha r) \\ pR - \alpha r \text{ if } p_0(\alpha r) \leq p \end{cases} \tag{12}$$

Firms with $p_1(\alpha r) \leq p < p_0(\alpha r)$ choose the safe project if $\mathbb{E}(safe) > \mathbb{E}(risky)$ which holds if their success probability is sufficiently small.

$$\frac{r(1-\alpha+\alpha\varepsilon)}{(1-\varepsilon)(R-\alpha r) + \alpha\varepsilon r} \equiv p_m > p \tag{13}$$

Some relations between the cutoff values always hold and are summarized next.

[3] If $\alpha R < r$, some firms choose the safe project and therefore the assumption of the maximization problem that all firms choose the risky project is not satisfied. Whenever $\alpha R < r$, αr is the only income-maximizing bribe.

Lemma 5. *Both* $p_m(\alpha r) \geq p_1(\alpha r)$ *and* $p_0(\alpha r) \geq p_1(\alpha r)$ *always hold.*

Proof. $p_1(\alpha r) \geq p_m(\alpha r)$ if $0 \geq (R-\alpha r)(1-\varepsilon)(1-\alpha)+\varepsilon\alpha r(1-\alpha)$. Because this never holds, $p_m(\alpha r) \geq p_1(\alpha r)$ always. $p_0(\alpha r) \geq p_1(\alpha r)$ because this can be simplified to $1 \geq 2\varepsilon$. □

The next proposition summarizes the firm's project choices if the bureaucrat asks for the small bribe.

Proposition 3. *Given that the bureaucrat chooses the small bribe* αr*, a firm chooses the risky project if* $p_0(\alpha r) > p \geq p_m$ *or* $p \geq max\left\{p_0(\alpha r), \frac{r}{R}\right\}$.

$$\tilde{\varepsilon} = \frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r}$$

If $\varepsilon < \tilde{\varepsilon}$*, firms choose the safe project if* $p < p_m(\alpha r)$ *and the project choice is inefficiently risky. If* $\tilde{\varepsilon} \leq \varepsilon$*, firms choose the safe project if* $p < \frac{r}{R}$ *and the project choice is optimal.*

Proof. See Appendix. □

If the bribe demand is independent of the firm's technology choice, the firm does not profit from a smaller bribe if it chooses the safe technology. Therefore, firms with an intermediate or high level of the success probability switch to the risky technology if the signal is sufficiently precise. In the case of everyday production, with ε small, corruption leads to inefficiently high levels of risk implying that many more projects fail in expectation. In the case of innovation, with ε high, corruption leads to the optimal level of risk because the bribe is similar to a fixed cost that always has to be paid. Note the underlying assumption of the firm comparing innovating to doing something else. The following example demonstrates the intuition for the two most extreme values of ε, $\varepsilon = 0$ and $\varepsilon = 1/2$.

Example 3. *For* $\varepsilon = 0$, $p_0(\alpha r) = 1$, $p_1(\alpha r) = 0$, $p_m(\alpha r) = \frac{r(1-\alpha)}{R-\alpha r}$ *and* $\mu = p$. *Because* $\frac{r}{R} > p_m(\alpha r)$*, there are too many risky choices. For* $\varepsilon = 1/2$, $p_0(\alpha r) = \frac{\alpha r}{R}$, $p_1(\alpha r) = \frac{\alpha r}{R}$, $p_m(\alpha r) = \frac{r(2-\alpha)}{R}$ *and* $\mu = 1/2$*. In this case,* p_m *and* μ *are irrelevant because* $p_0(\alpha r) = p_1(\alpha r)$*. Firms choose the risky project if* $p \geq \max\{\frac{r}{R}, \frac{\alpha r}{R}\} = \frac{r}{R}$*. Therefore, the project choice is efficient.*

For the bureaucrat choosing the low bribe αr, we summarize the firm's project choice based on the signal's precision and the firm's success probability in table 1.

Table 1: Firm's project choice if the bureaucrat demands the low bribe αr

firm	signal precision $\varepsilon < \tilde{\varepsilon}$	$\tilde{\varepsilon} \leq \varepsilon$
$p < p_1(\alpha r)$	safe	safe
$p_1(\alpha r) \leq p < p_0(\alpha r)$	risky if $p \geq p_m$	safe
$p_0(\alpha r) \leq p$	risky	risky if $p \geq r/R$

4.2.2 High bribe demand

If the bureaucrat chooses the high bribe, all firms choose the risky project. The more imprecise the signal, the fewer firms pay the bribe. Because the high bribe is higher than the safe return, $r < \alpha R$, no firm chooses the safe project, but firms might decide not to pay the bribe. Depending on the signal, a firm pays the bribe if

$$p \geq p_0(\alpha R) = \frac{(1-\varepsilon)\alpha}{\varepsilon(1-\alpha)+(1-\varepsilon)\alpha} \text{ if } s = 0$$
$$p \geq p_1(\alpha R) = \frac{\varepsilon\alpha}{(1-\varepsilon)(1-\alpha)+\alpha\varepsilon} \text{ if } s = 1$$

It can be shown that it always holds that $p_0(\alpha R) \geq p_1(\alpha R)$ because this can be simplified to $1 - 2\varepsilon \geq 0$. A firm with a small success probability, $p < p_1(\alpha R)$, never pays the bribe and therefore never produces. A firm with a high success probability, $p \geq p_0(\alpha R)$, always pays the bribe. A firm with an intermediate success probability, $p_1(\alpha R) \leq p < p_0(\alpha R)$, pays the bribe if it observes a signal indicating a successful project, i.e. if it reaches I_1. Because either some firms decide not to produce or too many firms choose the risky project, distortions always result.

Proposition 4. *All firms choose the risky project if the bureaucrat chooses the high bribe αR. The efficient level of risk is reached if $\varepsilon = \frac{r(1-\alpha)}{\alpha(R-r)+r(1-\alpha)}$.*

$$\hat{\varepsilon} \equiv \frac{r(1-\alpha)}{\alpha R + r(1-2\alpha)} \tag{14}$$

If $\varepsilon < \hat{\varepsilon}$, risk is inefficiently high. If $\varepsilon > \hat{\varepsilon}$, risk is inefficiently low.

Proof. See Appendix. □

If firms have to pay a bribe that is higher than the safe return, firms with a low success probability of the risky project never pay the bribe and therefore drop out of production. A more precise signal both increases risk and decreases the number of firms never producing.

4.2.3 Bureaucrat's decision

In this section, the bureaucrat's choice is derived. The bureaucrat chooses the bribe that maximizes his ex post income.[4] The bribe can be either payable with both the safe and risky return, $b \leq r$, or with the risky return only, $b > r$. In the first case, some firms choose the safe and some the risky project. In the second case, all firms choose the risky project. No matter the actual bribe demand, bribes can therefore be written as fractions of the safe and risky return, respectively. Hence, the low bribe can be written as βr and the high bribe as $\gamma R > r$. The high bribe γR maximizes $B(\gamma R) = \gamma R - \pi \frac{\gamma R}{R} \gamma R$. The low bribe βr maximizes $B(\beta r) = \beta r - \pi \frac{\beta r}{\rho} \beta r$ with $\rho = r, R$. The following assumption guarantees that these two bribe levels are indeed different.

Assumption 1.

$$\pi < \frac{R}{2r}$$

Lemma 6. *Under assumption 1, if the bribe demand is observable, all firms choose the risky project if the bureaucrat asks for the high bribe αR while some firms choose the safe project if the small bribe αr is demanded.*

$$\alpha = \frac{1}{2\pi}$$

Proof. High bribe: Maximizing $B(\gamma R)$ gives $\gamma = \frac{1}{2\pi}$. Low bribe: Maximizing $B(\beta r)$ gives $\beta = \frac{R}{2\pi r}$ for $\rho = r$ and $\beta = \frac{1}{2\pi}$ for $\rho = R$. The first solution results in a bribe $\beta r = \frac{R}{2\pi}$ which equals the high bribe. But since $\frac{R}{2\pi} > r$ by assumption, not all firms can pay the bribe. Only $\beta = \frac{1}{2\pi}$ can be paid by all firms. Hence, $\beta = \gamma = \alpha$. □

If assumption 1 is violated, there is no optimal high bribe demand and the bureaucrat always chooses the low bribe demand. This occurs if the government's detection efforts are high such that two optimal bribe demands only exist in very corrupt countries. Assumption 1 is equivalent to $\alpha R > r$ and will be assumed to hold in the following.

We first calculate the bureaucrat's payoff of choosing the low bribe. Firms with a low success probability, $p < p_1(\alpha r)$, choose the safe project and always pay the low bribe αr; their expected bribe payment is given by $\alpha r \int_0^{p_1(\alpha r)} f(p)dp$. Firms with a high success probability, $p \geq p_0(\alpha r)$, choose the risky project and always pay the low bribe αr leading to an expected payment of $\alpha r \int_{p_0(\alpha r)}^1 f(p)dp$. Firms with an

[4] In the appendix, the specification for maximizing ex ante income is shown.

intermediate success probability, $p_1(\alpha r) \leq p < p_0(\alpha r)$, choose the risky project if $p_m(\alpha r) \leq p$ and only pay the bribe if they reach I_1. The probability that they pay is given by $\mu = p(1-\varepsilon) + (1-p)\varepsilon$. The expected payoff of choosing the low bribe is then given by

$$\mathbb{E}(\alpha r) = \alpha r \left(\int_0^a f(p)dp + \int_a^b (p(1-\varepsilon) + (1-p)\varepsilon) f(p)dp + \int_b^1 f(p)dp \right) \quad (15)$$

with $a = min\{p_0(\alpha r), p_m(\alpha r)\}$ and $b = p_0(\alpha r)$. In order to evaluate this expression, we need to know which value of a applies. It has already been shown that $p_m(\alpha r) \geq p_0(\alpha r)$ if $\varepsilon \geq \tilde{\varepsilon}$. Hence, if $\varepsilon \geq \tilde{\varepsilon}$, then $p_m(\alpha r) \geq p_0(\alpha r)$ such that $a = p_0(\alpha r)$ and firms with $p \in [p_1(\alpha r), p_0(\alpha r)]$ do not choose the risky project but only the safe project. If, on the other hand, $\tilde{\varepsilon} > \varepsilon$, then $p_0(\alpha r) > p_m(\alpha r)$ such that $a = p_m(\alpha r)$ and firms with an intermediate success probability p choose the risky project. The bureaucrat's expected payoff of choosing the low bribe depends on the signal's precision.

$$\mathbb{E}(\alpha r) = \begin{cases} \alpha r \left(\int_0^{p_0(\alpha r)} f(p)dp + \int_{p_0(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^1 f(p)dp \right) = \alpha r \text{ if } \varepsilon \geq \tilde{\varepsilon} \\ \alpha r \left(\int_0^{p_m(\alpha r)} f(p)dp + \int_{p_m(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^1 f(p)dp \right) \text{ if } \tilde{\varepsilon} > \varepsilon \end{cases}$$

We continue by deriving the expected payoff if the high bribe is chosen. Firms with $p \in [p_1(\alpha R), p_0(\alpha R))$ only pay the bribe if they receive a signal of success. Firms with $p \geq p_0$ always pay the bribe. Remembering that $\mu = p(1-\varepsilon) + (1-p)\varepsilon$, the expected payoff is given by

$$\mathbb{E}(\alpha R) = \alpha R \left(\int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp + \int_{p_0(\alpha R)}^1 f(p)dp \right) \quad (16)$$

Proposition 5. *The bureaucrat chooses the small bribe if*

$$\frac{r}{R} \geq \int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp + \int_{p_0(\alpha R)}^1 f(p)dp \text{ if } \varepsilon \geq \tilde{\varepsilon}$$

$$\frac{r}{R} \left(F(p_m(\alpha r)) + \int_{p_m(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^1 f(p)dp \right)$$
$$\geq \int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp + \int_{p_0(\alpha R)}^1 f(p)dp \text{ if } \tilde{\varepsilon} > \varepsilon$$

If $\varepsilon \geq \tilde{\varepsilon}$, the bureaucrat is more likely to choose αr, the higher the safe return, the lower the risky return and the higher α. If $\varepsilon = 0$, the bureaucrat is more likely to choose αr, the higher the safe return, the lower the risky return and the lower α.

Proof. See Appendix. □

Generally speaking, lower distortions result if the bureaucrat chooses the low bribe. Choosing the low bribe becomes more beneficial for the bureaucrat the higher the safe return and the lower the risky return. Interestingly, the effect of governmental detection efforts on the bureaucrat's choice depends on the signal's precision. If the signal is very imprecise, a decrease in detection efforts raises both optimal bribe levels. An increase in the high bribe, however, implies that more firms choose not to produce and not to pay the bribe. This, in turn, lowers the income resulting from the high bribe such that the bureaucrat decides to ask for the low bribe instead. If, on the other hand, the signal is perfectly precise, while a decrease in detection efforts still raises both bribe levels, no firm drops out of production altogether if the bureaucrat chooses the high bribe. An increase in the optimal bribe level, however, increases distortions both for the low and the high bribe. Because distortions increase relatively more for the low bribe than for the high bribe, the bureaucrat chooses the high bribe if detection efforts decrease. The optimal detection effort of the government therefore depends on the signal precision if a change in detection efforts can induce a change in the bureaucrat's bribe demand.

Lastly, observe that the bureaucrat has a higher expected bribe income from being the second mover if the signal is perfectly precise and observing the firm's choice is costless.

Remark 4. *The bureaucrat prefers to be the second mover if $\varepsilon = 0$.*

Proof. The bureaucrat's expected payoff is given by $\mathbb{E}(b_2) = \alpha r F\left(\frac{r}{R}\right) + \alpha R \int_{\frac{r}{R}}^{1} p f(p) dp$ if he moves second, while his expected payoff of being first-mover is given by $\mathbb{E}(\alpha r) = \alpha r \left(F(p^*) + \int_{p^*}^{1} p f(p) dp\right)$ and $\mathbb{E}(\alpha R) = \alpha R \int_0^1 p f(p) dp$. Both $\mathbb{E}(b_2) \geq \mathbb{E}(\alpha r)$ (since $r \int_{p^*}^{\frac{r}{R}} f(p)(1-p) dp + (R-r) \int_{\frac{r}{R}}^{1} p f(p) dp \geq 0$) and $\mathbb{E}(b_2) \geq \mathbb{E}(\alpha R)$ (since $\int_0^{\frac{r}{R}} F(p) dp \geq 0$). □

Whenever possible, the bureaucrat will therefore try not to set the bribe before the firm chooses its technology if $\varepsilon = 0$. We should therefore observe both the firm and the bureaucrat as first mover in this case.

5 The model: Repeated game

This section shows that the timing of the bribe demand matters for the distorting effect of corruption. It extends the intuition derived before to repeated games where either the firm or the bureaucrat is the long-run player. If the firm is long-lived, corruption does not result in distortions if the firm is sufficiently patient and the signal is perfectly precise, $\varepsilon = 0$. If the bureaucrat is long-lived, corruption does not distort firms' technology choices if the low bribe αr is chosen and the signal is sufficiently imprecise.

While assuming that either the firm or the bureaucrat moves before the other player was helpful to show the role of information, one could find it easier to believe that the dynamics of information evolve over time. Returning to the examples, companies like Odebrecht or Schenker have been in business very long such that the bureaucrat in charge can base his bribe demand on past records. The bureaucrat can not do this if a politician like Sohel Rana applies for a permit. In this case it is easier to believe that the Bangladeshi bureaucrat has acquired a reputation for asking for a certain bribe.

In equilibrium, the long-run player's strategy choice maximizes the discounted sum of payoffs, the short-run player's strategy maximizes the one period payoff, and both strategies are mutually best replies. Beliefs are formed using Bayes' rule wherever possible.

5.1 Firm as long-lived player

If the firm can commit to a technology, corruption leads to inefficiently safe project choices unless the signal is perfectly precise. Section 4.1 has shown the intuition for this result when the firm moves before the bureaucrat sets the bribe. This section extends this intuition to the firm being the long-lived player in a repeated game. We will see that we arrive at the standard result that a sufficiently patient long-lived player can choose his most preferred equilibrium (e.g. Fudenberg, Kreps and Maskin 1990; Fudenberg and Levine 1989, 1992).[5] A sufficiently patient long-lived firm acquires a reputation for choosing a certain technology.

Suppose that the firm is the long-run player in a repeated game played against an infinite sequence of bureaucrats, the short-run players. Assume further that firm and bureaucrat choose bribe and project simultaneously and that the firm can condition the bribe payment on the observed signal and bribe demand. The history of play is known to all short-run bureaucrats. The firm discounts at rate δ. Let $\pi(b, p)$ denote the firm's profit if the bureaucrat chooses bribe $b = \{\alpha r, \alpha R\}$ and the firm chooses project $p = \{s, r\}$. Let $\gamma_t(\alpha r)$ be the probability the bureaucrat chooses the low bribe in period t.

The bureaucrats maximize the per-period payoff. Their belief that the firm chooses the safe project in period t is given by β_t. Let $\mathbb{E}(b, p)$ denote the bureaucrat's

[5] In this case, this equilibrium is also the most preferred by the short-run player.

expected payoff if the bureaucrat chooses bribe $b = \{\alpha r, \alpha R\}$ and the firm chooses project $p = \{s, r\}$. Expected payoffs of choosing the high and low bribe are given by

$$\begin{aligned}\mathbb{E}(\alpha r) =& \beta_t \mathbb{E}(\alpha r, s) + (1-\beta_t)\mathbb{E}(\alpha r, r)\\ \mathbb{E}(\alpha R) =& (1-\beta_t)\mathbb{E}(\alpha R, r)\end{aligned}$$

Obviously, the high bribe is the bureaucrats' best response to the firm choosing the risky project.[6] This gives rise to the following firm behavior.

Lemma 7. *If the firm chooses the risky project once, it chooses the risky project forever. This does not have to be true for the safe project.*

Proof. (i) Suppose the firm played risky once. If the bureaucrats expect the risky project, they choose αR in the next period. But with the safe return, the firm cannot pay the bribe. (ii) Suppose the firm played safe once. If the bureaucrats expect the safe project, they choose αr in the next period. The firm can choose risky to reap a higher return. □

We can distinguish between three different types of firms: firms that choose the safe project in every period, firms that choose the risky project in every period and firms that want to trick the bureaucrats into believing that they will choose the safe project but choose the risky project instead. If the last group is small, the bureaucrats' play is entirely determined by the firm's choice in the first period. If the last group is large, however, bureaucrats choose the high bribe if the firm plays risky once and randomize between the small and high bribe for n periods such that the tricking firm type is indifferent between the safe and risky project. After n periods, the bureaucrats are sufficiently sure that the firm will indeed continue choosing the safe project.[7] The tricking firm type randomizes between the safe and risky project such that the bureaucrat is indifferent between choosing both bribe demands.[8] The

[6] It is possible that the firm has a small success probability ($p < p_m$ or $p < p_h$ depending on ε) such that the firm choosing the safe project and the bureaucrats the low bribe would be an equilibrium, but myopic bureaucrats do not experiment. Ely and Välimäki (2003) obtain a similar result where additional information is not revealed because it does not benefit the short-run players.

[7] It is possible that the firm still belongs to this group and the bureaucrat's behavior is not optimal ex post, but the bureaucrats are sufficiently sure that this will not be the case when stopping to randomize.

[8] Value functions for the tricking firm type can be found in the appendix.

expected profits of choosing the safe and risky project are given by

$$\mathbb{E}(safe) = \pi(\alpha r, s)\sum_{t=0}^{n-1}\gamma_t(\alpha r)\delta^t + \pi(\alpha R, s)\sum_{t=0}^{n-1}(1-\gamma_t(\alpha r))\delta^t + \pi(\alpha r, s)\sum_{t=n}^{\infty}\delta^t \quad (17)$$

$$\mathbb{E}(risky) = \gamma_0(\alpha r)\pi(\alpha r, r) + (1-\gamma_0(\alpha r))\pi(\alpha R, r) + \sum_{t=1}^{\infty}\pi(\alpha R, r)\delta^t \quad (18)$$

with $\pi(\alpha R, s) = 0$.

We look at the equilibrium where both the bureaucrats and the tricking type of firms randomize to keep the other player indifferent. Whenever the firm chooses the risky technology, all bureaucrats choose the high bribe forever.

Proposition 6. *If the firm is the sufficiently patient long-lived player, distortions always result unless $\varepsilon = 0$. The firm chooses the safe project if*

$$\pi(\alpha r, s)\sum_{t=0}^{n-1}\gamma_t(\alpha r)\delta^t + \pi(\alpha r, s)\sum_{t=n}^{\infty}\delta^t \geq \gamma_0(\alpha r)\pi(\alpha r, r) + (1-\gamma_0(\alpha r))\pi(\alpha R, r) + \sum_{t=1}^{\infty}\pi(\alpha R, r)\delta^t$$

Proof. See Appendix. □

Corruption does not affect the firm's technology decision if the signal is perfectly precise, $\varepsilon = 0$, and the firm is sufficiently patient to build a reputation for choosing the efficient technology. Hence, firms that enter a corrupt country after acquiring a reputation for a certain technology and older firms in general should be less affected when there is corruption.

5.2 Bureaucrat as long-lived player

Corruption leads to distortions if the signal is precise and the bureaucrat sets the bribe before the firm decides which technology to use. Section 4.2 has provided the intuition for this result in the context of a game with the bureaucrat moving first. It is shown in the following that this intuition continues to hold if the bureaucrat can acquire a reputation for choosing a certain bribe. Suppose that the bureaucrat is the long-run player in a repeated game played against an infinite sequence of firms, the short-run players. Bureaucrat and firm move simultaneously and the history of play is known to the firms. The bureaucrat's discount rate is denoted by δ. Firms pay the bribe after observing both the signal and the bribe demand.

The bureaucrat's belief that the firm reacts to αr in period t is γ_t. Beliefs are updated every period. Firms react to the bribe demand, they expect. The bureaucrat's payoff of playing b is given by $\mathbb{E}(b, b')$ where b' is the bribe demand, the firms react to. The firm's project choice is entirely determined by the bureaucrat's bribe demand in the first period.

Lemma 8. *If the bureaucrat chooses the low bribe once, he chooses the low bribe forever. If the bureaucrat chooses the high bribe once, he chooses the high bribe forever.*

Proof. See Appendix. □

The expected payoffs of choosing the high and low bribe for the bureaucrat are given by

$$\mathbb{E}(\alpha r) = \gamma_0 \mathbb{E}(\alpha r, \alpha r) + (1-\gamma_0)\mathbb{E}(\alpha r, \alpha R) + \mathbb{E}(\alpha r, \alpha r)\sum_{t=1}^{\infty}\delta^t$$
$$\mathbb{E}(\alpha R) = \gamma_0 \mathbb{E}(\alpha R, \alpha r) + (1-\gamma_0)\mathbb{E}(\alpha R, \alpha R) + \mathbb{E}(\alpha R, \alpha R)\sum_{t=1}^{\infty}\delta^t$$

The bureaucrat chooses the low bribe if the expected payoff is higher. Intuitively, by influencing the firms' decisions, the bureaucrat can choose his preferred equilibrium. For example, by choosing the high bribe, he can induce all firms with a small success probability to choose the risky project, which leads to a higher expected payoff of the high bribe than without this possibility. The short-lived firms play a best response to the bureaucrat's action.

Proposition 7. *If the bureaucrat is the long-lived player, distortions result unless the small bribe is chosen and $\varepsilon \geq \tilde{\varepsilon}$ and the bureaucrat is sufficiently patient. The bureaucrat chooses the small bribe if*

$$\begin{aligned}&\gamma_0 \mathbb{E}(\alpha r, \alpha r) + (1-\gamma_0)\mathbb{E}(\alpha r, \alpha R) + \mathbb{E}(\alpha r, \alpha r)\sum_{t=1}^{\infty}\delta^t \\ &\quad \geq \gamma_0 \mathbb{E}(\alpha R, \alpha r) + (1-\gamma_0)\mathbb{E}(\alpha R, \alpha R) + \mathbb{E}(\alpha R, \alpha R)\sum_{t=1}^{\infty}\delta^t\end{aligned} \tag{19}$$

Proof. $\mathbb{E}(\alpha r) \geq \mathbb{E}(\alpha R)$ if

$$\begin{aligned}&\gamma_0(\mathbb{E}(\alpha r, \alpha r) - \mathbb{E}(\alpha R, \alpha r)) + (1-\gamma_0)(\mathbb{E}(\alpha r, \alpha R) - \mathbb{E}(\alpha R, \alpha R)) \\ &\quad + \frac{\delta}{1-\delta}(\mathbb{E}(\alpha r, \alpha r) - \mathbb{E}(\alpha R, \alpha R)) \geq 0\end{aligned}$$

For $\delta \to 1$, the bureaucrat compares $\mathbb{E}(\alpha r, \alpha r)$ and $\mathbb{E}(\alpha R, \alpha R)$. □

If the bureaucrat is the long-run player while the firms are short-run players, corruption results in distortions if ε is small. A possible interpretation is that one bureaucrat works in the same area for a long time period thereby establishing a certain reputation while firms tend to be relatively young, e.g. start-up firms.

6 Distortions resulting from corruption influencing the firms' decisions

In this section we investigate the effect of corruption on two different measures of distortion if the bribe is independent of the project choice. These measures are the resulting volatility and the effect on expected output. We also compare the resulting expected output for different distributions of the success probability characterized by first-order stochastic dominance.

The influence of the precision of the signal and the identity of the long-lived player on firm's project choices are summarized in the following table. The intermediate case does not have to exist as it is possible that $\hat{\varepsilon} < \tilde{\varepsilon}$ where $\hat{\varepsilon} = \frac{r(1-\alpha)}{\alpha R + r(1-2\alpha)}$ and $\tilde{\varepsilon} = \frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r}$.

Table 2: Project choices

			information	
committing		very high ($\varepsilon = 0$)	intermediate ($\tilde{\varepsilon} < \varepsilon < \hat{\varepsilon}$)	low ($\hat{\varepsilon} < \varepsilon$)
bureaucrat	α low	too risky	too risky	too safe
	α high	more too risky	efficient	efficient
firm	any α	efficient	too safe	more too safe

The firm's project choice is efficient if it is observed by the bureaucrat and the signal is perfectly precise. This case is similar to taxation, which, as has already been argued by Shleifer and Vishny (1993), results in less distortion than corruption. Interestingly, firms also choose the efficient project if α, the share the bureaucrat takes, is high and the signal is very imprecise. Interpreting α as a measure of how widespread corruption is in a country, and a very imprecise signal as related to an innovative project, corruption should not distort innovative choices. Innovative activity is reduced, however, if the firm's project choice is observed or if the bribe is independent of the project and corruption is not too severe corresponding to α being low. Corruption increases the extent of risky project choices if the bureaucrat does not condition his bribe on the firm's project and the signal is very precise, corresponding to a daily business project.

Next, we look at the influence of the distribution of success probabilities on the bureaucrat's bribe demand. If $\varepsilon = 0$ or $\varepsilon > \tilde{\varepsilon}$, the bureaucrat never chooses the low bribe for the first-order stochastically dominant distribution $f(p)$ and the high bribe for the dominated distribution $g(p)$. A distribution $f(p)$ first-order stochastically dominates (FOSD) another distribution $g(p)$ if $F(p) \leq G(p) \; \forall \; p$.[9] The following observation results from first-order stochastic dominance.

Proposition 8. *If the bureaucrat chooses different bribes for $f(p)$ and $g(p)$ with $F(p) \leq G(p) \forall \; p$, the bureaucrat chooses αR for $f(p)$ and αr for $g(p)$ if $\varepsilon \geq \tilde{\varepsilon}$ or $\varepsilon = 0$.*

Proof. See Appendix. □

First-order stochastic dominance makes the high bribe more likely.

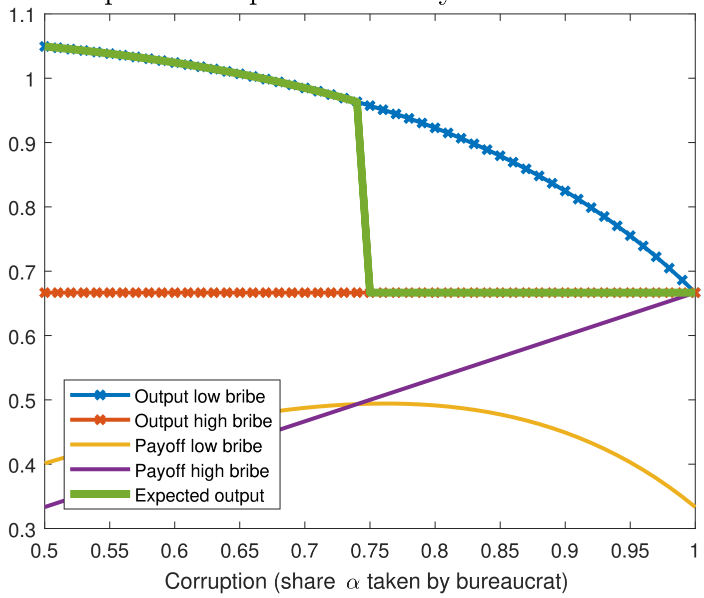
6.1 Effect on expected output

This section investigates the effect of corruption on expected output, y. Whenever corruption distorts the firms' behavior, expected output is reduced. When comparing expected output for two distributions, where distribution $f(p)$ first-order stochastically dominates distribution $g(p)$, expected output is higher for distribution $f(p)$ in the absence of corruption. If the bureaucrat chooses the high bribe for both distributions, this order is preserved and expected output is again higher for the first-order stochastically dominant distribution $f(p)$. If the low bribe αr is chosen for both $f(p)$ and $g(p)$, expected output can be higher or lower for $f(p)$ than for $g(p)$, depending on the parameter values. If the high bribe is chosen for $f(p)$ and the low bribe is chosen for $g(p)$, expected output is higher for $f(p)$ if the risky return is sufficiently large.

Figure 5 shows how the bribery choice of the bureaucrat translates into expected output. If the bureaucrat chooses the low bribe, an increase in the tolerance of corruption, i.e. an increase in α, first gradually increases and then decreases the payoff of the low bribe. If α increases, the low bribe increases leading to an increase in the bureaucrat's income. Simultaneously, however, increasing the low bribe raises the associated distortion leading to a decrease in the bureaucrat's income. Therefore, the bureaucrat's payoff of the low bribe first increases and then decreases in corruption, measured as the share α taken from the firm. Because an increase in corruption, the share α, increases distortions, expected output associated with the low bribe decrease as the low bribe increases. If the bureaucrat chooses the high

[9] In the following, $f(p)$ always denotes the first-order stochastic dominant distribution and $g(p)$ the dominated one.

Figure 5: Expected output induced by the bureaucrat's choice



Notes: $g(p) = 2(1-p)$, $\varepsilon = 0$, $r = 1$, $R = 2$

bribe, all firms choose the risky project in the example (because $\varepsilon = 0$). An increase in corruption therefore only increases the bureaucrat's bribe but does not increase distortions. Because firms do not change their project choices if corruption increases, expected output is independent of the level of corruption. Thus, the bureaucrat first chooses the low bribe and switches to ask for the high bribe if the tolerance for corruption is sufficiently high. This results in a sudden drop of expected output to its lowest level.

Expected output changes according to the proportion of firms operating with the safe and risky technology respectively. If there is no corruption, firms with $p < \frac{r}{R}$ choose the safe technology such that expected output is given by

$$y_f(0) = \int_0^{\frac{r}{R}} rf(p)dp + \int_{\frac{r}{R}}^1 Rpf(p)dp. \tag{20}$$

If the bureaucrat chooses the high bribe, a firm only ever pays the bribe if $p \geq p_1(\alpha R)$. But firms with $p < p_0(\alpha R)$ only continue production if they receive a positive signal. Expected output is given by

$$y_f(\alpha R) = R\left(\int_{p_1(\alpha R)}^{p_0(\alpha R)} p(1-\varepsilon)f(p)dp + \int_{p_0(\alpha R)}^1 pf(p)dp\right). \tag{21}$$

Lastly, if the bureaucrat chooses the low bribe, expected output depends on ε. If $\tilde{\varepsilon} > \varepsilon$, firms with $p_m(\alpha r) \leq p < p_0(\alpha r)$ only choose to continue with production if they receive a signal of success. Hence, their expected output is $p(1-\varepsilon)R$. If $\varepsilon \geq \tilde{\varepsilon}$, firms that choose the risky project always pay the bribe. Expected output is given by

$$y_f(\alpha r) = \begin{cases} \int_0^{p_m} rf(p)dp + \int_{p_m}^{p_0(\alpha r)} p(1-\varepsilon)Rf(p)dp + \int_{p_0(\alpha r)}^{1} pRf(p)dp \text{ if } \tilde{\varepsilon} > \varepsilon \\ \int_0^{\frac{r}{R}} rf(p)dp + \int_{\frac{r}{R}}^{1} pRf(p)dp \text{ if } \varepsilon \geq \tilde{\varepsilon} \end{cases}$$

First, we investigate which level of ε maximizes expected output given a bureaucrat's bribe demand. If the government has some influence on the projects available to firms, the government could try to encourage the choice of projects that maximize expected output. A perfectly precise signal maximizes output if the bureaucrat chooses the high bribe because in this case all firms start production. If the bureaucrat chooses the low bribe, on the other hand, expected output can be independent of the exact level of ε provided that $\varepsilon \geq \tilde{\varepsilon}$.

Lemma 9. *For $b = \alpha R$, expected output is maximized for $\varepsilon = 0$. For $b = \alpha r$, if $\varepsilon \geq \tilde{\varepsilon}$, expected output does not depend on ε. If $\varepsilon < \tilde{\varepsilon}$, there exists an interior solution for ε that maximizes output.*

Proof. See Appendix. □

The remainder of this section compares the effect of corruption on expected output for two distributions, where one, $f(p)$, first-order stochastically dominates the other, $g(p)$. Expected output is higher for the first-order stochastically dominant distribution $f(p)$ if there is no corruption or the bureaucrat chooses the high bribe for both distributions. These results follow directly from the assumed stochastic dominance.

Proposition 9. *Given that $f(p)$ first-order stochastically dominates $g(p)$, expected output is higher for $f(p)$ than for $g(p)$ if*

- *there is no corruption*
- *the high bribe αR is chosen for both $f(p)$ and $g(p)$ for all ε*
- *the low bribe αr is chosen for both $f(p)$ and $g(p)$ if $\tilde{\varepsilon} \leq \varepsilon$*

Proof. See Appendix. □

6.1.1 Corruption can lead to a reversal of expected output

Even though corruption reduces expected output for both $f(p)$ and $g(p)$, there are many instances where expected output continues to be higher for $f(p)$. We will see next that expected output can be lower for $f(p)$ than for $g(p)$ if the low bribe is chosen for both distributions and the firm is perfectly informed about the project's outcome, $\varepsilon = 0$.

In general, the effect of the low bribe on the order of output is ambiguous because for the first-order stochastically dominant distribution $f(p)$ the firms choosing the risky project generate a higher expected output while those choosing the safe project generate a higher expected output for $g(p)$. Intuitively, expected output is higher for $g(p)$ than for $f(p)$ if the mass of firms inefficiently switching to the risky project is higher under $f(p)$ than under $g(p)$ because this implies that corruption results in a larger distortion under $f(p)$ than under $g(p)$. Similarly, the risky return R has to be sufficiently small to induce the bureaucrat to choose the small bribe. Lastly, the share α has to fall into an intermediate range. The intuition is that for high values of α the bureaucrat chooses the high bribe, while for low values of α the distortion under $f(p)$ is not large enough to compensate for first-order stochastic dominance. Therefore, expected output is higher for $g(p)$ than for $f(p)$ and the bureaucrat chooses the low bribe for both distributions under a number of additional conditions. As the following result is shown for $\varepsilon = 0$, define $p^* = p_m(\alpha r) = \frac{r(1-\alpha)}{R-\alpha r}$ for $\varepsilon = 0$.[10] Note that p^* depends on α, r and R. To highlight this and minimize notation, in the following we will write $p^*(\alpha)$. Let $\bar{\alpha}$ denote the share α for which the bureaucrat is indifferent between choosing the low and the high bribe for $f(p)$, i.e. $F(p^*(\alpha))r + r\int_{p^*(\alpha)}^1 pf(p)dp = R\int_0^1 pf(p)dp$. Let $\underline{\alpha}$ and $\bar{\alpha}'$ denote the share α such that expected output of $g(p)$ with a low bribe equals expected output of $f(p)$ with a low bribe, $\int_0^{p^*(\alpha)} rg(p)dp + \int_{p^*(\alpha)}^1 pRg(p)dp = \int_0^{p^*(\alpha)} rf(p)dp + \int_{p^*(\alpha)}^1 pRf(p)dp$.[11]

Proposition 10. *For $F(p) \leq G(p)\ \forall p$, the low bribe is chosen for both $f(p)$ and $g(p)$ and expected output is higher for $g(p)$ than for $f(p)$ if $\varepsilon = 0$, $\underline{\alpha} < \alpha < \min\{\bar{\alpha}, \bar{\alpha}'\}$, $\frac{r}{R}$ sufficiently large, and $f(p) \geq g(p)$ over a sufficiently large range for $p \in [p^*(\alpha), \frac{r}{R}]$.*

Proof. See Appendix. □

This implies that corruption can reverse the order of expected output if some conditions are met. An economy that starts with a better distribution of success

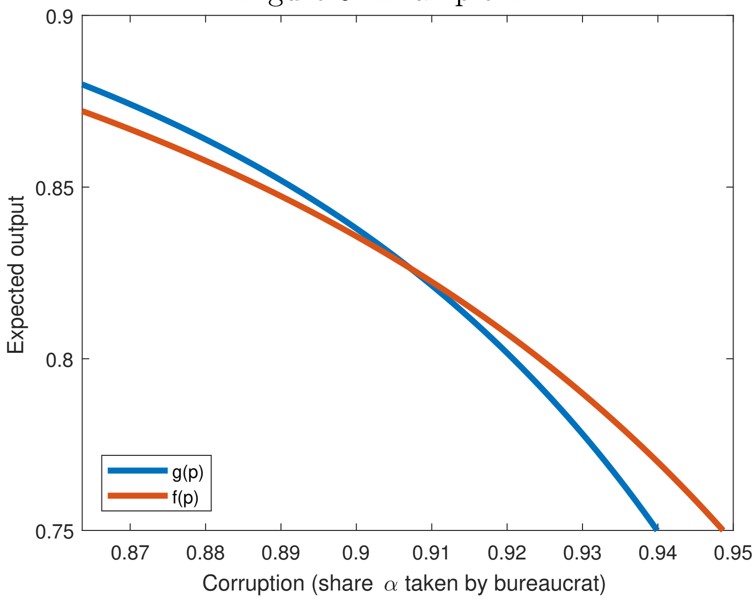
[10] An extension that the following statement holds for ε sufficiently close to 0 is shown in the appendix.

[11] The reason for two existing cutoff values of α is that $y_g(\alpha r) - y_f(\alpha r)$ first increases and then decreases in α.

probabilities, in terms of first-order stochastic dominance, can experience a worse outcome than the economy with the stochastically dominated distribution. The following example illustrates that output can be higher or lower for the first-order stochastically dominant distribution if the bureaucrat chooses the low bribe. As distributions, we use the triangular distribution with the mode at 0 for $g(p)$ and the uniform distribution for $f(p)$.

Example 1
Letting $\varepsilon = 0$, for $f(p) = 1$, we have $\mathbb{E}(\alpha r) = \alpha r(p^* + \frac{1}{2} - \frac{1}{2}(p^*)^2)$, $\mathbb{E}(\alpha R) = \frac{1}{2}\alpha R$ and $y_f(\alpha r) = rp^* + \frac{1}{2}R(1 - (p^*)^2)$. For $g(p)$ with $g(p) = 2(1-p)$, we have $\mathbb{E}(\alpha r) = 2\alpha r(p^* - p^{*2} + \frac{1}{6} + \frac{1}{3}p^{*3})$, $\mathbb{E}(\alpha R) = \frac{1}{3}\alpha R$ and $y_g(\alpha r) = 2rp^*(1 - \frac{1}{2}p^*) + R(\frac{1}{3} - p^{*2} + \frac{2}{3}p^{*3})$.

Figure 6: Example 1



Notes: $\varepsilon = 0$, $r = 0.95$, $R = 1.1$

Figure 6 depicts the evolution of expected output when α is sufficiently small to guarantee that the small bribe is chosen for both distributions. Expected output is higher for $g(p)$ than for $f(p)$ if α is not too large.

6.1.2 High bribe payments can be positively correlated with high output

If the bureaucrat chooses the high bribe for the first-order stochastically dominant distribution $f(p)$ and the low bribe for $g(p)$, expected output is higher for $f(p)$ if the risky return is sufficiently large. Intuitively, the choice of the high bribe imposes an additional distortion compared to the low bribe, which is mitigated for high values of the risky return.

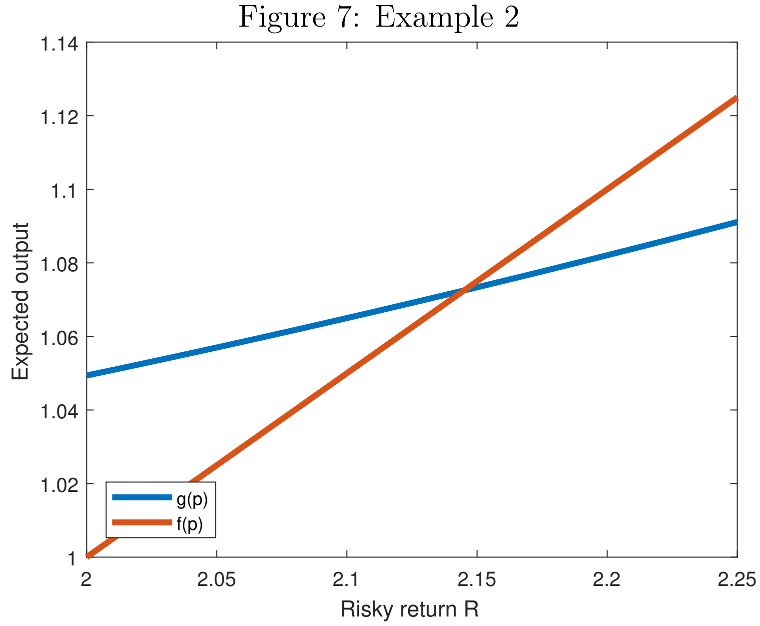
Proposition 11. *For $F(p) \leq G(p)\ \forall\ p$, if the bureaucrat chooses the high bribe for $f(p)$ and the low bribe for $g(p)$, there exists a R_{min} such that $\forall R > R_{min}$, expected output for $f(p)$ with a high bribe is higher than for $g(p)$ with a low bribe if $\varepsilon > \tilde{\varepsilon}$ or $\varepsilon = 0$.*

Proof. See Appendix. □

For sufficiently large values of the risky return, high bribes are positively correlated with expected output. This result does not imply that corruption improves economic outcomes because the level of both corruption and output results from the underlying distribution of success probabilities. It is therefore necessary to appropriately control for economic conditions when assessing the effect of corruption on output. The following example, using the two distributions from before, illustrates this observation:

Example 2
Letting $\varepsilon = 0$, for $g(p) = 2(1-p)$, we have $\mathbb{E}(\alpha r) = 2\alpha r(p^* - p^{*2} + \frac{1}{6} + \frac{1}{3}p^{*3})$, $\mathbb{E}(\alpha R) = \frac{1}{3}\alpha R$ and expected output $y_g(\alpha r) = rp^*(2 - p^*) + 2R(\frac{1}{6} - \frac{1}{2}p^{*2} + \frac{1}{3}p^{*3})$. For $f(p) = 1$, we have $\mathbb{E}(\alpha r) = \alpha r(p^* + \frac{1}{2} - \frac{1}{2}(p^*)^2)$, $\mathbb{E}(\alpha R) = \frac{1}{2}\alpha R$ and expected output $y_f(\alpha R) = \frac{1}{2}R$. Figure 7 depicts expected output for

Figure 7: Example 2



Notes: $\varepsilon = 0$, $r = 1$, $\alpha = 0.5$

different values of the risky return R where the bureaucrat chooses the high bribe for $f(p)$ and the low bribe for $g(p)$. As can be seen, once the risky return is higher than the threshold level R_{min}, expected output is higher for $f(p)$.

6.2 Effect on volatility

In this section, we investigate the effect of corruption on volatility. Generally speaking, corruption leads to an increase in volatility, compared to the situation without corruption, if the signal is very precise. If, on the other hand, the signal is very imprecise, corruption decreases volatility.[12]

In order to arrive at a measure of aggregate volatility, we aggregate the volatility resulting at the firm level. Firms that choose the safe investment do not contribute to volatility. Firms that choose the risky investment have an expected return of pR if they always pay the bribe. Each of these firms generates a volatility of $\sigma^2_{firm} = p(R-pR)^2 + (1-p)(0-pR)^2 = p(1-p)R^2$. Firms that choose the risky investment and only pay the bribe if they observe a signal of success have an expected payoff of $p(1-\varepsilon)R$. They generate a volatility of $\sigma^2_{firm} = p(1-\varepsilon)(1-p(1-\varepsilon))R^2$. Following this logic, aggregating volatility over all firms gives volatility without corruption as

$$\sigma^2(0) = R^2 \int_{\frac{r}{R}}^{1} p(1-p)f(p)dp.$$

If the bureaucrat chooses the high bribe αR, all firms choose the risky technology but only firms with $p \geq p_1(\alpha R)$ ever pay the bribe. Volatility is given by

$$\sigma^2(\alpha R) = \int_{p_1(\alpha R)}^{p_0(\alpha R)} p(1-\varepsilon)(1-p(1-\varepsilon))R^2 f(p)dp + \int_{p_0(\alpha R)}^{1} p(1-p)R^2 f(p)dp.$$

If the signal is very precise, more firms continue production with the risky project than if there was no corruption which increases volatility. If the signal is very imprecise, fewer firms continue production with the risky project resulting in a reduction in volatility.

Proposition 12. *If the bureaucrat chooses the high bribe αR, volatility increases if the signal is sufficiently precise and decreases if the signal is sufficiently imprecise.*

Proof. See Appendix. □

If the bureaucrat chooses the small bribe, volatility depends on the signal's precision. If the signal is very imprecise ($\varepsilon \geq \tilde{\varepsilon}$), the risky project is only chosen by firms with $r/R \leq p$. Hence, volatility is given by

$$\sigma^2(\alpha r) = R^2 \int_{\frac{r}{R}}^{1} p(1-p)f(p)dp.$$

[12]In the appendix, I compare the effect of corruption on additional volatility for two different distributions if $\varepsilon = 0$.

If the signal is very precise ($\tilde{\varepsilon} > \varepsilon$) and the bureaucrat chooses the small bribe, firms with $p \geq p_m(\alpha r)$ choose the risky project such that volatility is given by

$$\sigma^2(\alpha r) = R^2 \int_{p_m(\alpha r)}^{p_0(\alpha r)} p(1-\varepsilon)(1-p(1-\varepsilon))f(p)dp + R^2 \int_{p_0(\alpha r)}^{1} p(1-p)f(p)dp.$$

Proposition 13. *If the bureaucrat chooses the small bribe, volatility is at the efficient level if $\varepsilon > \tilde{\varepsilon}$. Volatility is inefficiently high if $\varepsilon = 0$.*

Proof. $\varepsilon > \tilde{\varepsilon}$: Excess volatility is given by $\sigma^2(\alpha r) - \sigma^2(0) = 0$. There is no excess volatility. $\tilde{\varepsilon} > \varepsilon$: Already shown that for $\tilde{\varepsilon} > \varepsilon$, $p_0 > r/R > p_m$. Excess volatility is given by $\sigma^2(\alpha r) - \sigma^2(0) = R^2(\int_{p_m}^{\frac{r}{R}} p(1-\varepsilon)(1-p(1-\varepsilon))f(p)dp + \int_{\frac{r}{R}}^{p_0(\alpha r)} p\varepsilon(2p-1-p\varepsilon)f(p)dp)$. The first term is positive, the second term can be positive or negative, but is zero if $\varepsilon = 0$. □

If the bureaucrat chooses the low bribe, volatility remains unchanged if the signal is imprecise and increases if the signal is very precise.

7 Discussion

In this section the predictions of the model for the effect of corruption on risk are discussed for both a precise signal as well as for the aggregate effects on output and volatility.

7.1 Empirical evidence for the model's prediction if the signal is perfectly precise

This section compares the model's predictions for the case of a perfectly precise signal, $\varepsilon = 0$, and the existing empirical evidence. According to the model, firm behavior is optimal if the firm is the long-run player, while it is distorted if the bureaucrat is the long-run player. One interpretation of the firm as long-lived player is that of an international firm entering a corrupt country after it has built a reputation abroad or that of an old firm having built a reputation over time. Similarly, a long-lived bureaucrat could be a corrupt bureaucrat operating in a certain district for years. The model then predicts that sufficiently patient firms, which have had an opportunity to build a reputation for a certain project over time, choose the efficient project. Corruption therefore only results in distortions if the bureaucrat is the long-run player or the long-lived firm is not sufficiently patient.

Under some conditions,[13] it also follows from the model that long-lived firms pay lower bribes. This should be the case for foreign firms, which usually operate in their home country before moving abroad, as well as for old firms, which have operated in the market for some time. O'Toole and Tarp (2014) indeed find that bribery has no negative effect on the investment efficiency of large or foreign-owned firms.[14] The negative effect of bribery is instead largest for domestic small and medium enterprises. That long-lived firms pay lower bribes is confirmed by Gauthier and Goyette (2014) who find a negative effect of firm age on the bribe level in Uganda. Hence, the model's predictions regarding the level of the bribe for long- and short-lived firms are confirmed empirically.

If there is more than one firm in the market, the model predicts a different spread of bribes depending on the identity of the long-run player. If the firms are the long-run players, the bureaucrats choose both the high and the low bribe depending on the firms' success probability. If the bureaucrat is the long-run player, he chooses the same bribe for all firms. Hence, the spread of bribes is higher if the firms are the long-run players. We have seen that firms choose the efficient project if they are the long-run player. Hence, according to the model, if firms are the long-run players, the spread of bribes is higher and firms' investment choices are efficient. If the longevity of players is unobservable, empirically a higher dispersion of bribes should be positively correlated with firm performance. This is confirmed by Hanousek and Kochanova (2016).

7.2 Empirical evidence for the model's predictions regarding output and volatility

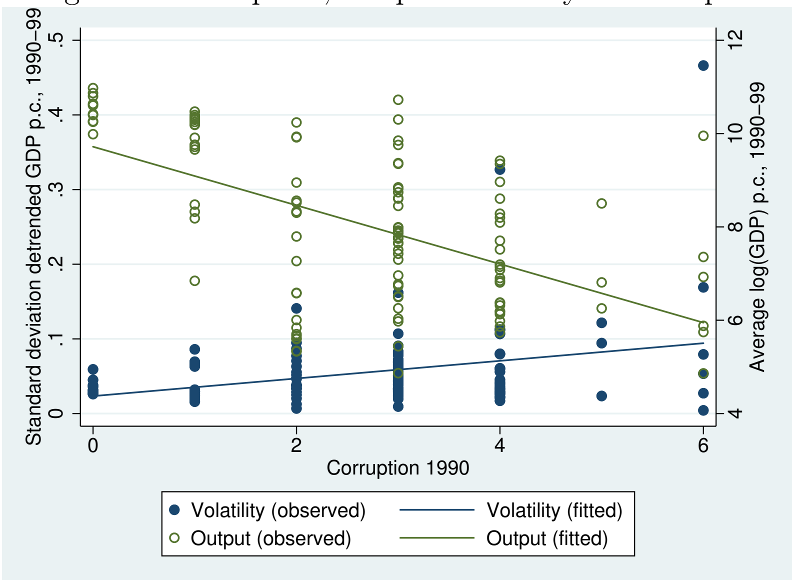
The model predicts that corruption leads to a fall in output and an increase in volatility if the firm has a lot of information about the project's success when paying the bribe.

Figure 8 provides suggestive evidence of a positive relation between corruption and risk. The measure of corruption is taken from the International Country Risk Guide of the Political Risk Services Group and has been recoded such that higher values of the variable depict higher levels of corruption. Output is measured as the average of the logarithm of real GDP per capita over the time period 1990-1999. The data on GDP comes from the World Bank. It can be seen that higher levels of corruption in the year 1990 are associated with lower average output levels. As

[13]If the bureaucrat is the long-run player and optimally chooses the high bribe while the safe project is efficient.

[14]It seems to be reasonable that large firms are, on average, older firms.

Figure 8: Corruption, Output Volatility and Output



Notes: The logarithm of real GDP per capita has been detrended using a quadratic trend over the period 1960-2014. Then the standard deviation of detrended GDP per capita has been computed over the period 1990-1999. The above regression lines are given by: volatility $= 0.019 + 0.014 *$ corruption, and output $= 9.915 - 0.727 *$ corruption.

a measure of risk, the standard deviation of detrended GDP per capita is used. It is easy to see that a higher level of corruption in the year 1990 is associated with higher volatility over the following two decades.

An alternative explanation for this observation could be that corruption leads to more volatile policy choices. Government spending, for example, is less pro-cyclical in industrialized countries (Lane 2003). Coate and Morris (1995) develop a model where politicians choose inefficient projects to disguise transfers to special interests. Similarly, corrupt politicians could choose more volatile policies in exchange for bribes. Given the empirical finding of corruption and lobbying being substitutes (Bennedsen, Feldmann and Lassen 2009), it is, however, unclear why volatile policies should only be used in the presence of corruption and not also in the presence of lobbying.

The prediction of the model that corruption decreases output is confirmed by the vast, primarily macroeconomic and empirical, literature relating corruption to growth. Mauro (1995) provides empirical evidence that corruption reduces growth by decreasing investment. O'Toole and Tarp (2014) empirically find that corruption reduces investment efficiency. A negative relation between corruption and growth is also uncovered by Ugur (2014) in a meta-analysis of 29 empirical studies. More recently, Hanousek and Kochanova (2016) find that firms which have to pay higher bribes grow more slowly. On the other hand, they also find that higher bribery dispersion positively contributes to firm growth. Also Evrensel (2010) finds that corruption increases growth rates. Méndez and Sepúlveda (2006) find that small levels of corruption maximize growth for, by their definition, "free" countries. This paper shows that corruption induces firms to choose riskier business projects characterized by the firm having a lot of information about the project's success when paying the bribe. Even though this usually leads to lower expected output, there can be instances where more risky projects turn out successful than expected. Hence, the occasional finding of a positive effect of corruption on growth can be explained by corruption inducing firms to follow inefficiently risky strategies which are successful despite the small chance.

The model provides a theoretical explanation for how institutions can influence volatility by illuminating the effect of a typical characteristic of weak institutions, corruption. Empirical findings regarding the impact of institutions on volatility are manifold. Weak institutions, as defined by a broad index, increase volatility (Malik and Temple 2009) as do more extractive institutions (Acemoglu, Johnson, Robinson and Thaicharoen 2003) and corruption (Evrensel 2010). On the other hand, democracy (Mobarak 2005; Klomp and de Haan 2009) and the number of interest groups can be related to lower volatility (Coates, Heckelman and Wilson 2007; Heckelman and Wilson 2014). Volatility is also higher in developing countries because they specialize in fewer and more volatile sectors, and are subject to larger aggregate shocks (Koren and Tenreyro 2007). Countries also experience higher volatility if they are more open to trade (Di Giovanni and Levchenko 2009), exporters of primary commodities, remote from the sea or in the tropics (Malik and Temple 2009). The model's prediction of high volatility and low output for business projects is also supported by the empirical finding of a negative relation between volatility and output growth. Higher volatility, measured as standard deviation of output growth, is associated with lower growth rates (Ramey and Ramey 1995). Lastly, Jetter (2014) finds a negative effect of volatility on growth in democracies.

8 Empirical analysis

In this section, I test the model's predictions empirically. Because the model makes a range of predictions about corruption and risk, I investigate a number of different hypotheses. First, I look at whether there is a relation between the level of corruption a firm experiences and the firm's risk, measured as change in sales. Second, I disentangle the effect of corruption on risk for different level of information of the bureaucrat and of the firm. Third, I look at the effect of corruption on firms' choices for different projects.

I use cross-sectional firm-level data from the World Bank Enterprise Survey. The data is collected using stratified random sampling. Firms are surveyed in different countries in different years and are stratified along industries. The sample period is from 2006 to 2018 and includes 138 countries. I restrict the analysis to manufacturing firms. The questionnaire also includes a set of questions about bribery and is therefore well suited to study corruption. Paunov (2016), for example, uses the same data set, supplemented with other data sources, to investigate the effect of corruption on patents and quality certificates.

I use two different measures to control for corruption. In order to measure the extent of corruption, I use the answer to the following question: "We've heard that establishments are sometimes required to make gifts or informal payments to public officials to 'get things done' with regard to customs, taxes, licenses, regulations, services etc. On average, what percent of total annual sales, or estimated total annual value, do establishments like this one pay in informal payments or gifts to public officials for this purpose?" In order to measure the incidence of corruption I use the number of bribe payments. Firms indicate whether a bribe payment was demanded or expected if they applied for e.g. a water connection or an operating license. I count the number of times where a firm declares that a bribe was expected in order to measure a firm's exposure to bribery. Firms applying for many licenses have a higher probability to be asked to pay a bribe because they interact with public officials more often. Because the model's predictions rest on the assumption that firms are faced with the same number of bribe demands, I control for the incidence of bribery measured as the number of expected bribe payments.

Table 3 shows some descriptive statistics. We see that firms reporting a positive informal payment declare that this amounts to 8% of annual sales. Many firms, however, report this payment to be non-existent decreasing the mean of all firms to below 1%.

The model predicts that corruption influences how much risk firms are willing to accept.

Hypothesis 1. *Corruption has an impact on the level of risk, firms are willing to take.*

In order to measure risk, it is necessary to identify a reasonable proxy for the level of risk a firm assumes. The data set includes information about firms' sales in the year of the survey and three years earlier.[15] A riskier strategy should result in larger

[15] Because firms are not surveyed every year, using the sales of the previous year is not possible.

	Mean	Std.Dev.	Min	Max
$\Delta(log(sales_{ijlct}))$	0.536	1.052	0.0	21.1
New product	0.663	0.473	0.0	1.0
Transport loss	0.983	4.224	0.0	100.0
Informal payments	1.214	5.517	0.0	100.0
Payment if positive	7.160	11.700	0.0	100.0
Times of bribe payment	1.005	0.701	0.0	8.0
Firm age	22.433	17.044	3.0	226.0
log(employment$_{ijlct}$)	3.542	1.411	0.0	10.3
% owned by largest owner	76.948	27.042	0.0	100.0
Share owned by government (1=yes)	0.016	0.124	0.0	1.0
% of sales exported	9.948	24.643	0.0	100.0
log(sales$_{ijlct}$)	17.215	3.138	3.9	33.8
N	34724			

Table 3: Summary Statistics

changes in sales. The model predicts that corruption should influence the change in sales both positively and negatively. Because I am interested in the magnitude of the change, I look at the absolute value of the change in sales. Dropping the absolute value would potentially pick up the effect of corruption on sales growth. The baseline specification is given by

$$\Delta(log(sales_{ijlct})) = cons + \beta_c corr_{ijlct} + \beta_x X_{ijlct} + \gamma_j + \gamma_l + \gamma_c + \gamma_t + \varepsilon_{ijlct} \quad (22)$$

where $\Delta(log(sales_{ijlct}))$ is the absolute value of the difference in the logarithm of sales of firm i in industry j in location l in country c from time $t-3$ to t. The variable of interest is $corr_{jlct}$ which is the percentage of total annual sales paid in informal payments. X_{ijlct} is a vector of firm control variables, γ_j is an industry fixed effect, γ_l is a location fixed effect, γ_c is a country fixed effect and γ_t is a time fixed effect.

A firm's industry is based on its main product and defined by ISIC Code Revision 3.1. and the firm's location is measured as the size of the location where the firm is situated.[16] Firm control variables include firm age because older firms are usually less volatile, firm size, measured by both the number of employees and the level of sales, and ownership structure. I also control for exports and product innovation as these could also impact firm risk. Because the model assumes that all firms pay a bribe exactly once while some firms are faced with bribe demands

[16] There are five different categories: main business city, cities with a population over 1 million, over 250,000 to 1 million, over 50,000 to 250,000 and less than 50,000.

more often, I also control for the incidence of bribery.[17] In the first specification, I only include fixed effects, in the second specification I add firm control variables and in the third specification, I use average values to measure the extent and intensity of corruption. The average level of corruption, the amount of informal payments, is calculated by taking the average of informal payments over all firms in country c, year t and location l. The average level for the incidence of bribe payments is calculated similarly. Following Paunov (2016), using the average values instead of the firm values is supposed to minimize endogeneity concerns.[18]

Table 4 shows that firms that pay higher informal payments experience larger changes in their sales growth in absolute terms. The effect is small but highly significant. The effect remains significant and positive when firm control variables are included. In the third specification, I look at whether the average level of corruption influences the absolute value of a firm's sales growth. We see that the coefficient decreases in size but remains significant.

According to the model, the effect of corruption on risk depends on information.

Hypothesis 2. *Corruption decreases risk whenever the bureaucrat has a lot of information about the firm. Corruption increases risk whenever the bureaucrat has little information about the firm and the firm has a lot of information about the project.*

In order to capture whether a firm has to make informal payments when more information about the project's success is already available to the firm, I construct a binary variable, *paylate*. Firms indicate how many inspections were conducted by public officials over the previous 12 months. At an inspection the firm should already have more information about the success of the project. Because firms only indicate whether an informal payment was expected at an inspection but not whether it was also made, I proxy actual payments by using whether a firm indicates that informal payments are usually positive. The number of firms claiming that a payment at inspection was expected or requested but that informal payments are zero is non-negligible.[19] The dummy variable *paylate* then indicates whether a firm made an informal payment at an inspection. It takes the value of 1 for firms who both report a positive fraction of annual sales usually paid by similar firms and that an informal

[17] Firms can report bribe demands for inspections in the previous 12 months and for license applications in the previous two years.

[18] One would prefer to use the values for corruption in the previous years but because of the long survey intervals, e.g. Albania is surveyed in 2007 and then again in 2013, this would presumably pick up the effect of corruption on growth and not on volatility.

[19] Restricting attention to firms that have existed for at least three years and for which control variables are available these are 1663 firms.

	(1) $\Delta(log(sales_{ijlct}))$	(2) $\Delta(log(sales_{ijlct}))$	(3) $\Delta(log(sales_{ijlct}))$
Informal payments	0.00881***	0.00837**	
	(2.62)	(2.48)	
Times of bribe payment		0.0214	
		(1.05)	
Informal payments (lct)			0.0240**
			(2.41)
Times of bribe payment (lct)			-0.193
			(-1.35)
log(age$_{ijlct}$)		-0.382***	-0.377***
		(-5.87)	(-5.81)
(log(age$_{ijlct}$))^2		0.0566***	0.0552***
		(5.27)	(5.16)
log(employment$_{ijlct}$)		-0.0140	-0.0122
		(-0.88)	(-0.77)
% owned by largest owner		-0.000207	-0.000224
		(-0.80)	(-0.87)
Share owned by government (1=yes)		0.128	0.132
		(1.05)	(1.08)
% of sales exported		0.000709**	0.000729**
		(2.06)	(2.13)
New product (1=yes)		0.0376**	0.0395***
		(2.47)	(2.60)
log(sales$_{ijlct}$)		0.0169	0.0157
		(1.33)	(1.24)
Constant	0.477***	0.782***	0.991***
	(4.05)	(3.16)	(3.48)
Industry fixed effects	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Location fixed effects	No	Yes	Yes
Observations	20238	20238	20238
R^2	0.199	0.203	0.202

t statistics in parentheses
Robust standard errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 4: Effect of informal payments on firm risk

payment was expected or requested upon inspection.

$$\Delta(log(sales_{ijlct})) = cons + \beta_l paylate + \beta_x X_{ijlct} + \gamma_j + \gamma_l + \gamma_c + \gamma_t + \varepsilon_{ijlct} \quad (23)$$

The bureaucrat can acquire information about a firm by inspecting this firm. Over time, it should also be possible to collect more information about a firm. On the other hand, the bureaucrat has only little information about firms that have not been operating very long and that he has not had the opportunity to inspect very often. The specification is therefore assessed for two different subsets of firms. One subset corresponds to the bureaucrat not being well informed about the firms' characteristics. This subset is characterized by young firms with few inspections.[20] The second subset corresponds to firms the bureaucrat has a lot of information about. This subset is characterized by old firms with many inspections.[21]

Table 5 shows the different effects of corruption on risk depending on the information, the bureaucrat has about the firm. Specification (1) shows that firms about which the bureaucrat lacks information experience more volatility when firms pay a bribe at an inspection. After including firm control variables, the coefficient still has the expected sign but is not significant. One reason for this could be that the number of firms for which all control variables are available is small. Specifications (3) and (4) show the relation between corruption and risk if both the firm is informed about the project and the bureaucrat is informed about the firm. The effect is of the predicted direction: If the bureaucrat has a lot of information about the firm, firms experience smaller changes in sales indicating that they choose safer strategies. The effect is significant and increases in size after including firm control variables.

We continue by looking at the effect of corruption on different project choices. Apart form the timing of a bribe demand within a project, information also differs for firms according to project type. In order to capture this, I use two different kinds of projects. First, I look at innovation as an example of a project where firms lack information about the final outcome. Then I look at transportation as an example for firms having more information about the project's result.

Whenever the firm has only little information about the project, the model predicts that the level of risk will either be unaffected or decrease.

Hypothesis 3. *Corruption either decreases innovation or does not influence innovative activity.*

[20] Firms are required to have: log(Age) $\leq 10^{\text{th}}$ percentile(log(Age)) and inspections $\leq 50^{\text{th}}$ percentile(inspections).

[21] Firms are required to have: log(Age) $\geq 50^{\text{th}}$ percentile(log(Age)) and inspections $\geq 90^{\text{th}}$ percentile(inspections).

dep.var. $\Delta(log(sales_{ijlct}))$	Young firms, few inspections (1)	(2)	Old firms, many inspections (3)	(4)
Pay at inspection	1.031*	0.962	-0.393**	-0.404**
	(1.67)	(1.22)	(-2.38)	(-2.06)
Informal payments		0.0353		-0.0119
		(1.45)		(-0.87)
Times of bribe payment		0.0931		0.0354
		(0.65)		(0.67)
log(age$_{ijlct}$)		-0.450		-0.0165
		(-1.03)		(-0.23)
log(employment$_{ijlct}$)		0.0617		0.0202
		(0.70)		(0.36)
% owned by largest owner		-0.000955		-0.00107
		(-0.54)		(-1.00)
Share owned by government (1=yes)		-0.0303		-0.0191
		(-0.03)		(-0.05)
% of sales exported		-0.00234		0.00295
		(-1.32)		(1.15)
New product (1=yes)		0.0525		0.0753
		(0.40)		(0.79)
log(sales$_{ijlct}$)		-0.00997		-0.0339
		(-0.18)		(-0.87)
Constant	4.471***	0.565	0.324	0.942
	(5.15)	(0.46)	(0.88)	(1.48)
Industry fixed effects	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes
Observations	670	597	907	907
R^2	0.435	0.354	0.322	0.327

t statistics in parentheses

Robust standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 5: Effect of timing of payment

As a measure for innovation I use the introduction of a new product by the firm. The data set differentiates between new products that are only new to the firm and product that are also new to the market. I focus on the second category because this entails more risk than copying an already existing product. Because the dependent variable is a dummy variable, taking the value of 1 if the firm introduced a new product in the past three years, I use a probit model to estimate the effect of corruption on innovation. The specification is given by

$$new\ product = cons + \beta_c corr_{ijlct} + \beta_x X_{ijlct} + \gamma_j + \gamma_l + \gamma_c + \gamma_t + \varepsilon_{ijlct} \quad (24)$$

where *new product* indicates whether the firm issued a new product to the market. Specification (1) in table 6 shows that there is no effect of the level of informal payments on the development of new products. When including firm control variables in specification (2), this continues to be true. It also seems that firms that have to pay bribes very often, i.e. experience a higher incidence of bribery, innovate more often. Because the calculation of the incidence variable is based on the number of applications for licenses in the two years before the survey, it could be that firms that plan to develop a new product also need more new licenses.[22] The model does not make a prediction about the effect of a higher incidence of bribery on innovation and we will therefore not explore this further. Specification (3) shows the effect of average values for bribe level and bribery incidence on firm innovation. We see that an environment where high bribes are paid decreases firms' propensity to introduce new products. Nevertheless, both no effect and a negative effect are consistent with the model's prediction. We also observe that firms in an environment where bribes have to be paid very often are more likely to introduce a new product. This could again be related to the way the variable indicating bribery incidence is constructed.[23] The predictions of the model regarding the effect of corruption on innovation are thus confirmed.

As an example for a project about which the firm has a lot of information I use transportation. A riskier transportation strategy should result in more broken goods. The dependent variable is therefore the percentage of goods lost due to spoilage or leakage during transportation. Because I consider manufacturing firms whose primary business is not transportation, the bureaucrat should not have very precise information about the firm's profit in this case.

[22] In order to construct the incidence variable, I count incidences of bribe payment. If a firm did not apply for a license, there was no scope to ask for a bribe and this is treated as if the firm was not asked for a bribe after applying.

[23] It could be that firms introduce innovations at a similar time.

	(1) New product	(2) New product	(3) New product
Informal payments	-0.000207	-0.000893	
	(-0.07)	(-0.29)	
Times of bribe payment		0.0478**	
		(2.07)	
Informal payments (lct)			-0.0337*
			(-1.69)
Times of bribe payment (lct)			0.939***
			(3.52)
Foreign technology (1=yes)		0.265***	0.263***
		(7.16)	(7.10)
% of foreign ownership		0.0000874	0.000102
		(0.15)	(0.18)
log(age$_{ijlct}$)		0.0344*	0.0338*
		(1.78)	(1.75)
log(sales$_{ijlct}$)		0.0159**	0.0151**
		(2.21)	(2.11)
Constant	0.463	0.00125	-0.958
	(0.94)	(0.00)	(-1.59)
Industry fixed effects	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes
Observations	9451	9451	9451
R^2			

t statistics in parentheses

Robust standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 6: Effect on innovation

	(1) Transport loss	(2) Transport loss	(3) Transport loss
Informal payments	0.0495***	0.0358***	
	(4.69)	(3.84)	
Times of bribe payment		0.102**	
		(1.98)	
Informal payments (lct)			0.130***
			(3.26)
Times of bribe payment (lct)			-0.503
			(-1.01)
Transport theft		0.354***	0.357***
		(7.37)	(7.42)
log(sales$_{ijlct}$)		-0.0567***	-0.0583***
		(-4.89)	(-5.09)
% of sales exported		-0.00234***	-0.00219***
		(-3.17)	(-2.95)
Constant	4.308**	3.612**	3.909**
	(2.20)	(2.55)	(2.56)
Industry fixed effects	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes
Observations	27494	27494	27494
R^2	0.066	0.155	0.153

t statistics in parentheses
Robust standard errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 7: Effect on transportation

Hypothesis 4. *Corruption increases risk in transportation resulting in more spoilage.*

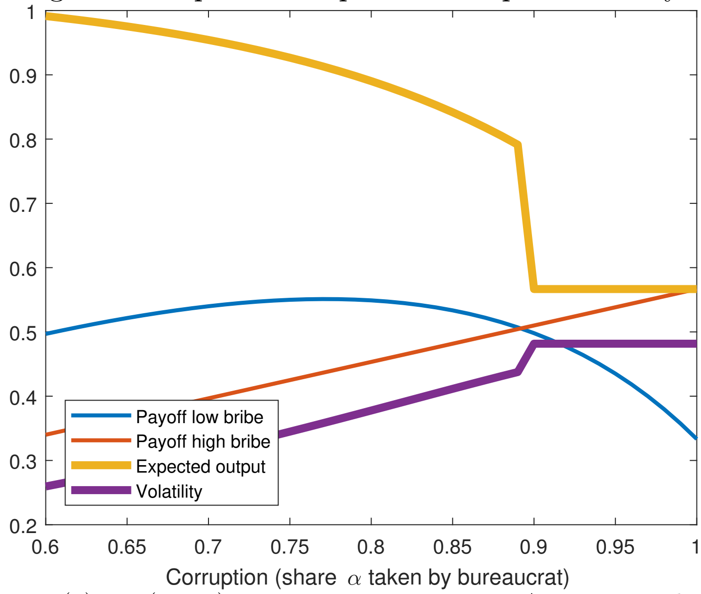
The specification to be estimated is given by

$$transport\ loss = cons + \beta_c corr_{ijlct} + \beta_x X_{ijlct} + \gamma_j + \gamma_l + \gamma_c + \gamma_t + \varepsilon_{ijlct} \quad (25)$$

where $transport\ loss$ indicates the percentage of goods lost during transport. We see in table 7 that higher bribe payments lead to more goods lost during transportation. This confirms the model's prediction.

9 Conclusion

Figure 9: Expected output and output volatility



Notes: $g(p) = 2(1-p)$, $r = 1$, $R = 1.7$, $\varepsilon = 0$. Aggregate volatility for the small bribe is given by $\sigma^2(\alpha r) = 2R^2(\frac{1}{12} - \frac{1}{2}(p^*)^2 + \frac{2}{3}(p^*)^3 - \frac{1}{4}(p^*)^4)$ and for the high bribe is given by $\sigma^2(\alpha R) = \frac{1}{6}R^2$.

This paper has addressed the influence of corruption in risk. We have seen that corruption usually decreases the level of risk if the firm's project choice is observed by the bureaucrat. An increase in risk results if the bureaucrat's bribe demand is independent of the firm's project and the firm has very precise information about the project's success before paying the bribe. The differing extent of information the firm has drives the different effects of corruption on risk if the bureaucrat's bribe is independent of the firms project choice. If the firm has precise information, the bribe is paid conditional on the project's success leading to excessive risk. If the firm has very imprecise information, the bribe is paid independently of the project's success. This leads to excessively safe project choices if the bribe is high and efficient choices if the bribe is low. The effect of corruption therefore depends both on the amount of information the firm has and on the identity of the long-lived agent.

The model shows that high levels of the bribe can be positively correlated with high levels of output. The reason for this finding, however, is that underlying conditions, modeled as distributions of the firms' success probabilities, influence the level of both bribe and output. When assessing the effects of corruption on output empirically, it is therefore necessary to appropriately control for economic conditions.

The model also shows a potential tradeoff between fostering innovation or production that arises from a reduction in the acceptable level of corruption. The current Chinese approach in the words of Xi Jinping is the "fighting of tigers and flies at the same time, resolutely investigating law-breaking cases of leading officials and also earnestly resolving the unhealthy tendencies and corruption problems which happen all around people" (Branigan 2013). If small domestic firms account for the majority of innovations, reducing corruption can lead to an initial drop in innovative activity. Innovations, however, will recover after this initial drop.

The predictions of the model regarding corruption and risk were tested using data from the World Bank Enterprise Survey. We have seen that corruption affects the firms' level of risk. If firms are informed about the outcome of a project, firms choose more risk when the bureaucrat is ill informed and less risk when the bureaucrat is well informed about their profit. Similarly, firms choose riskier transportation strategies while their propensity to introduce a new product to the market is unaffected. The predictions of the model are thus confirmed empirically.

We have also seen that corruption usually leads to lower expected output. Volatility increases if the firm observes a sufficiently precise signal before paying the bribe. The effect of corruption on expected output and output volatility for a perfectly precise signal is depicted in Figure 9. As the extent of corruption, measured as the share α, increases, expected output falls while output volatility increases. Once the bureaucrat switches to ask for the high bribe, expected output reaches its lowest level and volatility reaches its highest level. This fits the empirical evidence of a negative relation between the quality of institutions and volatility (Acemoglu et al. 2003; Malik and Temple 2009; Evrensel 2010). Given that corruption is more widespread in poor countries (Treisman 2000), the model additionally provides a possible explanation for the observed specialization of developing countries in fewer and more volatile sectors (Koren and Tenreyro 2007). According to the model, combatting corruption reduces risk in everyday production which both stabilizes and increases output.

References

Acemoglu, D., Johnson, S., Robinson, J., Thaicharoen, Y. (2003). Institutional causes, macroeconomic symptoms: Volatility, crises and growth, *Journal of Monetary Economics 50*, 49-123.

Ahlin, C., Bose, P. (2007). Bribery, inefficiency, and bureaucratic delay, *Journal of Development Economics 84*, 465-486.

Amir, R., Burr, C. (2015). Corruption and socially optimal entry, *Journal of Public Economics 123*, 30-41.

Bennedsen, M., Feldman, S. E., Lassen, D. D. (2009). Strong firms lobby, weak firms bribe: A survey-based analysis of the demand for influence and corruption, *SSRN: https://ssrn.com/abstract=1503028*.

Blackburn, K., Forgues-Puccio, G. F., (2009). Why is corruption less harmful in some countries than in others?, *Journal of Economic Behavior & Organization 72*, 797-810.

Bliss, C., Di Tella, R. (1997). Does competition kill corruption?, *Journal of Political Economy 105*(5), 1001-1023.

Branigan, T. (22 January 2013). "Xi Jinping vows to fight 'tigers' and 'flies' in anti-corruption drive", *The Guardian*, `https://www.theguardian.com/world/2013/jan/22/xi-jinping-tigers-flies-corruption` (accessed 10 November 2017).

Cadot, O. (1987). Corruption as a gamble, *Journal of Public Economics 33*, 223-244.

Célimène, F., Dufrénot, G., Mophou, G., N'Guérékata, G. (2016). Tax evasion, tax corruption and stochastic growth, *Economic Modelling 52*, 251-258.

Choi, J. P., Thum, M. (2003). The dynamics of corruption with the ratchet effect, *Journal of Public Economics 87*, 427-443.

Choi, J. P., Thum, M. (2004). The economics of repeated extortion, *The RAND Journal of Economics 35*(2), 203-223.

Coate, S., Morris, S. (1995). On the form of transfers to special interests, *Journal of Political Economy 103*(6), 1210-1235.

Coates, D., Heckelman, J., Wilson, B. (2007). Special-interest groups and volatility, *Economics Bulletin 15*(18), 1-13.

Dhingra, A. (13 May 2012). "'Publishers said paanwalas not supposed to write'", *The Sunday Guardian*, http://www.sunday-guardian.com/investigation/publishers-said-paanwalas-not-supposed-to-write (accessed 10 November 2017).

Di Giovanni, J., Levchenko, A. A. (2009). Trade openness and volatility, *The Review of Economics and Statistics 91*(3), 558-585.

Ely, J. C., Välimäki, J. (2003). Bad reputation, *The Quarterly Journal of Economics 118*(3), 785-814.

Evrensel, A. Y. (2010). Corruption, growth and growth volatility, *International Review of Economics and Finance 19*, 501-514.

Fudenberg, D., Levine, D. K. (1989). Reputation and equilibrium selection in games with a patient player, *Econometrica 57*(4), 759-778.

Fudenberg, D., Levine, D. K. (1992). Maintaining a reputation when strategies are imperfectly observed, *The Review of Economic Studies 59*(3), 561-579.

Fudenberg, D., Kreps, D. M., Maskin, E. S. (1990). Repeated games with long-run and short-run players, *The Review of Economic Studies 57*(4), 555-573.

Gauthier, B., Goyette, J. (2014). Taxation and corruption: Theory and firm-level evidence from Uganda, *Applied Economics 46*(23), 2755-2765.

"Bangladesh factory collapse blamed on swampy ground and heavy machinery" (23 May 2013), *The Guardian*, https://www.theguardian.com/world/2013/may/23/bangladesh-factory-collapse-rana-plaza (accessed 12 June 2017).

Hanousek, J, Kochanova, A. (2016). Bribery environments and firm performance: Evidence from CEE Countries, *European Journal of Political Economy 43*, 14-28.

Harstad, B., Svensson, J. (2011). Bribes, lobbying, and development, *The American Political Science Review 105*(1), 46-63.

Heckelman, J. C., Wilson, B. (2014). Interest groups and the "rise and decline" of growth, *Southern Economic Journal 81*, 435-456.

Hodal, K. (27 September 2013). "How eating dog became big business in Vietnam", *The Guardian*, https://www.theguardian.com/world/2013/sep/27/eating-dog-vietnam-thailand-kate-hodal (accessed 10 November 2017).

Jetter, M. (2014). Volatility and growth: Governments are key, *European Journal of Political Economy 36*, 71-88.

Klomp, J., de Haan, J. (2009). Political institutions and economic volatility, *European Journal of Political Economy 25*, 311-326.

Koren, M., Tenreyro, S. (2007). Volatility and development, *The Quarterly Journal of Economics 122*(1), 243-287.

Lane, P. R. (2003). Business cycles and macroeconomic policy in emerging market economies, *International Finance 6*(1), 89-108.

Leung, C. K. Y., Tang, S. H. K., Groenewold, N. (2006). Growth volatility and technical progress: A simple rent-seeking model, *Journal of Economics 88*(2), 159-178.

Malik, A., Temple, J. R. W. (2009). The geography of output volatility, *Journal of Development Economics 90*, 163-178.

Mauro, P. (1995). Corruption and growth, *The Quarterly Journal of Economics 110*(3), 681-712.

Méndez, F., Sepúlveda, F. (2006). Corruption, growth and political regimes: Cross country evidence, *European Journal of Political Economy 22*, 82-98.

Mobarak, A. M. (2005). Democracy, volatility, and economic development, *The Review of Economics and Statistics 87*(2), 348-361.

Mulye, P. (28 November 2014). "Mumbai street vendors: 'Cops work in shifts, so I have to bribe each one'", *The Guardian*, `https://www.theguardian.com/cities/2014/nov/28/mumbai-street-vendors-cops-bribery` (accessed 10 November 2017).

Murphy, K. M., Shleifer, A., Vishny, R. W. (1993). Why is rent-seeking so costly to growth?, *The American Economic Review Papers and Proceedings 83*(2), 409-414.

O'Toole, C. M., Tarp, F. (2014). Corruption and the efficiency of capital investment in developing countries, *Journal of International Development 26*, 567-597.

Ott, K. (18 August 2015). "Dunkle Geschäfte in Sankt Petersburg", *Süddeutsche Zeitung*, `http://www.sueddeutsche.de/wirtschaft/verdacht-auf-schmiergeld-dunkle-geschaefte-in-st-petersburg-1.2610300` (accessed 12 June 2017).

Paunov, C. (2016). Corruption's asymmetric impacts on firm innovation, *Journal of Development Economics 118*, 216-231.

Ramey, G., Ramey, V. A. (1995). Cross-country evidence on the link between volatility and growth, *The American Economic Review 85*(5), 1138-1151.

Shleifer, A., Vishny, R. W. (1993). Corruption, *The Quarterly Journal of Economics 108*(3), 599-617.

Søreide, T. (2009). Too risk averse to stay honest? Business corruption, uncertainty and attitudes toward risk, *International Review of Law and Economics 29*, 388-395.

Svensson, J. (2003). Who must pay bribes and how much? Evidence from a cross section of firms, *The Quarterly Journal of Economics 118*(1), 207-230.

Treisman, D. (2000). The causes of corruption: a cross-national study, *Journal of Public Economics 76*, 399-457.

Ugur, M. (2014). Corruption's direct effects on per-capita income growth: A meta-analysis, *Journal of Economic Surveys 28*(3), 472-490.

Yoo, S. H. (2008). Petty corruption, *Economic Theory 37*(2), 267-280.

A Proofs

A.1 Proof of Proposition 2

Proof. The effect of a change in ε on the cutoff-values is given by:

$$\begin{aligned}
\frac{\partial p_0}{\partial \varepsilon} =& -\frac{\alpha(1-\alpha)}{(\varepsilon(1-\alpha)+(1-\varepsilon)\alpha)^2} < 0 \\
\frac{\partial p_1}{\partial \varepsilon} =& \frac{\alpha(1-\alpha)}{((1-\varepsilon)(1-\alpha)+\alpha\varepsilon)^2} > 0 \\
\frac{\partial p_h}{\partial \varepsilon} =& 0 \\
\frac{\partial p_m}{\partial \varepsilon} =& \frac{R(1-\alpha)(\alpha(R-r)+(1-\alpha)r)}{((1-\varepsilon)R(1-\alpha)+\alpha\varepsilon R)^2} > 0
\end{aligned}$$

$p \leq p_1$: The efficient project is chosen if $\frac{r}{R} \geq p_1$. This inequality is satisfied if $\frac{r(1-\alpha)}{\alpha(R-r)+r(1-\alpha)} \geq \varepsilon$ which is more likely to hold for small ε.

$p_1 < p \leq p_0$: For $\varepsilon > \bar{\varepsilon}$, all firms choose safe. This is efficient if $\frac{r}{R} \geq p_0$. $\frac{r}{R} \geq p_0$ if $\varepsilon > \frac{\alpha(R-r)}{\alpha R+r(1-2\alpha)}$. Because $\bar{\varepsilon} < \frac{\alpha(R-r)}{\alpha R+r(1-2\alpha)}$, project choices are only efficient if $\varepsilon > \frac{\alpha(R-r)}{\alpha R+r(1-2\alpha)}$. For $\varepsilon < \bar{\varepsilon}$, firms choose risky if $p \geq p_m$. There are no inefficiently safe choices if $\frac{r}{R} \geq p_m$. This can be rewritten as $\varepsilon(\alpha(r-R)-r(1-\alpha)) \geq 0$. Hence, risk is at the efficient level if $\varepsilon^* = 0$. For $\varepsilon > 0$, this equation becomes $\alpha(r-R)-r(1-\alpha) < 0$. Hence, $p_m > \frac{r}{R}$. Because $\frac{\partial p_m}{\partial \varepsilon} > 0$, an increase in ε increases the extent of inefficient choices.

$p_0 < p$: For $\varepsilon > \bar{\varepsilon}$, firms choose risky if $p \geq p_h$. There are two many safe choices because $p_h > \frac{r}{R}$. Because $\frac{\partial p_0}{\partial \varepsilon} < 0$ and $\frac{\partial p_h}{\partial \varepsilon} = 0$, increasing ε induces more firms to pay the bribe at both I_0 and I_1, but does not change the threshold level where they choose the risky project. More firms make inefficiently safe choices. For $\varepsilon < \bar{\varepsilon}$, firms choose risky if $p > p_0$. There are too many risky choices if $\frac{r}{R} > p_0$. This holds if $\varepsilon > \frac{\alpha(R-r)}{\alpha R+r(1-2\alpha)}$. But $\varepsilon < \bar{\varepsilon} < \frac{\alpha(R-r)}{\alpha R+r(1-2\alpha)}$. Hence, there are too many safe choices. Because $\frac{\partial p_0}{\partial \varepsilon} < 0$, an increase in ε leads to more inefficiently safe choices.

□

A.2 Proof of Proposition 3

Proof. $p \geq max\{p_0(\alpha r), \frac{r}{R}\}$: Firms always pay if $p > p_0(\alpha r)$. They choose risky if $p \geq \frac{r}{R}$. All firms, that always pay, always choose risky if $p_0(\alpha r) > \frac{r}{R}$. This is the case if $\tilde{\varepsilon} > \varepsilon$. We can check whether $\tilde{\varepsilon} > \frac{1}{2}$, this can be simplified to $\alpha > 1$ and therefore does not always hold.

Firms with $p_0(\alpha r) < p$ choose the risky project if $\mathbb{E}(risky) \geq \mathbb{E}(safe)$. This can be simplified to $pR - \alpha r > r - \alpha r$ which holds if $pR > r$ and gives the efficient project choice.

Firms with $p < p_1(\alpha r)$ always choose the safe project. The safe choice is efficient if $\frac{r}{R} > p$. If $p_1(\alpha r) > \frac{r}{R}$, too many safe choices are made.

This can be rewritten as

$$\varepsilon \geq \frac{R - \alpha r}{R(1+\alpha) - 2\alpha r}$$

This condition is never satisfied because $\frac{R-\alpha r}{R(1+\alpha)-2\alpha r} > \frac{1}{2} \geq \varepsilon$.

$p_m < p < p_0(\alpha r)$: $p_0(\alpha r) > p_m$ if $\tilde{\varepsilon} \geq \varepsilon$. We can check whether $\frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r} \geq \frac{1}{2}$. Because this inequality can be rewritten as $\alpha \geq 1$, both $p_0(\alpha r) > p_m$ and $p_0(\alpha r) < p_m$ are possible.

All firms that pay the bribe if at I_1 choose the safe project if $p_m > p_0(\alpha r)$ or $\varepsilon > \tilde{\varepsilon}$. If $p_0 \geq \frac{r}{R}$, there would be too many safe choices. But since $\varepsilon > \tilde{\varepsilon}$ implies $\frac{r}{R} > p_0(\alpha r)$, firms' choices are efficient. Note that this group does not exist if $\varepsilon = 1/2$ because then $p_1 = p_0$. If $\tilde{\varepsilon} > \varepsilon$, firms that pay the bribe if at I_1 choose the efficient project if $p_m(\alpha r) = \frac{r}{R}$. Lastly, $p_m(\alpha r) \geq \frac{r}{R}$ if $\varepsilon \geq \frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r}$. If $\tilde{\varepsilon} \equiv \frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r} \geq \varepsilon$, $p_0(\alpha r) > r/R$ and $p_0(\alpha r) > p_m$. □

A.3 Proof of Proposition 4

Proof. The risky project is chosen by all firms but only with $p \geq p_1(\alpha R) = \frac{\alpha\varepsilon}{(1-\alpha)(1-\varepsilon)+\alpha\varepsilon}$ ever pay the bribe. Corruption leads to more risk if $\frac{r}{R} > p_1(\alpha R)$ which holds if $\hat{\varepsilon} > \varepsilon$. The efficient risky choice is made if $\frac{r}{R} = \frac{\alpha\varepsilon}{(1-\alpha)(1-\varepsilon)+\alpha\varepsilon}$. This holds if

$$\frac{r(1-\alpha)}{\alpha(R-r) + r(1-\alpha)} = \varepsilon^*$$

with $\frac{\partial \varepsilon^*}{\partial \alpha} = \frac{r(r-R)}{(\alpha(R-r)+r(1-\alpha))^2} < 0$. The optimal ε^* decreases in α. With a higher α, firm needs to be surer that the project is successful to be willing to bribe. We still have that $p_0(\alpha R) \geq p_1(\alpha R)$ because this can be simplified to $1 \geq 2\varepsilon$.

$$\frac{\partial p_1(\alpha R)}{\partial \varepsilon} = \frac{\alpha(1-\alpha)(1-\varepsilon)}{((1-\varepsilon)(1-\alpha)+\alpha\varepsilon)^2} > 0$$

A higher ε decreases risk but at the same time fewer firms still produce.

It always holds that $p_0(\alpha R) > \frac{r}{R}$ because this can be rewritten as

$$\frac{\alpha(R-r)}{r(1-2\alpha) + \alpha R} > \varepsilon$$

This is satisfied for all ε because $\frac{\alpha(R-r)}{r(1-2\alpha)+\alpha R} > \frac{1}{2}$ holds if $\alpha R > r$. $\square$

A.4 Proof of Proposition 5

Proof. $\varepsilon > \tilde{\varepsilon}$: Define $A = \frac{r}{R} - \int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp - \int_{p_0(\alpha R)}^{1} f(p)dp$. Remember that $\frac{\partial p_0}{\partial \varepsilon} < 0$, $\frac{\partial p_1}{\partial \varepsilon} > 0$ and $\frac{\partial \mu}{\partial \varepsilon} = 1 - 2p$. $\frac{\partial p_0}{\partial R} = \frac{\partial p_1}{\partial R} = \frac{\partial p_0}{\partial r} = \frac{\partial p_1}{\partial r} = 0$.

$$\begin{aligned}
\frac{\partial A}{\partial r} =& \frac{1}{R} > 0 \text{ and } \frac{\partial A}{\partial R} = -\frac{r}{R}^2 < 0 \\
\frac{\partial p_0}{\partial \alpha} =& \frac{(1-\varepsilon)\varepsilon}{(\varepsilon(1-\alpha)+(1-\varepsilon)\alpha)^2} \geq 0 \\
\frac{\partial p_1}{\partial \alpha} =& \frac{(1-\varepsilon)\varepsilon}{((1-\varepsilon)(1-\alpha)+\alpha\varepsilon)^2} \geq 0
\end{aligned}$$

$$\frac{\partial A}{\partial \alpha} = f(p_0)\frac{\partial p_0}{\partial \alpha}(1 - p_0(1-\varepsilon) - (1-p_0)\varepsilon) + (p_1(1-\varepsilon) + (1-p_1)\varepsilon)f(p_1)\frac{\partial p_1}{\partial \alpha} > 0$$

because $1 - p_0(1-\varepsilon) - (1-p_0)\varepsilon = \frac{\varepsilon(1-\varepsilon)}{\varepsilon(1-\alpha)+(1-\varepsilon)\alpha} \geq 0$.
$\tilde{\varepsilon} > \varepsilon$: Define $B = \frac{r}{R}(F(p_m(\alpha r)) + \int_{p_m(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^{1} f(p)dp) - (\int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp + \int_{p_0(\alpha R)}^{1} f(p)dp)$.

$$\begin{aligned}
\frac{\partial p_0}{\partial r} =& \frac{(1-\varepsilon)\alpha\varepsilon R}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} \geq 0 \\
\frac{\partial p_0}{\partial R} =& -\frac{(1-\varepsilon)\alpha\varepsilon r}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} \leq 0 \\
\frac{\partial p_0}{\partial \alpha} =& \frac{(1-\varepsilon)R\varepsilon r}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} \geq 0 \\
\frac{\partial p_m}{\partial r} =& \frac{(1-\alpha+\alpha\varepsilon)(1-\varepsilon)R}{((1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r)^2} > 0 \\
\frac{\partial p_m}{\partial R} =& -\frac{(1-\alpha+\alpha\varepsilon)(1-\varepsilon)r}{((1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r)^2} < 0 \\
\frac{\partial p_m}{\partial \alpha} =& \frac{r(-(1-2\varepsilon)(R-r)-\varepsilon^2 R)}{((1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r)^2} < 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial B}{\partial r} =& \frac{1}{R}(F(p_m(\alpha r)) + \int_{p_m(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^{1} f(p)dp) \\
&+ \frac{r}{R}(f(p_m(\alpha r))\frac{\partial p_m(\alpha r)}{\partial r}(1-\mu(p_m(\alpha r))) - f(p_0(\alpha r))\frac{\partial p_0(\alpha r)}{\partial r}(1-\mu(p_0(\alpha r)))
\end{aligned}$$

The first four terms are positive. The last term is negative and converges to zero as $\varepsilon \to 0$. Define $C = F(p_m(\alpha r)) + \int_{p_m(\alpha r)}^{p_0(\alpha r)} \mu f(p)dp + \int_{p_0(\alpha r)}^{1} f(p)dp - \frac{R}{r}(\int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp +$

$\int_{p_0(\alpha R)}^1 f(p)dp)$.

$$\frac{\partial C}{\partial R} = f(p_m(\alpha r))\frac{\partial p_m(\alpha r)}{\partial R}(1-\mu(p_m(\alpha r))) - f(p_0(\alpha r))\frac{\partial p_0(\alpha r)}{\partial R}(1-\mu(p_0(\alpha r))) - \frac{1}{r}(\int_{p_1(\alpha R)}^{p_0(\alpha R)} \mu f(p)dp + \int_{p_0(\alpha R)}^{1} f(p)dp)$$

The first and last terms are negative. The second term is positive and converges to zero as $\varepsilon \to 0$.

$$\frac{\partial B}{\partial \alpha} = \frac{r}{R}(f(p_m(\alpha r))\frac{\partial p_m(\alpha r)}{\partial \alpha}(1-\mu(p_m(\alpha r))) - f(p_0(\alpha r))(1-\mu(p_0(\alpha r)))\frac{\partial p_0(\alpha r)}{\partial \alpha}) - (f(p_0(\alpha R))\frac{\partial p_0(\alpha R)}{\partial \alpha}(\mu(p_0(\alpha R))-1) - \mu(p_1(\alpha R))f(p_1(\alpha R))\frac{\partial p_1(\alpha R)}{\partial \alpha})$$

with $1-\mu(p_m(\alpha r)) = \frac{(1-2\varepsilon)(R-r)+\varepsilon^2 R}{(1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r} > 0$ and $1-\mu(p_0(\alpha r)) = \frac{\varepsilon R(1-\varepsilon)}{\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r} \geq 0$. The first two terms are negative. The last two terms are positive. These latter two terms converge to zero if ε goes to zero.

If α increases, more firms always pay if αr chosen, this makes the small bribe more attractive. There is no similar change in p_0 for the high bribe. □

A.5 Proof of Proposition 6

Proof. There are four possible combinations of bribe and project choice.

- B: αR, F: safe; $\pi(\alpha R, s) = 0$
- B: αr, F: safe; $\pi(\alpha r, s) = (1-\alpha)r$
- B: αR, F: risky; $\pi(\alpha R, r) = (p-\alpha)R$ if $p > p_0(\alpha R)$, $\pi(\alpha R, r) = p(1-\varepsilon)R - \alpha R((1-p)\varepsilon + p(1-\varepsilon))$ if $p_1(\alpha R) < p < p_0(\alpha R)$, $\pi(\alpha R, r) = 0$ if $p < p_1(\alpha R)$.
- B: αr, F: risky; $\pi(\alpha r, r) = pR - \alpha r$ if $p > p_0$, $\pi(\alpha r, r) = p(1-\varepsilon)R - \alpha r((1-p)\varepsilon + p(1-\varepsilon))$ if $p_1 < p < p_0$, $\pi(\alpha r, r) = 0$ if $p < p_1$ (these firms never pay the bribe).

The bureaucrat chooses the low bribe if

$$\beta_t \geq \frac{\mathbb{E}(\alpha R, r) - \mathbb{E}(\alpha r, r)}{\mathbb{E}(\alpha R, r) + \mathbb{E}(\alpha r, s) - \mathbb{E}(\alpha r, r)}$$

There exist parameter combinations, for which there does not exist an equilibrium in pure strategies. Proof by contradiction. Suppose an equilibrium in pure strategies exists. From the firm's behavior, the only problematic case is when $(\alpha R, s)$ is played in first period. First show that it cannot be an equilibrium that after $(\alpha R, s)$ is played in the first period, the bureaucrats play αR for n periods. Suppose this was the case. Then only firms that intend to play safe for at least n periods, choose safe in the first period and for all bureaucrats with $t \in [2, n]$, the expected probability that the firm chooses the risky project is zero, $\beta_t = 0$. Hence, $\mathbb{E}(\alpha R) = 0$ and already the bureaucrat in the second period would not follow the equilibrium strategy.

Second, show that it cannot be an equilibrium that after $(\alpha R, s)$, the bureaucrat chooses αr immediately for all parameter values. Suppose this was the case, then firms that intend to trick the bureaucrat play risky after the first period. The bureaucrat's belief that safe will be played in the next period if it has been played in this period is given by

$$p(safe|safe) = \frac{p(p < \bar{p}|safe)}{p(p < \bar{p}|safe) + p(p > \bar{p}|safe)}$$

where only firms with $p < \bar{p}$ choose to play safe the following period. If the group with $p > \bar{p}$ is large enough, $p(safe|safe)$ is so small that the bureaucrat would not choose the low bribe but the high bribe. An alternative way to guarantee equilibrium would be to impose restrictions on the parameter values to guarantee that the deviating group is not large enough to destroy this equilibrium.

In equilibrium, the firm chooses the safe project if $\mathbb{E}(safe) \geq \mathbb{E}(risky)$ which holds if $\pi(\alpha r, s)\gamma_0 - \gamma_0 \pi(\alpha r, r) - (1-\gamma_0)\pi(\alpha R, r) + \sum_{t=1}^{n-1}(\pi(\alpha r, s)\gamma_t - \pi(\alpha R, r))\delta^t + (\pi(\alpha r, s) - \pi(\alpha R, r))\sum_{t=n}^{\infty}\delta^t \geq 0$. As $\delta \to 1$, the comparison is made between $\pi(\alpha r, s)$ and $\pi(\alpha R, r)$.

After period n, bureaucrats play αr for sure. The expected profit of playing the safe and risky project, respectively, for firms of the deviating type are given by

$$\begin{aligned}
\mathbb{E}(dev, safe) =& \gamma_t \pi(\alpha r, s) + (1-\gamma_t)\pi(\alpha R, s) + \delta(\beta_{t+1} V(dev, s) + (1-\beta_{t+1})V(dev, r)) \\
\mathbb{E}(dev, risky) =& \gamma_t \pi(\alpha r, r) + (1-\gamma_t)\pi(\alpha R, r) + \delta V(dev, r)
\end{aligned}$$

with $V(dev, r) = \pi(\alpha R, r)\sum_{t=0}^{\infty}\delta^t$. The bureaucrats randomize such that in each period, $\mathbb{E}(dev, safe) = \mathbb{E}(dev, risky)$. Deviating firm-type randomizes such that $\beta = \frac{\mathbb{E}(\alpha R, r) - \mathbb{E}(\alpha r, r)}{\mathbb{E}(\alpha R, r) + \mathbb{E}(\alpha r, s) - \mathbb{E}(\alpha r, r)}$.

$\mathbb{E}(risky) \geq \mathbb{E}(dev, risky)$ if $\gamma_t \pi(\alpha r, r) + (1-\gamma)\pi(\alpha R, r) + \delta\pi(\alpha R, r)\sum_{t=0}^{\infty}\delta^t \geq \gamma_t \pi(\alpha r, r) + (1-\gamma)\pi(\alpha R, r) + \delta\pi(\alpha R, r)\sum_{t=0}^{\infty}\delta^t$ is identical. $\mathbb{E}(safe) \geq \mathbb{E}(dev, risky)$ can be reduced to $\mathbb{E}(safe) \geq \mathbb{E}(risky)$ and has been shown already.

□

A.6 Proof of Lemma 8

Proof. (i) If the bureaucrat chooses the low bribe once, he chooses the low bribe forever.

If $\varepsilon < \tilde{\varepsilon}$: assuming αr, firms choose risky if $p > p_m(\alpha r)$. Observing the real bribe αR, firms pay if $I_0 : p \geq p_0(\alpha R)$ and if $I_1 : p \geq p_1(\alpha R)$. If $p_m(\alpha r) < p_1(\alpha R)$, more firms choose risky but only the same firms pay the bribe αR as if αR was truthfully announced. If $p_1(\alpha R) < p_m(\alpha r)$, fewer firms choose risky and because of that, even at I_1, only firms with $p > p_m(\alpha r)$ can pay the bribe. Expected payoff is smaller.

If $\tilde{\varepsilon} < \varepsilon$: assuming αr, firms choose risky if $p > r/R$. Observing the real bribe, they pay according to $p_1(\alpha R)$ and $p_0(\alpha R)$. If $r/R < p_1(\alpha R)$, more firms choose risky but only firms with $p > p_1(\alpha R)$ pay the bribe. If $p_1(\alpha R) < r/R$, fewer firms choose risky and only firms with $p > r/R > p_1(\alpha R)$ can ever pay the bribe. This decreases the expected payoff.

(ii) If the bureaucrat chooses the high bribe once, he chooses the high bribe forever.

If $\varepsilon < \tilde{\varepsilon}$: firms choose risky if $p \geq 0$. They pay if $p > p_0(\alpha r)$ or $p > p_1(\alpha r)$. Remember that $p_1(\alpha R) > p_1(\alpha r)$. If firms assume that αr is the bribe demand, they choose risky if $p > p_m(\alpha r)$. Because $p_m(\alpha r) > 0$, more firms choose the risky project and may not be able to pay the bribe.

If $\tilde{\varepsilon} < \varepsilon$: firms choose risky if $p > 0$. Without deceit, they choose risky if $p > \frac{r}{R}$, all firms pay the bribe. Because $0 < \frac{r}{R}$, more firms choose risky and the outcome can be worse. □

A.7 Proof of Proposition 8

Proof. The statement has to be shown for both $\varepsilon > \tilde{\varepsilon}$ and $\varepsilon < \tilde{\varepsilon}$.

(i) $\varepsilon > \tilde{\varepsilon}$: The bureaucrat is more likely to choose αR for $f(p)$ if $\mathbb{E}_f(\alpha R) - \alpha r \geq \mathbb{E}_g(\alpha R) - \alpha r$ which holds if $\mathbb{E}_f(\alpha R) \geq \mathbb{E}_g(\alpha R)$. Rewrite $\mathbb{E}(\alpha R)$ as

$$\mathbb{E}(\alpha R) = \alpha R(1-(1-2\varepsilon)\int_{p_1}^{p_0} F(p)dp + F(p_0)(\varepsilon-1+p_0(1-2\varepsilon)) - F(p_1)(\varepsilon+p_1(1-2\varepsilon))$$

Then, $\mathbb{E}_f(\alpha R) \geq \mathbb{E}_g(\alpha R)$ if

$$(1-2\varepsilon)\int_{p_1}^{p_0}(G(p)-F(p))dp + (F(p_0)-G(p_0))(\varepsilon-1+p_0(1-2\varepsilon)) + (G(p_1)-F(p_1))(\varepsilon+p_1(1-2\varepsilon)) \geq 0$$

With $f(p)$ FOSD $g(p)$, this always holds because $\varepsilon - 1 + p_0(1-2\varepsilon) = \frac{-\varepsilon(1-\varepsilon)}{\varepsilon(1-\alpha)+(1-\varepsilon)\alpha} \leq 0$ and $\varepsilon + p_1(1-2\varepsilon) = \frac{\varepsilon(1-\varepsilon)}{(1-\varepsilon)(1-\alpha)+\alpha\varepsilon} \geq 0$

(ii) $\varepsilon < \tilde{\varepsilon}$: Need to show that $\mathbb{E}_f(\alpha R) - \mathbb{E}_f(\alpha r) \geq \mathbb{E}_g(\alpha R) - \mathbb{E}_g(\alpha r)$. Rewrite $\mathbb{E}_f(\alpha r) = \alpha r(1 - (1-2\varepsilon)\int_{p_m}^{p_0} F(p)dp + \varepsilon \int_{p_m}^{p_0} f(p)dp + F(p_0)(-1 + p_0(1-2\varepsilon)) + F(p_m)(1 - p_m(1-2\varepsilon)))$. Then,

$$\begin{aligned}\mathbb{E}_f(\alpha r) - \mathbb{E}_g(\alpha r) =& \alpha r((1-2\varepsilon)\int_{p_m}^{p_0} (G(p) - F(p))dp \\ &+ (F(p_0) - G(p_0))(-1 + p_0(1-2\varepsilon) + \varepsilon) \\ &+ (F(p_m) - G(p_m))(1 - p_m(1-2\varepsilon)))\end{aligned}$$

With $f(p)$ FOSD $g(p)$, the first two terms are positive, the third is negative. $1 - \varepsilon - p_m(1-2\varepsilon) = \frac{(1-\varepsilon)^2 R + r(-1+2\varepsilon)}{(1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r} > 0$ and $-1 + p_0(1-2\varepsilon) + \varepsilon = \frac{-\varepsilon(1-\varepsilon)R}{\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r} < 0$.

Then, $\mathbb{E}_f(\alpha R) - \mathbb{E}_f(\alpha r) \geq \mathbb{E}_g(\alpha R) - \mathbb{E}_g(\alpha r)$ is given by

$$\begin{aligned}&R((1-2\varepsilon)\int_{p_1(\alpha R)}^{p_0(\alpha R)} (G(p) - F(p))dp \\ &\quad + (F(p_0(\alpha R)) - G(p_0(\alpha R)))(\varepsilon - 1 + p_0(\alpha R)(1-2\varepsilon)) \\ &\quad + (G(p_1(\alpha R)) - F(p_1(\alpha R)))(\varepsilon + p_1(\alpha R)(1-2\varepsilon))) \\ &\quad - r((1-2\varepsilon)\int_{p_m}^{p_0(\alpha r)} (G(p) - F(p))dp + (F(p_m) - G(p_m))(1 - p_m(1-2\varepsilon))) \\ &\quad + (F(p_0(\alpha r)) - G(p_0(\alpha r)))(-1 + p_0(1-2\varepsilon) + \varepsilon) \geq 0\end{aligned}$$

This equation holds if $R\int_{p_1(\alpha R)}^{p_0(\alpha R)}(G(p) - F(p))dp \geq r\int_{p_m(\alpha r)}^{p_0(\alpha r)}(G(p) - F(p))dp$ and $R(F(p_0(\alpha R)) - G(p_0(\alpha R)))(\varepsilon - 1 + p_0(\alpha R)(1-2\varepsilon)) \geq r(F(p_0(\alpha r)) - G(p_0(\alpha r)))(-1 + p_0(\alpha r)(1-2\varepsilon) + \varepsilon)$. Both conditions are satisfied if ε is sufficiently small. □

A.8 Proof of Lemma 9

Proof. High bribe:

$$\begin{aligned}\frac{\partial y(\alpha R)}{\partial \varepsilon} =& Rf(p_0)p_0(\alpha R)(1-\varepsilon)\frac{\partial p_0(\alpha R)}{\partial \varepsilon} - \int_{p_1(\alpha R)}^{p_0(\alpha R)} pRf(p)dp \\ &- \frac{\partial p_1(\alpha R)}{\partial \varepsilon} Rf(p_1(\alpha R))p_1(\alpha R)(2-\varepsilon) < 0\end{aligned}$$

because $\frac{\partial p_0(\alpha R)}{\partial \varepsilon} < 0$ and $\frac{\partial p_1(\alpha R)}{\partial \varepsilon} > 0$.

Low bribe: Given $\varepsilon > \tilde{\varepsilon}$,

$$\frac{\partial y(\alpha r)}{\partial \varepsilon} = 0$$

Given $\tilde{\varepsilon} > \varepsilon$,

$$\frac{\partial y(\alpha r)}{\partial \varepsilon} = \frac{\partial p_m}{\partial \varepsilon} f(p_m)(r - p_m(1-\varepsilon)R) - \int_{p_m}^{p_0(\alpha r)} pRf(p)dp - \varepsilon \frac{\partial p_0(\alpha r)}{\partial \varepsilon} Rf(p_0(\alpha r))p_0(\alpha r)$$

The first term is positive, the second term is negative and the third term is positive. $\frac{\partial p_m}{\partial \varepsilon} = \frac{r(R-\alpha^2 r(1-2\varepsilon)}{((1-\varepsilon)(R-\alpha r)+\alpha \varepsilon r)^2} > 0$ and $\frac{\partial p_0}{\partial \varepsilon} = -\frac{\alpha r(R-\alpha r)}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} < 0$. □

A.9 Proof of Proposition 9

Proof. No corruption: Expected output is higher for $f(p)$ than for $g(p)$ because $y_f(0) - y_g(0) = r\int_0^{\frac{r}{R}}(f(p) - g(p))dp + R\int_{\frac{r}{R}}^1 p(f(p) - g(p))dp$ can be simplified to $R\int_0^1(G(p) - F(p)) \geq 0$ which holds by FOSD.

High bribe: $y_f(\alpha R)$ can be rewritten as $y_f(\alpha R) = R(1 - \int_{p_1(\alpha R)}^1 F(p)dp - p_1(\alpha R)F(p_1(\alpha R)) - \varepsilon(p_0(\alpha R)F(p_0(\alpha R)) - p_1(\alpha R)F(p_1(\alpha R)) - \int_{p_1(\alpha R)}^{p_0(\alpha R)} F(p)dp))$. Then, $y_f(\alpha R) \geq y_g(\alpha R)$ if

$$(1-\varepsilon)\int_{p_1(\alpha R)}^{p_0(\alpha R)}(G(p) - F(p))dp + \int_{p_0(\alpha R)}^1 (G(p) - F(p))dp$$
$$+ p_1(\alpha R)(G(p) - F(p))(1-\varepsilon) + \varepsilon p_0(\alpha R)(G(p_0(\alpha R)) - F(p_0(\alpha R))) \geq 0$$

which is always satisfied.

Low bribe: Expected output can be rewritten as $y(\alpha r) = R(1 - \int_{\frac{r}{R}}^1 F(p)dp)$. Then, $y_f(\alpha r) > y_g(\alpha r)$ if $R(1 - \int_{\frac{r}{R}}^1 F(p)dp) \geq R(1 - \int_{\frac{r}{R}}^1 G(p)dp)$. This can be simplified to $\int_{\frac{r}{R}}^1(G(p) - F(p))dp \geq 0$. □

A.10 Proof of Proposition 10

Proof. Because the low bribe is chosen for $g(p)$ if it is chosen for $f(p)$, the two conditions guaranteeing that the low bribe is chosen for both distributions and expected output is higher for $g(p)$ than for $f(p)$ are:

$$\alpha r \text{ for } f(p): F(p^*)r + r\int_{p^*}^1 pf(p)dp \geq R\int_0^1 pf(p)dp \tag{26}$$

$$y_g \geq y_f: \int_0^{p^*} rg(p)dp + \int_{p^*}^1 pRg(p)dp \geq \int_0^{p^*} rf(p)dp + \int_{p^*}^1 pRf(p)dp \tag{27}$$

Condition $f(p) \geq g(p)$ over a sufficiently large range for $p \in [p^*, \frac{r}{R}]$. Expected output can be rewritten as $\int_0^{p^*} rf(p)dp + \int_{p^*}^1 Rpf(p)dp = \int_0^{\frac{r}{R}} rf(p)dp + \int_{\frac{r}{R}}^1 Rpf(p)dp + \int_{p^*}^{\frac{r}{R}}(Rp - r)dp$. Then, $y_g \geq y_f$ can be written as $\int_0^{\frac{r}{R}} r(g(p) - f(p))dp + \int_{\frac{r}{R}}^1 Rp(g(p) -$

$f(p))dp + \int_{p^*}^{\frac{r}{R}}(Rp-r)(g(p)-f(p))dp \geq 0$. Rewriting gives $R(\int_{\frac{r}{R}}^{1}(F(p)-G(p))dp) \geq \int_{p^*}^{\frac{r}{R}}(r-Rp)(g(p)-f(p))dp$. Because the left-hand side is always negative, this inequality can only hold if $f(p) \geq g(p)$ over a sufficiently large range. Condition $\underline{\alpha} < \alpha < \min\{\bar{\alpha}, \bar{\alpha}'\}$. First, note that $\frac{\partial p^*}{\partial \alpha} < 0$. Then, it can be shown that equation (26) decreases in α: $\frac{\partial(26)}{\alpha} = \frac{\partial p^*}{\partial \alpha} f(p^*)r(1-p^*) < 0$. Equation (27) increases in α, $\frac{\partial(27)}{\partial \alpha} = \frac{\partial p^*}{\partial \alpha}(g(p^*)-f(p^*))(r-Rp^*)$, if $f(p^*) > g(p^*)$ and decreases in α if $g(p^*) > f(p^*)$. Because $\frac{\partial p^*}{\partial \alpha} < 0$, at low levels of α, $f(p^*) > g(p^*)$, while $g(p^*) > f(p^*)$ at high levels of α. Therefore, (27) first increases and then decreases in α, giving the two cutoff levels, $\underline{\alpha}$ and $\bar{\alpha}'$. Condition $\frac{r}{R}$ sufficiently large. Equation (26) increases in $\frac{r}{R}$: $\frac{\partial(26)}{\partial(\frac{r}{R})} = F(p^*) + \int_{p^*}^{1} pf(p)dp + p^* \frac{r}{R}\frac{\partial p^*}{\partial(\frac{r}{R})}(1-p^*) > 0$. □

A.11 Proof of Proposition 11

Proof. $\varepsilon > \tilde{\varepsilon}$.: $y_f(\alpha R) > y_g(\alpha r)$ if

$$A \equiv \int_{\frac{r}{R}}^{1} G(p)dp - p_1(\alpha R)F(p_1(\alpha R))(1-\varepsilon) - \varepsilon p_0(\alpha R)F(p_0(\alpha R))$$
$$- \int_{p_0(\alpha R)}^{1} F(p)dp - (1-\varepsilon)\int_{p_1(\alpha R)}^{p_0(\alpha R)} F(p)dp \geq 0$$

$$\frac{\partial A}{\partial R} = G(\omega)\frac{r}{R^2} > 0$$

because

$$\frac{\partial p_0(\alpha R)}{\partial R} = 0 \text{ and } \frac{\partial p_1(\alpha R)}{\partial R} = 0$$

$\tilde{\varepsilon} > \varepsilon$: $y_f(\alpha r)$ can be rewritten as $y_f(\alpha r) = rF(p_m) + R(1-\varepsilon)(p_0(\alpha r)F(p_0(\alpha r)) - p_m F(p_m) - \int_{p_m}^{p_0(\alpha r)} F(p)dp) + R(1 - p_0(\alpha r)F(p_0(\alpha r)) - \int_{p_0(\alpha r)}^{1} F(p)dp)$ if $\tilde{\varepsilon} > \varepsilon$. Then, $y_f(\alpha R) \geq y_g(\alpha r)$ if

$$R(-(1-\varepsilon)\int_{p_1(\alpha R)}^{p_0(\alpha R)} F(p)dp - \int_{p_0(\alpha R)}^{p_1(\alpha R)} F(p)dp - p_1(\alpha R)F(p_1(\alpha R))(1-\varepsilon)$$
$$- \varepsilon p_0(\alpha R)F(p_0(\alpha R)) + \varepsilon p_0(\alpha r)G(p_0(\alpha r))$$
$$+ \int_{p_m}^{1} G(p)dp - \varepsilon \int_{p_m}^{p_0(\alpha r)} G(p)dp) - G(p_m)(r - Rp_m(1-\varepsilon)) \geq 0$$

Obviously, $\frac{\partial p_0(\alpha R)}{\partial R} = 0$ and $\frac{\partial p_1(\alpha R)}{\partial R} = 0$.

$$\frac{\partial p_m}{\partial R} = -\frac{r(1-\alpha+\alpha\varepsilon)(1-\varepsilon)}{((1-\varepsilon)(R-\alpha r)+\alpha\varepsilon r)^2} < 0$$

$\frac{\partial p_0(\alpha r)}{\partial R} = -\frac{(1-\varepsilon)\alpha\varepsilon r}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} < 0$. Define $B = -(1-\varepsilon)\int_{p_1(\alpha R)}^{p_0(\alpha R)} F(p)dp - \int_{p_0(\alpha R)}^{p_1(\alpha R)} F(p)dp - p_1(\alpha R)F(p_1(\alpha R))(1-\varepsilon) - \varepsilon p_0(\alpha R)F(p_0(\alpha R)) + \varepsilon p_0(\alpha r)G(p_0(\alpha r)) + \int_{p_m}^{1} G(p)dp - \varepsilon\int_{p_m}^{p_0(\alpha r)} G(p)dp - G(p_m)(r/R - p_m(1-\varepsilon))$. Then,

$$\frac{\partial B}{\partial R} = \varepsilon p_0(\alpha r)g(p_0(\alpha r))\frac{\partial p_0(\alpha r)}{\partial R} - g(p_m)\frac{\partial p_m}{\partial R}(\frac{r}{R} - p_m(1-\varepsilon)) + G(p_m)\frac{r}{R^2}$$

Already shown that $p_m < r/R$ if $\varepsilon < \tilde{\varepsilon}$, and therefore $p_m(1-\varepsilon) < p_m < r/R$. The first term is negative, but the second two terms are positive. The negative converges to zero as $\varepsilon \to 0$.

$$\frac{\partial p_0(\alpha r)}{\partial \varepsilon} = -\frac{\alpha r}{(\varepsilon(R-\alpha r)+(1-\varepsilon)\alpha r)^2} < 0$$

□

A.12 Proof of Proposition 12

Proof. $b = \alpha R$: If $\varepsilon < \hat{\varepsilon}$, it follows that $p_1(\alpha R) < \frac{r}{R} < p_0(\alpha R)$. Then we can write $\sigma^2(\alpha R)$ as

$$\sigma^2(\alpha R) = R^2\left(\int_{p_1(\alpha R)}^{\frac{r}{R}} p(1-\varepsilon)(1-p(1-\varepsilon))f(p)dp + \int_{\frac{r}{R}}^{p_0(\alpha R)} p(1-\varepsilon)(1-p(1-\varepsilon))f(p)dp + \int_{p_0(\alpha R)}^{1} p(1-p)f(p)dp\right)$$

and excess volatility is given by

$$\sigma^2(\alpha R) - \sigma^2(0) = R^2\left(\int_{p_1(\alpha R)}^{p_0(\alpha R)} p\varepsilon(-1+2p-p\varepsilon)f(p)dp + \int_{p_1(\alpha R)}^{\frac{r}{R}} p(1-p)f(p)dp\right).$$

The first term can be positive or negative. The second term is positive. We can take the limit: $\lim_{\varepsilon\to 0}\int_{p_1}^{p_0} p\varepsilon(-1+2p-p\varepsilon)f(p)dp = 0$

If $\hat{\varepsilon} < \varepsilon$, $\frac{r}{R} < p_1(\alpha R) < p_0(\alpha R)$ such that $\sigma^2(\alpha R)$ can be rewritten as

$$\sigma^2(\alpha R) = R^2\left(\int_{p_1(\alpha R)}^{p_0(\alpha R)} p(1-\varepsilon)(1-p(1-\varepsilon))f(p)dp + \int_{p_0(\alpha R)}^{1} p(1-p)f(p)dp\right)$$

and excess volatility is given by

$$\sigma^2(\alpha R) - \sigma^2(0) = R^2\left(\int_{p_1(\alpha R)}^{p_0(\alpha R)} p\varepsilon(-1+2p-p\varepsilon)f(p)dp - \int_{\frac{r}{R}}^{p_1(\alpha R)} p(1-p)f(p)dp\right).$$

The first term can be positive or negative. The second term is negative. We can take the limit $lim_{\varepsilon \to \frac{1}{2}} \int_{p_1}^{p_0} p\varepsilon(-1 + 2p - p\varepsilon)f(p)dp = 0$ (because $p_1(\alpha R) = p_0(\alpha R)$). $\square$

B Extensions

B.1 Optimal bribe if the bribe demand influences the firm's project choice

If the return is not observable and the bribe influences the project choice, an alternative to the one presented in the text is that the bureaucrat maximizes his ex ante income. In this case, the expected values of the bribe payment and firm's return matter. The optimal bribe can be found as follows:

$$B = \mathbb{E}(b) - \pi \frac{\mathbb{E}(b)}{\mathbb{E}(return)} \mathbb{E}(b)$$

with $\mathbb{E}(b) = bn(b)$ where $n(b)$ is the probability that the bribe will be paid.

First, look at the high bribe $b > r$, then $\mathbb{E}(return) = R\bar{n}(b)$ where $\bar{n}(b)$ captures that all firms choose the risky project but not all firms' project succeed. Inserting gives the maximization problem as

$$B = \alpha R n(\alpha R) - \pi \alpha^2 R \frac{n(\alpha R)^2}{\bar{n}(\alpha R)}$$

which gives the first-order condition as

$$n(\alpha R) + \alpha n^{'}(\alpha R) = \pi \left(2\alpha \frac{n(\alpha R)^2}{\bar{n}(\alpha R)} + \alpha^2 \frac{2n(\alpha R)n^{'}(\alpha R)\bar{n}(\alpha R) - n(\alpha R)^2 \bar{n}^{'}(\alpha R)}{\bar{n}(\alpha R)^2} \right)$$

This condition defines the optimal value of α, but does usually not result in a closed-form solution. For $\varepsilon = 0$, $n(\alpha R) = \bar{n}(\alpha R)$ and $n^{'}(\alpha R) = \bar{n}^{'}(\alpha R) = 0$. Then, this equation can be reduced to

$$\alpha = \frac{1}{2\pi}$$

giving just the same result as in the case of observable profit. For $\varepsilon > 0$, however, this need not hold.

Second, look at the low bribe $b \leq r$. Then, $\mathbb{E}(return)$ depends on the proportions of firms choosing the risky project and paying the bribe.

$$B = \alpha r n(\alpha r) - \pi \frac{r^2 \alpha^2 n(\alpha r)^2}{\mathbb{E}(return)}$$

The first order condition is given by

$$n(\alpha r)+\alpha n^{'}(\alpha r) = \pi r \left(\frac{\mathbb{E}(return)(2\alpha n(\alpha r)^2 + \alpha^2 2n(\alpha r)n^{'}(\alpha r) - \alpha^2 n(\alpha r)^2\mathbb{E}(return)}{\mathbb{E}(return)^2} \right)$$

B.2 Reversal of expected output if ε small and αr chosen for both distributions

Proposition 14. *If αr is chosen for both distributions $f(p)$ and $g(p)$ and $\varepsilon < \tilde{\varepsilon}$, $y_g(\alpha r) > y_f(\alpha r)$ if $f(p) \geq g(p)$ over a sufficiently large range of $[p_m, p_0]$, $\underline{\alpha} < \alpha < \bar{\alpha}$ and ε sufficiently small.*

Proof. $y_f(\alpha r) \geq y_g(\alpha r)$ if

$$(G(p_m) - F(p_m))(R(1-\varepsilon)p_m - r) + R(1-\varepsilon)\int_{p_m}^{p_0}(G(p) - F(p))dp$$
$$+ R\int_{p_0}^{1}(G(p) - F(p))dp + Rp_0(G(p_0) - F(p_0))dp \geq 0$$

It can be shown that $R(1-\varepsilon)p_m - r < 0$ because this can be rewritten as $0 < (1-2\varepsilon)(R-r) + \varepsilon^2 R$. The first term is negative while all the other terms are positive. The relation can therefore be either way.

$$\frac{\partial p_m}{\partial \alpha} = \frac{r((R-r)(2\varepsilon - 1) - \varepsilon^2 R)}{((1-\varepsilon)(R-\alpha r) + \alpha\varepsilon r)^2} < 0$$

$$\frac{\partial p_0}{\partial \alpha} = \frac{(1-\varepsilon)r\varepsilon R}{(\varepsilon(R-\alpha r) + (1-\varepsilon)\alpha r)^2} > 0$$

We can now take the derivative of $y_g(\alpha r) - y_f(\alpha r)$ with respect to α.

$$\frac{\partial(y_g(\alpha r) - y_f(\alpha r))}{\partial \alpha} = \frac{\partial p_m}{\partial \alpha}(g(p_m) - f(p_m))(r - R(1-\varepsilon)p_m) - \varepsilon\frac{\partial p_0}{\partial \alpha}p_0(g(p_0) - f(p_0))R$$

This derivative is positive if $f(p_m) > g(p_m)$ and $f(p_0) > g(p_0)$. For small ε, the second term becomes negligible and only first is relevant for the sign.

Expected output can be rewritten: We know that $p_0 > p_m$ and $p_0 > r/R$. We have that $p_m > r/R$ if $\varepsilon > \frac{\alpha(R-r)}{R(1+\alpha)-2\alpha r} = \tilde{\varepsilon}$. And by assumption, $\varepsilon < \tilde{\varepsilon}$. Hence, $r/R > p_m$ and we have $p_m < r/R = \omega < p_0$. Then we can rewrite expected output as

$$y(\alpha r) = \int_0^\omega rf(p)dp + \int_{p_m}^\omega (p(1-\varepsilon)R - r)f(p)dp$$
$$+ \int_\omega^{p_0} p(1-\varepsilon)Rf(p)dp + \int_{p_0}^1 pRf(p)dp$$

Then, $y_g(\alpha r) - y_f(\alpha r) \geq 0$ if

$$-R\int_\omega^1 (G(p) - F(p))dp \geq \int_{p_m}^\omega (r - pR)(g(p) - f(p))dp + \varepsilon R \int_{p_m}^{p_0} p(g(p) - f(p))dp$$

LHS is negative, RHS is negative if $f(p) \geq g(p)$ over a sufficiently large range of $[p_m, \omega]$ and $[\omega, p_0]$. If α is small, p_m is high such that $f(p_m) > g(p_m)$ most likely, and p_0 low such that $g(p_0) > f(p_0)$ most likely. □

B.3 Proportion of firms behaving inefficiently for $\varepsilon = 0$

In the following, we measure the extent of the distortion as the proportion of firms, M, which choose the inefficiently risky project. When comparing the resulting distortion for $f(p)$ and $g(p)$ where $f(p)$ first-order stochastically dominates $g(p)$, we have to distinguish three different cases depending on whether the same, either high or low, or different bribe levels are chosen for the two distributions. Because of first-order stochastic dominance, the resulting corruption-induced distortion is more likely to be higher for $g(p)$ than for $f(p)$, especially if the same bribe level is chosen for both distributions.

The proportion of firms switching to the inefficiently risky technology depends on the bribe level. If the high bribe is chosen, all firms with success probability $p \in [0, \frac{r}{R}]$ switch, $M_f(\alpha R) = \int_0^{\frac{r}{R}} f(p)dp$. If the low bribe is chosen, only firms with success probability $p \in [\frac{(1-\alpha)r}{R-\alpha r}, \frac{r}{R}]$ switch, $M_f(\alpha r) = \int_{p^*}^{\frac{r}{R}} f(p)dp$ with $p^* = \frac{(1-\alpha)r}{R-\alpha r}$. If the high bribe is chosen for both distributions, more firms divert from their optimal behavior for the first-order stochastically dominated distribution. This follows directly from the assumed stochastic dominance. If the low bribe is chosen for either one or both distributions, additional conditions are needed to guarantee that the distortion is smaller for $f(p)$. If the low bribe is chosen for $g(p)$ and the high bribe is chosen for $f(p)$, the proportion of firms behaving inefficiently is larger for the stochastically dominated distribution if p^* is sufficiently small. The next proposition summarizes conditions that guarantee that the distortion is higher for $g(p)$ than for $f(p)$.

Proposition 15. *The proportion of firms behaving inefficiently under corruption is higher for $g(p)$ than for $f(p)$ $[M_g \geq M_f]$*

- *with αR for $f(p)$ and $g(p)$ if $f(p)$ FOSD $g(p)$,*
- *with αr for $f(p)$ and $g(p)$ if $g(p) \geq f(p)$ for $p \in [p^*, \frac{r}{R}]$,*
- *with αr for $g(p)$ and αR for $f(p)$ if $p^* < \bar{p}^*$ where $\bar{p}^*$ is defined by $G(\bar{p}^*) = G(\frac{r}{R}) - F(\frac{r}{R})$ and $f(p)$ FOSD $g(p)$.*

Proof. High bribe: It follows from FOSD that $M_g(\alpha R) = \int_0^{\frac{r}{R}} g(p)dp = G\left(\frac{r}{R}\right) \geq F\left(\frac{r}{R}\right) = \int_0^{\frac{r}{R}} f(p)dp = M_f(\alpha R)$. Low bribe: The distortion is larger for $g(p)$ than for $f(p)$ if $M_g(\alpha r) = \int_{p^*}^{\frac{r}{R}} g(p)dp \geq \int_{p^*}^{\frac{r}{R}} f(p)dp = M_f(\alpha r)$. Different bribes: The distortion is larger for $g(p)$ than for $f(p)$ if

$$M_g(\alpha r) = \int_{p^*}^{\frac{r}{R}} g(p)dp \geq \int_0^{\frac{r}{R}} f(p)dp = M_f(\alpha R). \tag{28}$$

Taking the derivative of equation (28) with respect to p^* gives $\frac{\partial}{\partial p^*} = g(p^*) > 0$. Because equation (28) becomes $\int_0^{\frac{r}{R}}(g(p) - f(p))dp = G(\frac{r}{R}) - F(\frac{r}{R}) > 0$ for $p^* = 0$, $-\int_0^{\frac{r}{R}} f(p)dp < 0$ for $p^* = \frac{r}{R}$ and decreases in p^*, we can conclude that $\exists \bar{p}^*$ such that $G(\bar{p}^*) = G(\frac{r}{R}) - F(\frac{r}{R})$. Hence, $\forall p^* < \bar{p}^*$, the distortion is larger under $g(p)$. □

In general, corruption results in fewer firms diverting from their optimal behavior for a first-order stochastically dominant distribution compared to the dominated one.

B.4 Additional volatility for $\varepsilon = 0$

In this section the corruption-induced distortion is measured in terms of the excess volatility resulting from corruption. Because corruption leads to firms choosing the riskier technology, output volatility increases.

In order to arrive at a measure of additional aggregate volatility, we aggregate the additional volatility at the firm level. Firms, which choose the safe investment, do not contribute to volatility. Firms, which choose the risky investment, have an expected return of pR. Hence, each firm choosing the risky investment generates a volatility of $\sigma^2_{firm} = p(R-pR)^2+(1-p)(0-pR)^2 = p(1-p)R^2$. Aggregating volatility over all firms gives aggregate volatility without corruption as $\sigma^2 = R^2 \int_{\frac{r}{R}}^1 p(1-p)f(p)dp$. Then the additional volatility resulting from the high bribe being chosen for distribution $f(p)$ is given by $\Delta_f(\alpha R) = R^2 \int_0^{\frac{r}{R}} p(1-p)f(p)dp$. If the bureaucrat sets the low bribe, additional volatility is given by $\Delta_f(\alpha r) = R^2 \int_{p^*}^{\frac{r}{R}} p(1-p)f(p)dp$. Denote the ratio of the safe to the risky return by $\omega = \frac{r}{R}$ and define $\bar{\omega}$ as the cutoff value where $\Delta_f(\alpha R) = \Delta_g(\alpha R)$, $R^2 \int_0^{\bar{\omega}} p(1-p)g(p)dp = R^2 \int_0^{\bar{\omega}} p(1-p)f(p)dp$.

If there is no corruption, more firms choose the risky technology for the first-order stochastically dominant distribution. Interestingly, the additional, corruption-induced, volatility can be smaller for the stochastically dominant distribution, especially if the high bribe is chosen for both distributions. If the low bribe is chosen for both distributions, the distortion is smaller if the mass of firms in the relevant interval is smaller.

Proposition 16. *The additional volatility resulting from corruption is higher for $g(p)$ than for $f(p)$ $[\Delta_g \geq \Delta_f]$*

- *with αR for $f(p)$ and $g(p)$*
  1. *if $g(p) \geq f(p)$ $\forall p \in [0, \frac{r}{R}]$ or*
  2. *if $f(p)$ FOSD $g(p)$ and $\bar{\omega} \geq 1$*
- *with αr for $f(p)$ and $g(p)$ if $g(p) \geq f(p)$ $\forall p \in [p^*, \frac{r}{R}]$*

Proof. High bribe: The distortion is larger for $g(p)$ than for $f(p)$ if

$$\Delta_g(\alpha R) = R^2 \int_0^{\frac{r}{R}} p(1-p)g(p)dp \geq \Delta_f(\alpha R) = R^2 \int_0^{\frac{r}{R}} p(1-p)f(p)dp. \tag{29}$$

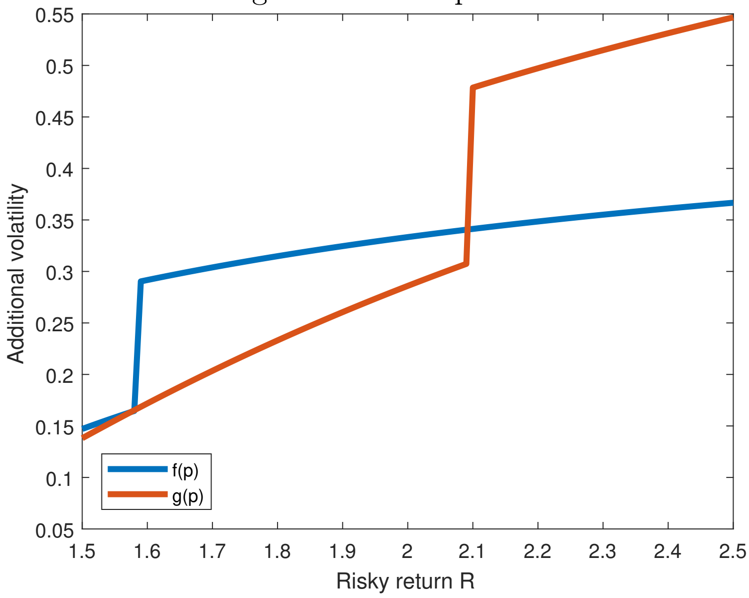
(1) obvious. (2) Taking the derivative of equation (29) w.r.t. ω, $\frac{\partial(29)}{\partial\omega} = \omega(1-\omega)(g(\omega) - f(\omega)) < 0$ if $f(\omega) > g(\omega)$. Hence, $\exists\omega$ s.t. $\Delta_g - \Delta_f < 0$ but $R > r$ by assumption. Low bribe: The distortion is larger for $g(p)$ than for $f(p)$ if $\Delta_g(\alpha r) = R^2 \int_{p^*}^{\frac{r}{R}} p(1-p)g(p)dp \geq \Delta_f(\alpha r) = R^2 \int_{p^*}^{\frac{r}{R}} p(1-p)f(p)dp$. □

The following example illustrates the above observations. As distributions, we use the triangular distribution with the mode at 0 for $g(p)$ and the uniform distribution for $f(p)$.

Example 3
Letting $\varepsilon = 0$, for $f(p) = 1$, we have $\Delta_f(\alpha R) = R^2 \int_0^{\frac{r}{R}} p(1-p)f(p)dp = r^2(\frac{1}{2} - \frac{1}{3}\frac{r}{R})$ and $\Delta_f(\alpha r) = R^2 \int_{p^*}^{\frac{r}{R}} p(1-p)f(p)dp = R^2(\frac{1}{2}((\frac{r}{R})^2 - (p^*)^2) - \frac{1}{3}((\frac{r}{R})^3 - (p^*)^3))$. For $g(p) = 2(1-p)$, we have $\Delta_g(\alpha R) = R^2 \int_0^{\frac{r}{R}} p(1-p)g(p)dp = 2r^2(\frac{1}{2} - \frac{2}{3}\frac{r}{R} + \frac{1}{4}(\frac{r}{R})^2)$ and $\Delta_g(\alpha r) = R^2 \int_{p^*}^{\frac{r}{R}} p(1-p)g(p)dp = 2R^2(\frac{1}{2}((\frac{r}{R})^2 - p^{*2}) - \frac{2}{3}((\frac{r}{R})^3 - p^{*3}) + \frac{1}{4}((\frac{r}{R})^4 - p^{*4}))$.

- $\Delta_f(\alpha r) > \Delta_g(\alpha r)$ because $g(p) \ngeq f(p)$ $\forall p \in [p^*, \frac{r}{R}]$
- $\Delta_g(\alpha R) > \Delta_f(\alpha R)$ because $\bar{\omega} = 1$

Figure 10: Example 3



Notes: $r = 1$, $\alpha = 0.67$

Figure 10 shows that there is a sudden rise in output volatility when the bureaucrat switches from the low to the high bribe demand.

Corruption increases volatility for both distributions. This increase tends to be larger for the first-order stochastically dominated distribution. An economy starting with a worse distribution of success probabilities, in terms of stochastic dominance, is therefore faced with larger distortions resulting from corruption.

Chapter 2

Corruption and Regulation: Choosing between Misallocation and Bribery

1 Introduction

Paying bribes in order to speed up administrative processes is a common action in many corrupt countries. While some authors argue that corruption serves a useful purpose in these cases, other authors have raised the concern that corrupt bureaucrats impose barriers and threaten to delay solely for the purpose of extracting higher bribes. The standard policy recommendation has therefore been that regulation be decreased in order to fight corruption (e.g. International Monetary Fund 2017). On the other hand, maintaining regulation could also contribute to a better allocation of goods. This therefore raises the question of what a benevolent government faced with corruption should do since a reduction in regulation will most likely reduce corruption yet increase misallocation. The present paper attempts to answer this question.

When an agent decides to apply for a driver's license or a business permit, this request is usually handled by a bureaucrat. The rules governing who should obtain a license or permit, however, are usually laid down by the government. In doing so, the government faces two problems: first, the bureaucrat administering the process is corrupt; second, agents value the scarce good to a different extent and might have insufficient income to pay for it. If the bureaucrat was not corrupt, any rule set by the government would be administered correctly. If all agents had the same valuation for the good, no rule to allocate the good would be required, and if all agents could pay an amount equal to their valuation, bribes would lead to the efficient allocation.

The government knows that the corrupt bureaucrat does not follow the official regulations to allocate the goods but rather designs and applies a new mechanism in order to maximize his income. This implies that the official rules, set by the government, are overwritten and not actually applied by the bureaucrat. There is, however, one caveat. The bureaucrat, when devising the income-maximizing mechanism to allocate the good, has to ensure that all agents prefer to participate in the mechanism and pay the bribe rather than follow the outside option of not paying the bribe.

If an agent decides not to pay a bribe to obtain the good more quickly, that agent can either follow the official regulation or decide not to try to get the good at all. If the agent follows the official regulation, then there is a certain probability that she will get the good. This refers to the idea of paying bribes to speed up administrative processes where the consideration is when and not whether the good is assigned. One reason for obtaining the good without paying the bribe could be that the permit is assigned to that applicant and cannot be given to someone else. Another reason for obtaining the good after the official waiting time could be that the applicant meets an honest bureaucrat when returning. In terms of the model, we will assume that some goods have not been distributed after every agent had the chance to bribe. These goods are then given to agents who followed the official regulation.

The bureaucrat has only limited influence on the agent's outside option of following the official regulation. This outside option determines how much an agent is willing to pay. The government can therefore impact whom the bureaucrat collects bribes from and allocates the good to by changing this outside option. Installing higher official regulation raises the cost of following the regulation because this becomes more time-consuming and therefore increases the bribes applicants are willing to pay. If the government wants to allocate goods in limited supply to agents who value them more, increasing the bribes they are willing to pay allows the bureaucrat to discriminate between agents with different valuations for the permit. Because higher bribes translate into a higher probability of obtaining the good, this contributes to a better allocation.

The model rests on a number of assumptions. There is a government that wants to distribute a good in limited supply and entrusts a corrupt bureaucrat with administering the application process.[1] Agents differ in their valuation of the good and, because of the scarcity of the good, the government wants to allocate the good to the applicants who value it the most. Agents can decide between paying the bribe or following the official regulation. If they pay a bribe, the intended regulation of the government is not followed. If they do not pay a bribe, they have to obey the entire official regulation. While the government decides on the rules and regulations for allocating goods, the corrupt bureaucracy may not apply them. Instead these official regulations enter the agents' consideration when evaluating whether to pay

[1] Assuming that the good is in limited supply implies that the good is not supposed to be given to every agent.

a bribe to speed up the process.

The bureaucrat cannot directly observe the applicants' valuation for the good. In order to maximize his income there are two strategies, he can follow. First, he can demand a low bribe which applicants with different valuations for the good can pay. Second, he can demand a bribe that is so high that only applicants who place a high value on the good pay. In terms of allocation, the government prefers the second approach. Applicants who value the good highly, however, only pay such a high bribe if they otherwise have to follow an overly cumbersome regulation.

The main results of the model are as follows: firstly, the optimal level of regulation differs depending on whether there is corruption or not. Secondly, because a corrupt bureaucracy does not rigorously apply official regulation, the government can choose a level of regulation that seems inefficiently high without having a negative impact on the allocation. Lastly, imposing a very high official level of regulation in the presence of a corrupt bureaucracy can improve the allocation of goods.

It is a common finding that regulation and corruption are positively correlated. Countries that have higher levels of corruption tend to be more strictly regulated and vice versa. Djankov et al. (2002) show for a sample of 85 countries that countries regulating entry more strictly, measured as the number of steps required to start a business, are more corrupt. Faría et al. (2013), using data from 169 countries and instrumental variables, find that a higher level of corruption leads to more business regulation that is also less transparent.

Figure 1 plots a measure of corruption, the Corruption Perceptions Index from Transparency International, against different measures of regulation, taken from the Doing Business project of the World Bank, for 2015. We see that a higher level of corruption is linked to both more days and a larger number of procedures that are needed to start a business. Also, the costs associated with starting a business are higher if there is more corruption. Lastly, a higher level of corruption is associated with a more difficult business climate.

Some authors study the quality of regulation if there is corruption. Méon and Sekkat (2005) find that corruption is even more harmful to growth and investments if governmental quality is low. Regulation, on the other hand, does not change the impact of corruption on investment. Breen and Gillanders (2012) use instrumental variables for corruption and the quality of institutions, and they find that institutions do not affect regulation once corruption is controlled for. They conclude that corruption determines regulatory quality.

Some, especially early, authors argue that corruption can be useful in overcoming regulation that is either overly complex or not very well suited for the economy. Leff

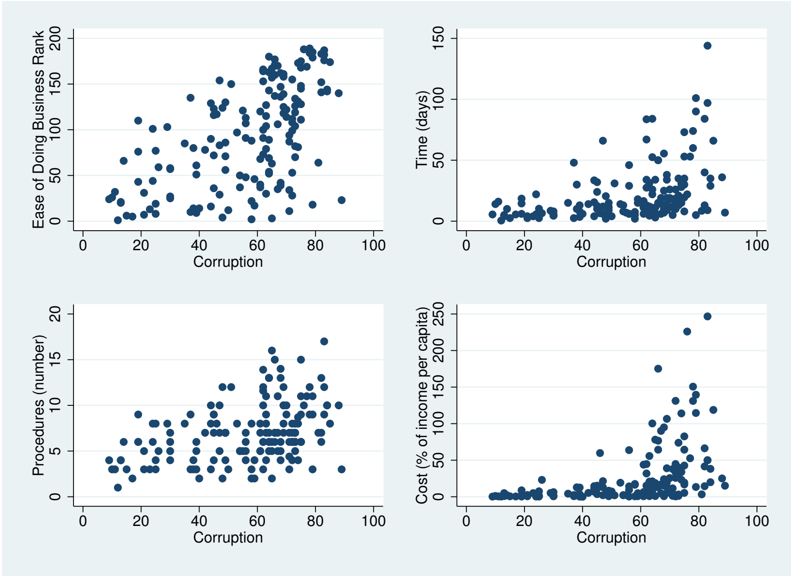


Figure 1: Corruption and Regulation

(1964) argues that if governments install wrong policies, graft helps to reduce these failures because giving favors to entrepreneurs paying the highest bribes is efficient. Also Huntington (1968) takes a positive stance on corruption if it helps to reduce the administrative burden. He claims that "the only thing worse than a society with a rigid, overcentralized, dishonest bureaucracy is one with a rigid, overcentralized, honest bureaucracy." While Nye (1967) also evaluates corruption positively if it is used to reduce red tape, he restricts this claim to being only true if the reduction in administrative procedures outweighs the cost of distorting rationally set selection criteria. Myrdal (1972) was among the first to point out that cumbersome administrative regulation might be introduced by corrupt officials in order to extract bribes. According to him, corruption does not increase administrative processing but rather decreases it by using delay as a threat. Lui (1985) investigates this claim in a queueing model where waiting time is determined by queue length and clients are placed in the queue according to their bribe payment. He shows that, contrary to Myrdal's hypothesis, allowing bribery can lead to the bureaucrat speeding up the process. Aidt (2003) points out that corrupt bureaucrats both create barriers and waste resources to conceal corrupt activities, but that corrupt politicians might also choose policies that more easily allow for corruption. Beck and Maher (1986), on the other hand, show that both bribery and bidding result in the same supplier being awarded a procurement contract and the government paying the same price, net of bribes.

There are several papers studying the relation between endogenously created red tape and regulatory outcomes. Banerjee (1997) and Banerjee et al. (2012)

are probably the most closely related papers to this one. Banerjee (1997) uses a principal-agent framework to show that a corrupt bureaucrat uses red tape in order to screen applicants and maximize his own income. The bureaucrat who can observe neither wealth nor type of the agents is punished by the government for allocating too many goods to the wrong type of agents. The poorer the agents and the scarcer the good, the more red tape the bureaucrat imposes. Banerjee et al. (2012) generalize this framework. They maintain the assumption that a corrupt bureaucrat has to allocate goods to agents who differ both in their types and wealth levels. While they allow the government to define a set of rules to be used for allocating the good, this is not treated as an outside option for the agents to follow. Instead they define corruption as the bureaucrat breaking these rules. This occurs when the rules chosen by the government do not guarantee him the highest surplus. Guriev (2004) assumes that a government wants to allocate a good such that negative externalities imposed by one type of agent are minimized. A corrupt bureaucrat whose income depends both on bribes and the allocation of goods can administer tests to reveal the agents' unobservable types. Guriev (2004) shows that the socially optimally amount of testing is implemented if bribes are used to speed up testing but not if they are used to hide bad test results. In all of these papers, the government, unable to control corruption, can perfectly observe the resulting allocation and alter the bureaucrat's payment accordingly. I do not assume that bureaucrats are paid according to performance.

There is a vast theoretical literature regarding corruption and regulation. The paper most closely related to this one is by Méndez and Sepúlveda (2013) who also endogenize the government's choice of regulation in the presence of corruption. They assume that a benevolent government uses regulation and red tape to limit negative externalities where entrepreneurs can circumvent regulation but not red tape. They show that the government increases red tape if negative externalities are high. This result, however, is driven by the assumption that entrepreneurs cannot circumvent red tape. This appears to be a very restrictive assumption as corrupt officials will try to maximize their income from bribery also by circumventing red tape even if this is more difficult. I do not make the assumption that the government has a type of regulation that can never be circumvented even if there is corruption. Instead, I allow the corrupt bureaucrat to maximize his income given the regulation put in place.

Other topics in this strand of the literature are the influence of bribery on waiting time and the choice of regulation. Bose (2004) shows that in order to decrease the threat of delay by public officials, punishments and rewards need to be based on the number of processed applications and not on the collected bribes. Assuming that bribes are used to decrease an exogenously given level of red tape, Saha (2011) shows that the rent from paying a bribe accrues to different types of agents depending on the level of red tape. Ahlin and Bose (2007) show that the possibility of extracting bribes creates an incentive for corrupt officials to delay. They also show that efficient producers are less willing to bribe if they can re-apply for a license and if there are many honest officials. This can result in a worse allocation where inefficient producers both get a license more often than and also ahead of their efficient counterparts. Kulshreshta (2007) shows that the effect of bribery on waiting time and allocative efficiency depends on whether the good the agents have to wait for, and the other good, which they consume, are substitutes or complements. Acemoglu and Verdier (2000) show how market and government failures interact. A government wants to limit pollution but corrupt bureaucrats inspect and fine. They show that the government might choose to reduce the fine and hire more bureaucrats to raise inspection rates if monitoring bureaucrats becomes more difficult. Esteban and Ray (2006) show that misallocation of goods can also result if a government lacks information about the productivity and wealth of agents. If the government can only use a price mechanism to allocate goods, agents that are either very productive or very rich pay the price such that the government allocates the good also to rich but unproductive agents. Aghion et al. (2010) study the interconnection between regulation, corruption and trust. Agents can decide whether to become civic and whether to become an entrepreneur. An uncivic entrepreneur imposes a negative externality and an uncivic bureaucrat is corrupt. There are two equilibria: one with many civic individuals and low levels of regulation and corruption and one with many uncivic individuals and high levels of regulation and corruption.

2 The model

In the economy there are three types of agents: the government, a corrupt bureaucrat and a large number of heterogenous agents. The government wants to allocate N goods to the agents but can only do so with the bureaucrat's help. Agents differ both in their valuation for the good and in their income. The agents' valuation can be either high, H, or low, L. Agents are either rich and have an income y_r or poor and have an income y_p.[2] The number of agents with high valuation is denoted by n_H, comprised of rich high-valuation agents, n_{Hr}, and poor high-valuation agents, n_{Hp}.

[2]Because of the potential existence of budget-constrained agents selling the good at the market price is usually not optimal.

There are only rich low-valuation agents denoted by n_L. Agents apply to get one of the goods and can use their income to bribe the bureaucrat. An agent's valuation and income are private information. The number of goods, the government wants to allocate, equals the number of high-valuation agents and we assume that there are more goods than rich agents, $n_L + n_H > N = n_H > n_L + n_{Hr}$.[3]

The bureaucrat derives utility from increasing his income by collecting bribes but incurs a small cost $\tau > 0$ if he imposes waiting time on an agent and an even smaller cost $\varepsilon > 0$ if he does not allocate a good optimally. The bureaucrat chooses an optimal mechanism in order to maximize his revenue from bribery by upholding a positive amount of waiting time on some agents, by choosing the probability, an agent gets a good, and by setting the bribe to be paid. For example, he intends the rich high-valuation agents to choose bribe b_{Hr} and waiting time T_{Hr} in return for getting the good with probability π_{Hr}.

The government maximizes social welfare which depends on the allocation of goods, the waiting time endured by the agents and the government's taste for bribes being paid by the agents to the bureaucrat. The government can impose official waiting time T as only means to control whom the goods are allocated to.

Agents receive a utility equal to their valuation if they get a good. They incur a cost of $\delta \geq 0$ for each unit of time they have to wait. Agents can bribe the bureaucrat, follow the official waiting time or do nothing. Whenever an agent is indifferent between two actions, she does what the government or the bureaucrat prefers her to do.[4] If an agent decides to bribe, that agent always has to pay the bribe and wait for the time imposed by the bureaucrat no matter whether the good is assigned at the end of the waiting time.[5] If agents decide not to bribe, they have to obey the official waiting time to get the good. While the agent waits, the bureaucrat continues to distribute goods in exchange for bribes. After all agents had the opportunity to bribe the bureaucrat, there is a certain probability that a good is still left. The agent who decided not to bribe gets the good with this probability.[6] Because the records show that the applicant waited for a certain time, she gets the

[3]It would be possible to extent this to assuming that there are more goods than agents with a high valuation, $N > n_H$. This would, however, require a more sophisticated governmental policy in the absence of corruption, but presumably not reveal more insights into the underlying mechanism.

[4]As an example consider the following: If the government imposes regulation $\delta T = L$ the low-valuation agents do not wait. But if the bureaucrat asks for a bribe that extracts all surplus from these agents, they pay the bribe.

[5]This can be changed to the case where the bribe is only paid and the agents only wait if they get the good.

[6]An alternative assumption would be that agents get the good for sure after waiting the official time. This increases the agent's outside option. Assuming that the bureaucrat can choose between these two procedures, he chooses the one which results in a lower outside option for the agent.

good if there is one left. If agents decide to bribe, but are not assigned a good through bribing, they never get the good. The timeline of the game is summarized below

1. the government sets the official regulation T
2. agents meet the bureaucrat
3. the bureaucrat offers each agent a menu of options consisting of waiting time T_i, bribe b_i and probability to obtain the good π_i with $i = Hr, Hp, L$
4. agents decide whether to bribe and which offer to take
   (a) if the agent bribes, she receives the good according to the offer
   (b) if the agent does not bribe, she is registered for the official waiting time
5. at the end of the official waiting time, the agent gets the good if a good is still available

2.1 The bureaucrat's problem

In this section we compute the mechanism chosen by the bureaucrat for a given level of official regulation. The mechanism, the bureaucrat optimally designs, varies with the official regulation chosen by the government. Therefore, after specifying the bureaucrat's maximization problem in its general form, we first identify the parts of the bureaucrat's mechanism that are independent of governmental regulation. We then solve the bureaucrat's problem for different levels of the official regulation.

The bureaucrat maximizes his utility from bribery subject to the individual rationality and incentive constraints of the high- and low-valuation agents. If an agent bribes, she pays the bribe irrespective of whether she receives the good. Bribery income therefore depends on the number of agents of each type and the type-specific bribe paid by them. The bureaucrat incurs a cost $\varepsilon > 0$ for each good allocated to a low-valuation agent instead of a high-valuation agent. The best allocation of goods to agents results if all goods are given to the high-valuation agents. Therefore, the total cost resulting from misallocation for the bureaucrat is the number of low-valuation agents getting the good weighted with the associated cost parameter. He also incurs a cost from imposing waiting time $\tau > 0$.[7] The total cost associated with waiting time is the entire waiting time imposed on the agents times the cost

[7]These assumptions, $\varepsilon > 0$ and $\tau > 0$, guarantee that the equilibrium is unique because the bureaucrat always chooses the smallest amount of waiting time satisfying the agents' constraints and is not indifferent between several allocations because he does not like misallocation.

parameter τ. For each agent i the bureaucrat chooses π_i, the probability the agent gets the good, b_i, the bribe the agent has to pay, and T_i, the waiting time for that agent. He also has to ensure that the agent prefers to bribe instead of following the official regulation or not attempting to get the good. In doing so, the bureaucrat has to consider the individual rationality constraints of the agents. These depend on the agent's total cost of following the official regulation, δT, and the probability that agent i gets the good by obeying the official waiting time, p_i. The agent can also decide to neither bribe nor wait and do nothing instead. The agent's outside option is therefore given by $max\{0, p_i V_i - \delta T\}$ with $V = L, H$. The bureaucrat solves the following problem:

$$\max_{\substack{\{b_{Hr}, b_{Hp}, b_L, \\ \pi_{Hr}, \pi_{Hp}, \pi_L, \\ T_{Hr}, T_{Hp}, T_L\}}} B = n_{Hr} b_{Hr} + n_{Hp} b_{Hp} + n_L b_L - \varepsilon \pi_L n_L - \tau(T_{Hp} n_{Hp} + T_{Hr} n_{Hr} + T_L n_L)$$

$$\begin{aligned}
& s.t. \\
(IR_{Hr})\ & \pi_{Hr} H - b_{Hr} - \delta T_{Hr} \geq max\{0, p_{Hr} H - \delta T\} \\
(IR_{Hp})\ & \pi_{Hp} H - b_{Hp} - \delta T_{Hp} \geq max\{0, p_{Hp} H - \delta T\} \\
(IR_L)\ & \pi_L L - b_L - \delta T_L \geq max\{0, p_L L - \delta T\} \\
(Hr \geq Hp)\ & \pi_{Hr} H - b_{Hr} - \delta T_{Hr} \geq \pi_{Hp} H - b_{Hp} - \delta T_{Hp} \\
(Hp \geq Hr)\ & \pi_{Hp} H - b_{Hp} - \delta T_{Hp} \geq \pi_{Hr} H - b_{Hr} - \delta T_{Hr} \\
(Hr \geq L)\ & \pi_{Hr} H - b_{Hr} - \delta T_{Hr} \geq \pi_L H - b_L - \delta T_L \\
(L \geq Hr)\ & \pi_L L - b_L - \delta T_L \geq \pi_{Hr} L - b_{Hr} - \delta T_{Hr} \\
(L \geq Hp)\ & \pi_L L - b_L - \delta T_L \geq \pi_{Hp} L - b_{Hp} - \delta T_{Hp} \\
(Hp \geq L)\ & \pi_{Hp} H - b_{Hp} - \delta T_{Hp} \geq \pi_L H - b_L - \delta T_L \\
& 1 \geq \pi_{Hr}, \pi_{Hp}, \pi_L \geq 0 \\
& T_{Hr}, T_{Hp}, T_L \geq 0 \\
& y_r \geq b_{Hr}, b_L \\
& y_p \geq b_{Hp}
\end{aligned} \tag{1}$$

The first three conditions guarantee that all agents want to bribe instead of either following the official regulation or not doing anything. The following six conditions guarantee that every agent prefers the option designed for her instead of mimicking another type. Lastly, it is required that agents have a sufficiently high income to pay their bribe, that they cannot wait a negative amount of time and that they cannot get the good with a negative probability.

We first observe that the probability an agent gets the good when obeying the official waiting time is the same as the probability that this agent gets a good when bribing. This follows from the bureaucrat offering each agent a menu of options including the probability of getting the good. If the agent decides not to bribe, then the good, she should have gotten, is left after every agent had the possibility to bribe. We make the following observations about the incentive constraints.

Remark 1. *($Hp \geq Hr$) can be dropped. The following incentive constraints cannot bind simultaneously: ($Hr \geq L$) and ($L \geq Hr$); ($L \geq Hp$) and ($Hp \geq L$).*

Proof. 1. ($Hp \geq Hr$) is either irrelevant or implied by ($Hr \geq Hp$): To see this note that $b_{Hr} \geq b_{Hp}$ because $y_r \geq b_{Hr} \geq y_p \geq b_{Hp}$. For $b_{Hr} > y_p$, ($Hp \geq Hr$) is irrelevant because Hp cannot mimic Hr; for $y_p \geq b_{Hr}$, $b_{Hr} = b_{Hp}$ such that $\pi_{Hr}(H - b_{Hr} - \delta T_{Hr}) = \pi_{Hp}(H - b_{Hp} - \delta T_{Hp})$ and both ($Hp \geq Hr$) and ($Hr \geq Hp$) bind. Because the objective function decreases in T_{Hp} and T_{Hr}, differentiating between the high types does not make sense if they can pay the same bribe. 2. ($Hr \geq L$) and ($L \geq Hr$) cannot bind simultaneously. Suppose they did. Then,

$$(\pi_L - \pi_{Hr})L = b_L + \delta T_L - b_{Hr} - \delta T_{Hp} = (\pi_L - \pi_{Hr})H$$

A contradiction. The same argument can be made for ($Hp \geq L$) and ($L \geq Hp$). □

There are six different combinations of wealth and valuation. Three of these six possible combinations of wealth and valuation result in the efficient outcome if the goods are sold at a sufficiently high price.[8] In all of these the poor high-valuation agents have an income that is higher than the valuation of the low-valuation agents and are therefore able to pay more than these agents. In the other three cases, however, budget constraints for the poor high-valuation agents bind such that the allocation can be improved by installing official waiting time.[9] In these cases the rich low-valuation agents can pay higher bribes than the poor high-valuation agents. In the following analysis we focus on the latter three combinations of wealth and valuation.

The bureaucrat tries to extract bribes that are as high as possible. In order to prevent agents who have to pay very high bribes from mimicking another type of agent, he can impose waiting time on agents paying smaller bribes. In some cases, the amount of waiting time, the bureaucrat chooses, is zero. For each type of agent the bureaucrat wants to lower the waiting time and increase the bribe whenever this is possible as this increases profits from this type. On the other hand, the bureaucrat also needs to construct an incentive-compatible mechanism. If the poor

[8]These are $L < y_p < y_r < H$, $L < y_p < H < y_r$ and $L < H < y_p < y_r$.
[9]These are $y_p < L < y_r < H$, $y_p < L < H < y_r$ and $y_p < y_r < L < H$.

agents are very poor, they can only pay a very small bribe. In order to prevent the rich agents from mimicking the poor agents, the bureaucrat introduces waiting time for the poor agents.

The bureaucrat never wants to impose waiting time on the rich low-valuation agents. He has to consider two cases. First, if no other type of agent wants to mimic the rich low-valuation agents, imposing additional waiting time is not necessary. Because waiting time might reduce the bribe, the bureaucrat reduces waiting time in this case. Second, if another type of agent wants to mimic the rich high-valuation agents, optimal waiting time imposed on these agents is still zero. This follows from the bureaucrat incurring a cost from giving the good to the low-valuation agents and incurring a cost from waiting time imposed on them. Thus, he can reduce both waiting time and the probability that the rich low-valuation agents get the good in order to increase his utility. When considering waiting time imposed on the rich high-valuation agents, there can again be the case that an increase in waiting time only reduces the bribe payment such that the bureaucrat reduces waiting time for these agents. If, however, the rich agents are relatively poor, meaning that their income is lower than the low valuation for the good, a positive amount of waiting time imposed on the high-valuation agents could help the bureaucrat to distinguish the rich high- and low-valuation agents. Whenever the bureaucrat dislikes imposing waiting time relatively more than allocating the good to the low-valuation agents, waiting time for the rich high-valuation agent is zero. The corresponding condition is given in the following assumption.

Assumption 1 (Waiting is worse than misallocation for the bureaucrat)**.**

$$\tau(n_L + n_{Hr})L > \varepsilon n_L \delta$$

The assumption guarantees that the bureaucrat prefers to give the good to the rich low-valuation agents instead of imposing a sufficient amount of waiting time to deter these agents from mimicking the rich high-valuation agents. If there is no need for the bureaucrat to distinguish between the rich low- and high-valuation agents in order to prevent misallocation, he does not uphold waiting time for the rich high-valuation agents. This implies that there is no waiting time for the rich agents.

Remark 2. *$T_L = 0$ and $T_{Hr} = 0$ if assumption 1.*

Proof. See Appendix. □

Another way to see that the rich low-valuation agents do not have to wait is to consider under which conditions the high-valuation agents would want to mimic them. The rich high-valuation agents would only want to pose as rich low-valuation agents in a partial pooling equilibrium with pooling between the rich agents.[10] In this case, however, there is no need for waiting time to be positive for the rich agents. The poor high-valuation agents usually cannot mimic the low-valuation agents because of their insufficiently small income.

We continue by investigating the probability that the rich high-valuation agents obtain the good. If both regulation and the income of the rich agents are higher than the low valuation, the rich high-valuation agents both want to and can pay the most for the good. This also implies that their probability of obtaining the good is the highest or among the highest. The bureaucrat decides whether a partial pooling equilibrium of the rich agents or a separating equilibrium results. In a separating equilibrium, the rich high-valuation agents pay such a high bribe that an increase in the probability that they get the good can only raise the bribe they pay. In a pooling equilibrium between the rich high- and low-valuation agents, both types of agents obtain the good with the same probability. Increasing the probability the rich agents get the good therefore raises the costs from misallocation for the bureaucrat. In order to determine whether the bureaucrat should raise or decrease the probability that the rich agents obtain the good, two effects are important. First, increasing this probability raises misallocation. Second, increasing this probability can potentially lower the amount of waiting time imposed on the poor agents. Which effect dominates depends on the bribe which the poor agents pay.

There are two possibilities for the bribe level of the poor agents. If the poor agents are relatively rich, their bribe will be limited by the probability that they are assigned a good. If, on the other hand, they are relatively poor, their bribe will be limited by their income. In this case, an increase in the probability that they get the good is accompanied by an increase in their waiting time. The bureaucrat seeks to reduce the probability they get the good if he dislikes waiting more than misallocation. The next assumption guarantees that poor agents are so poor that an increase in their probability of getting the good never increases the bribe they pay.[11] Moreover, this assumption guarantees that waiting time imposed on poor

[10]In a separating equilibrium the rich high-valuation agents pay higher bribes resulting in a higher probability of getting the good. This is preferred because of their higher valuation for the good.

[11]In order to guarantee this, $y_p < \pi_{Hp} H$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$ suffices. Suppose this was not satisfied such that $y_p > \pi_{Hp} H$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$. Then, increasing π_{Hp} from this minimum value increases b_{Hp} and reduces misallocation. In this case $\pi_{Hr} = \pi_L < 1$ would be possible in a pooling equilibrium also if $min\{y_r, \delta T\} \geq L$.

agents is always positive. Because the bureaucrat incurs a higher cost from waiting than from misallocation, he increases the probability that the rich agents get the good in order to reduce the waiting time imposed on the poor agents.

Assumption 2 (Poor agents are very poor)**.**

$$\pi_{Hp} L \geq y_p \text{ with } \pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$$

The next assumption is a slight variation of assumption 1 and guarantees that the bureaucrat dislikes waiting more than misallocation.

Assumption 3 (Waiting worse than misallocation for the bureaucrat)**.**

$$\tau(n_H + n_L)L > \varepsilon n_L$$

The following, obviously not very restrictive, assumption guarantees that the bureaucrat increases the bribe paid by the low-valuation agents by raising the probability that the low-valuation agents get the good even if this also increases misallocation.

Assumption 4 (Small cost of misallocation)**.**

$$n_L > \varepsilon$$

With these assumptions in place, the rich high-valuation agents obtain the good for sure if the rich low-valuation agents do not want to mimic them.

Remark 3. $\pi_{Hr} = 1$ *if* $min\{y_r, \delta T\} \geq L$ *and assumptions 2, 3 and 4.*

Proof. See Appendix. □

The assumptions thus ensure that the bureaucrat is as corrupt as possible. He first maximizes his income from bribery. If several mechanisms result in the same income, he chooses the one with the lowest amount of waiting time. Last, if several options result in the same income and the same amount of waiting time, he chooses the one that leads to the lowest level of misallocation.

We are now in a position to derive the mechanism chosen by the bureaucrat for different levels of the official regulation. The next assumption guarantees that the cost of an increase in waiting time is so small that this cost never changes the bureaucrat's optimal behavior determined by income maximization.

Assumption 5 (Small waiting cost)**.**

$$\left| -\frac{n_{Hr}}{n_L}\frac{H-L}{H} + \frac{L-\varepsilon}{H} \right| > \frac{\tau}{\delta}$$

The waiting time of the poor agents depends on which type of rich agent wants to deviate and pretend to be a poor agent. This, in turn, depends on the level of income of the rich agents. For high levels of the high income, the rich high-valuation agents want to mimic the poor agents. In this case changes in the probability that the low-valuation agents obtain the good, π_L, induce a large change in waiting time T_{Hp}. For low levels of the high income, the low-valuation agents want to pretend to be a poor agent. In this case a change in probability π_L induces a small change in waiting time T_{Hp}. Because the bureaucrat does not like an increase in T_{Hp} but the magnitude depends on the high income y_r, it would be possible that the chosen equilibrium depends on income y_r conditional on $y_r > L$. This assumption excludes the case that the bureaucrat's behavior changes for small changes in the high income y_r. If this assumption was not satisfied, for $y_r \in (L, H)$ and under some very rare parameter constellations, the bureaucrat would increase π_L for high levels of y_r and decrease π_L for low levels of y_r. A similar reasoning applies for $y_r > H$.[12] Because the bureaucrat's costs of waiting are small, this only occurs for very special parameter conditions and is therefore excluded in order to reduce the different number of cases to be investigated.[13]

From here onwards, we impose assumptions 1 to 5. We can now compute the solution to the bureaucrat's problem if regulation is high, $\delta T \in [L, H]$, and without the participation constraints of the high-valuation agents, (IR_{Hp}) and (IR_{Hr}). If the official regulation is high only high-valuation agents wait. Low-valuation agents either get the good by bribing or not at all.[14]

Proposition 1. *For $\delta T \in [L, H]$ the solution to problem (1) without (IR_{Hp}) and (IR_{Hr}) is given by*

1. *Case 1: $y_p < L < y_r < H$:*
   (a) *if $-n_{Hr}(H-L)+n_L L-\varepsilon n_L < 0$: $b_{Hr} = y_r$, $b_L = \pi_L L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = \frac{H-y_r}{H-L}$, $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, $\delta T_{Hp} = max\{\pi_{Hp}H - y_p - H + y_r, \pi_{Hp}L - y_p\}$*

[12]This excludes that π_L has an interior solution for very special parameter constellations, $\tau\frac{H}{\delta}n_L > (H-L)n_{Hr} - n_L L + \varepsilon n_L > \tau\frac{L}{\delta}n_L$.

[13]Under this assumption it is therefore not necessary to include the τ term as condition in the proposition.

[14]For official regulation, high-valuation agents wait if they are indifferent while low-valuation agents do not.

(b) *else:* $b_{Hr} = L$, $b_L = L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = 1$, $\pi_{Hp} = \frac{N - n_L - n_{Hr}}{n_{Hp}}$, $\delta T_{Hp} = max\{\pi_{Hp} H - y_p - H + L, \pi_{Hp} L - y_p\}$

2. *Case 2:* $y_p < L < H < y_r$:

   (a) *if* $-n_{Hr}(H - L) + n_L L - \varepsilon n_L > 0$; $b_{Hr} = L$, $b_L = L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = 1$, $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$, $\delta T_{Hp} = \pi_{Hp} L - y_p$

   (b) *else;* $b_{Hr} = H$, $b_L = 0$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = 0$, $\pi_{Hp} = 1$, $\delta T_{Hp} = H - y_p$

3. *Case 3:* $y_p < y_r < L < H$:

   (a) *if* $y_r > y_p + L(1 - \pi_{Hp})$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$: $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = 1$, $\pi_{Hp} = \frac{N - n_L - n_{Hr}}{n_{Hp}}$, $\delta T_{Hp} = y_r - y_p - L(1 - \pi_{Hp})$

   (b) *else:* $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_L - n_{Hr})}{n_{Hp}}$, $\delta T_{Hp} = 0$

Proof. See Appendix. □

The bureaucrat uses waiting time for the poor agents and different bribe demands to discriminate between the different types of agents and to maximize his income. Whenever the rich high- and low-valuation agents are relatively poor, meaning that their income is lower than the valuation of the low type, $y_r < L$, we arrive at a partial pooling equilibrium where the rich agents pay the same bribe and get the good with the same probability. This equilibrium results because the rich high-valuation agents are not able to pay more than the rich low-valuation agents and because the bureaucrat only incurs a very small cost of misallocation. In order to determine the exact probability that the rich agents get the good, the bureaucrat follows two considerations. First, he wants to reduce the waiting time of the poor agents because he incurs a cost from waiting. Second, he wants to limit the extent of misallocation because misallocation is also costly for him. Because he incurs a higher cost from waiting than from misallocation, he raises the probability that the rich agents get the good until either waiting cost equals zero or the probability that the rich agents get the good equals one. Which of these two cases arises depends on how rich the rich agents are relative to the poor agents. If the rich agents are very rich, they have to pay a high bribe resulting in a positive amount of waiting time to deter them from mimicking the poor agents. If they are only slightly richer than the poor agents, the waiting time needed to deter the rich agents from mimicking the poor agents can be negative. In this case the bureaucrat increases the waiting time to zero by lowering the probability the rich agents get the good thereby reducing the cost, he incurs from giving the good to the rich low-valuation agents.

Whenever the rich agents have a high income, i.e. an income higher than the low valuation, $y_r > L$, the rich high-valuation agents are both willing and able to pay more than the rich low-valuation agents. The bureaucrat can therefore decide to create a fully separating equilibrium or an equilibrium in which he pools the rich agents. Which of the two options he chooses depends on the parameter values. Whenever the rich high-valuation agents are both many and have a much higher valuation than the rich low-valuation agents, the bureaucrat chooses a fully separating equilibrium and demands the highest bribe, the rich high-valuation agents can and want to pay. Because the rich high-valuation agents pay a very high bribe in these cases, the bureaucrat has to ensure that they do not mimic the rich low-valuation agents. This results in a low probability that the low-valuation agents get the good. This in turn implies that the poor agents' probability of getting the good increases leading to a more efficient allocation of goods.

The next step is to include (IR_{Hr}) and (IR_{Hp}) in the solution derived before without these constraints. We then check whether the participation constraints of the high-valuation agents are satisfied. If this is the case, the solution remains the same. If this is not the case, the bureaucrat derives a new optimal mechanism.

Proposition 2. *For $\delta T \in [L, H]$ the solution to problem (1) with (IR_{Hp}) and (IR_{Hr}) is given by*

1. *Case 1: $y_p < L < y_r < H$:*
   (a) *if $-n_{Hr}(H - L) + n_L L - \varepsilon n_L < 0$:*
      i. *if $\delta T \geq y_r$: $b_{Hr} = y_r$, $b_L = L\pi_L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = \frac{H-y_r}{H-L}$, $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, $\delta T_{Hp} = max\{\pi_{Hp}H - y_p - H + y_r, \pi_{Hp}L - y_p\}$*
      ii. *$y_r > \delta T$: $b_{Hr} = \delta T$, $b_L = \pi_L L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = \frac{H-\delta T}{H-L}$, $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, $\delta T_{Hp} = max\{H(\pi_{Hp} - 1) + \delta T - y_p, \pi_{Hp}L - y_p\}$*

   (b) *else: $b_{Hr} = L$, $b_L = L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = 1$, $\pi_{Hp} = \frac{N-n_L-n_{Hr}}{n_{Hp}}$, $\delta T_{Hp} = max\{\pi_{Hp}H - y_p - H + y_r, \pi_{Hp}L - y_p\}$*

2. *Case 2: $y_p < L < H < y_r$:*
   (a) *if $-n_{Hr}(H - L) + n_L L - \varepsilon n_L > 0$; $b_{Hr} = L$, $b_L = L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = 1$, $\pi_{Hp} = \frac{N-n_{Hr}-n_L}{n_{Hp}}$, $\delta T_{Hp} = \pi_{Hp}L - y_p$*

   (b) *else; $b_{Hr} = \delta T$, $b_L = \pi_L L$, $b_{Hp} = y_p$, $\pi_{Hr} = 1$, $\pi_L = \frac{H-\delta T}{H-L}$, $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, $\delta T_{Hp} = max\{H(\pi_{Hp} - 1) - y_p + \delta T, \pi_{Hp}L - y_p\}$*

3. *Case 3:* $y_p < y_r < L < H$:

   (a) *if* $y_r > y_p + L(1 - \pi_{Hp})$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$:
   $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = 1$, $\pi_{Hp} = \frac{N - n_L - n_{Hr}}{n_{Hp}}$, $\delta T_{Hp} = y_r - y_p - L(1 - \pi_{Hp})$

   (b) *else:* $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_L - n_{Hr})}{n_{Hp}}$, $\delta T_{Hp} = 0$

Proof. See Appendix. □

When introducing the participation constraints for the high-valuation agents, the bureaucrat first checks whether the previously derived mechanism violates the participation constraint and leads to that type of agent dropping out. If it does not, the previously derived mechanism is retained. If it does, a new mechanism is derived. The participation constraint of the rich high-valuation agents is violated whenever these agents have to pay a very high bribe but official regulation is relatively low, i.e. lower than the bribe demand. In this case the bureaucrat chooses the highest bribe demand for the rich high-valuation agent that keeps these agents participating. Because a rich high-valuation agent pays a lower bribe, the bureaucrat can increase the probability, a rich low-valuation agent gets the good. Thereby he can increase the bribe paid by the low-valuation agents without inducing the rich high-valuation agents to mimic the low-valuation agents. This results in a higher income for the bureaucrat but also in an inferior allocation of goods because fewer goods are given to the poor high-valuation agents.

We continue by computing the bureaucrat's optimal mechanism if official regulation is low, $\delta T < L$. In this case every agent receives a positive utility from waiting such that the individual rationality constraints are given by:

$$(IR_L)\ \pi_L L - b_L \geq p_L L - \delta T$$
$$(IR_{Hr})\ \pi_{Hr} H - b_{Hr} \geq p_{Hr} H - \delta T$$
$$(IR_{Hp})\ \pi_{Hp} H - b_{Hp} - \delta T_{Hp} \geq p_{Hp} H - \delta T$$

We immediately include all constraints in the calculation.

Proposition 3. *For* $\delta T < L$ *the solution to problem (1) is given by*

1. *Case 1:* $y_p < L < y_r < H$:

   (a) *if* $y_p \geq \delta T$: $b_{Hr} = \delta T$, $b_L = \delta T$, $b_{Hp} = \delta T$, $\pi_{Hr} = \pi_L = \pi_{Hp} = \frac{N}{n_H + n_L}$, $\delta T_{Hp} = 0$

(b) *else:* $b_{Hr} = \delta T$, $b_L = \delta T$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L$

  i. *if* $\delta T > y_p + (1 - \pi_{Hp})L$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$: $\pi_L = 1$, $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$, $\delta T_{Hp} = (\pi_{Hp} - 1)L - y_p + \delta T$

  ii. *else:* $\pi_L = \frac{n_H L + n_{Hp}(\delta T - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$, $\delta T_{Hp} = 0$

2. *Case 2:* $y_p < L < H < y_r$*: same as Case 1*

3. *Case 3:* $y_p < y_r < L < H$*:*

   (a) $\delta T \geq y_r$*:*

   i. *if* $y_r > y_p + L(1 - \pi_{Hp})$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$: $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = 1$, $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$, $\delta T_{Hp} = y_r - y_p - L(1 - \pi_{Hp})$

   ii. *else:* $b_{Hr} = y_r$, $b_L = y_r$, $b_{Hp} = y_p$, $\pi_{Hr} = \pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_L - n_{Hr})}{n_{Hp}}$, $\delta T_{Hp} = 0$

   (b) *else:*

   i. *if* $y_p \geq \delta T$: $b_{Hr} = \delta T$, $b_L = \delta T$, $b_{Hp} = \delta T$, $\pi_{Hr} = \pi_L = \pi_{Hp} = \frac{N}{n_H + n_L}$, $\delta T_{Hp} = 0$

   ii. *else:* $b_{Hr} = \delta T$, $b_L = \delta T$, $\pi_{Hr} = \pi_L$, $b_{Hp} = y_p$

   A. *if* $\delta T > y_p + (1 - \pi_{Hp})L$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$: $\pi_L = 1$, $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$, $\delta T_{Hp} = (\pi_{Hp} - 1)L - y_p + \delta T$

   B. *else:* $\pi_L = \frac{n_H L + n_{Hp}(\delta T - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$, $\delta T_{Hp} = 0$

Proof. See Appendix. □

If all agents obey the official waiting time to obtain the good, rich agents do not pay a bribe that is higher than the level of official regulation. Therefore, the highest bribe the bureaucrat can extract from the rich agents equals the level of official regulation such that a pooling equilibrium, either partial or complete, always results. If the poor agents' income is higher than the official regulation, all agents pay the same bribe and receive the good with the same probability in a pooling equilibrium of all agents. If the poor agents' income is smaller than the official regulation, only the rich high- and low-valuation agents are pooled in equilibrium. In this case the bureaucrat has two objectives when determining the probability that the rich agents obtain the good and the waiting time imposed on the poor agents: he wants to reduce both the probability that the low-valuation agents get the good, and the waiting time of the poor agents. Because he incurs a higher cost from waiting than from misallocation, the bureaucrat reduces waiting time by raising the probability that the low-valuation agents get the good. He increases this probability until either waiting time equals zero or this probability equals one. Poor agents do not wait if they are sufficiently rich but they have to wait if they are so poor that the rich agents would mimic them if there was no waiting time.

2.2 Regulation in the absence of corruption

In this section we derive the level of official regulation chosen by the government if there is no corruption. The government wants to maximize allocative efficiency and minimize the agents' waiting time. We assume that the private and social valuations of the agents coincide. Note that regulation would become only more worthwhile if we changed this to assume that low-valuation agents impose a negative externality. Agents who want to get the good have to obey the entire official waiting time. Goods are allocated randomly at the end of the official waiting time among all agents waiting. The agent's decision to wait depends on the level of official regulation T. Social welfare is given by

$$S_T = (\pi_{Hp} n_{Hp} + \pi_{Hr} n_{Hr})H + n_L \pi_L L - \delta T(n_{Hr} + n_{Hp} + n_L|T). \tag{2}$$

Waiting time influences who decides to wait. There can be three cases: either all agents wait (if $\delta T < L$), only the high-valuation agents wait (if $\delta T \in [L, H]$) or no agent waits to get the good (if $\delta T > H$). As the last case is clearly not optimal, the government's decision reduces to choosing whether to allocate the goods to all agents with equal probability or whether to only allocate them to the agents with a high valuation.

If official regulation is so low that all agents wait, the probabilities of getting the good are given by $\pi_{Hr} = \pi_{Hp} = \pi_L = \frac{n_H}{n_H + n_L}$ such that social welfare is given by

$$S_{\delta T < L} = \frac{n_H}{n_H + n_L}(n_H H + n_L L) - \delta T(n_H + n_L). \tag{3}$$

If official regulation is so high that only the agents with a high valuation wait, the probabilities of getting the good are given by $\pi_{Hr} = \pi_{Hp} = 1$ and $\pi_L = 0$ such that social welfare is given by

$$S_{\delta T \in [L,H]} = n_H H - \delta T n_H. \tag{4}$$

For both low and high regulation, welfare decreases in official waiting time T. Hence, the lowest amount of regulation still guaranteeing the respective outcome will be chosen by the government. If regulation is low, the government chooses $T = 0$. If regulation is high, the government chooses $T = L$.

Proposition 4. *In the absence of corruption, the government chooses the high regulation, $T = L$, instead of the low regulation, $T = 0$, if $S_{T=L} \geq S_{T=0}$ which holds if*

$$\frac{n_L}{n_L + n_H} \frac{H - L}{L} \equiv \delta_{nc} \geq \delta. \tag{5}$$

Proof. $S_{T=L} \geq S_{T=0}$ if $n_H H(1 - \frac{n_H}{n_H + n_L}) - \frac{n_H}{n_H + n_L} n_L L \geq \delta(T_H n_H - T_L(n_H + n_L))$. Rearranging gives the result above. □

The government faces a tradeoff: higher regulation results in a better allocation of goods but also increases the agents' costs from waiting. Therefore, the government chooses to impose the high waiting time if agents incur only a small waiting cost.

2.3 Regulation in the presence of corruption

In this section we compute the level of official regulation chosen by the government if there is corruption. Because the level of official regulation results in different outside options for the agents and therefore different mechanisms designed by the bureaucrat, we proceed as follows: we first determine the level of regulation, the government chooses, conditional on the regulation being either high ($\delta T \in [L, H]$) or low ($\delta T < L$). We then compare the optimal level of the high and the optimal level of the low regulation in order to find the level of regulation that the government chooses overall. We also need to assess how the payment of bribes influences social welfare. There are three different possibilities: Firstly, one can view bribes as neutral monetary transfers from agent to bureaucrat without any impact on social welfare. Secondly, a rapacious government could look positively upon bribes paid, for example because it later receives a share thereof from the bureaucrat. Lastly, a government might view bribe payments as bad per se because it wants to eradicate that type of corruption. Let g capture the government's perception of bribery. If $g = 0$, bribes are perceived as neutral monetary transfers. If $g > 0$, bribes are viewed as bad in themselves. If $g < 0$, the government perceives bribes as good.

We start by computing the government's solution if regulation is high ($\delta T \in [L, H]$). If there is corruption, no agent follows the official regulation T. Instead the government considers T_{Hp}, the waiting time imposed by the bureaucrat on the poor agents.

$$\begin{aligned} S_{\delta T \in [L,H]} = (\pi_{Hr} n_{Hr} + \pi_{Hp} n_{Hp}) H + \pi_L n_L L - \delta T_{Hp} n_{Hp} \\ - g(n_{Hr} b_{Hr} + n_{Hp} b_{Hp} + n_L b_L) \end{aligned} \tag{6}$$

Corruption overrides the intended official regulation. Instead the bureaucrat uses the mechanism that he derived to maximize his income. Therefore, when the government wants to assess the effect of the official regulation on social welfare, the waiting time and probabilities to consider are those derived and applied by the bureaucrat.

As an example, we derive the government's choice of official regulation if $\delta T \in [L, H]$ for Case 2.(b).[15] Social welfare in this case is given by

$$S_{\delta T\in[L,H]} = n_{Hr}H + \pi_{Hp}n_{Hp}H + \pi_L n_L L - \delta T_{Hp}n_{Hp} - g(n_{Hr}\delta T + n_{Hp}y_p + n_L\pi_L L)$$

where $\delta T_{Hp} = \max\{\pi_{Hp}L - y_p, \pi_{Hp}H - y_p - H + \delta T\}$, $\pi_{Hp} = \frac{n_{Hp}-\pi_L n_L}{n_{Hp}}$ and $\pi_L = \frac{H-\delta T}{H-L}$. We first need to determine the chosen level of waiting time depending on the official regulation. The bureaucrat chooses $\delta T_{Hp} = \pi_{Hp}L - y_p$ if $H - \pi_{Hp}(H - L) \geq \delta T$. We calculate the official level of waiting time at which both values of waiting time for the poor agents, δT_{Hp}, are equalized by inserting π_L and π_{Hp} in $\delta T = (1 - \pi_{Hp})H + \pi_{Hp}L$.[16]

$$\delta T = \frac{n_{Hp}L + n_L H}{n_{Hp} + n_L} \tag{7}$$

In order to compute the change in social welfare induced by a change in official regulation, we take the derivative of social welfare with respect to official regulation

$$\frac{\partial S_{\delta T\in[L,H]}}{\partial T} = n_{Hp}\frac{\partial \pi_{Hp}}{\partial \pi_L}\frac{\partial \pi_L}{\partial T}H + n_L\frac{\partial \pi_L}{\partial T}L - n_{Hp}\frac{\partial \delta T_{Hp}}{\partial T} - g\left(\delta n_{Hr} + n_L\frac{\partial \pi_L}{\partial T}L\right)$$

For the exposition let us consider a level of official regulation smaller than the one defined in equation (7). In this case, the relevant level of waiting time imposed by the bureaucrat is $\delta T_{Hp} = \pi_{Hp}L - y_p$. We insert $\frac{\partial \pi_{Hp}}{\partial \pi_L} = -\frac{n_L}{n_{Hp}}$, $\frac{\partial \pi_L}{\partial T} = -\frac{\delta}{H-L}$ and $\frac{\partial \delta T_{Hp}}{\partial T} = L\frac{\partial \pi_{Hp}}{\partial \pi_L}\frac{\partial \pi_L}{\partial T} = \delta\frac{n_L}{n_{Hp}}\frac{L}{H-L}$ in the above equation. Social welfare increases in official regulation if

$$\frac{\partial S_{\delta T\in[L,H]}}{\partial T} = \delta\left(n_L\frac{H-2L}{H-L} - g\left(n_{Hr} - n_L\frac{L}{H-L}\right)\right) \geq 0.$$

We can conduct this analysis for every possible mechanism chosen by the bureaucrat if regulation is high, $\delta T \in [L, H]$, in order to derive the government's choice of official regulation.

Proposition 5. *If regulation T is restricted to be high, $\delta T \in [L, H]$, the government chooses the following regulation:*

1. *Case 1: $y_p < L < y_r < H$:*

[15] $y_p < L < H < y_r$ and $-n_{Hr}(H - L) + n_L L - \varepsilon n_L \leq 0$

[16] At this level of δT, $\delta T_{Hp} = \pi_{Hp}L - y_p = \pi_{Hp}H - y_p - H + \delta T$.

(a) *if* $-n_{Hr}(H-L)+n_L L-\varepsilon n_L<0$:

i. *if* $\delta T \geq y_r$: $\delta T \in [y_r, H]$

ii. $y_r > \delta T$:

A. *if* $\delta T \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

I. *if* $-n_L \frac{L}{H-L} - n_{Hp} \geq g\left(n_{Hr} - n_L \frac{L}{H-L}\right)$: $\delta T = y_r$

II. *else:* $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

B. *if* $\delta T \leq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

I. *if* $n_L \frac{H-2L}{H-L} \geq g\left(n_{Hr} - n_L \frac{L}{H-L}\right)$: $\delta T = \min\left\{\frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}, y_r\right\}$

II. *else:* $\delta T = L$

(b) *else:* $\delta T \in [L, H]$

2. *Case 2:* $y_p < L < H < y_r$:

(a) *if* $-n_{Hr}(H-L)+n_L L-\varepsilon n_L \geq 0$: $\delta T \in [L, H]$

(b) *else:*

i. *if* $\delta T \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

A. *if* $-n_L \frac{L}{H-L} - n_{Hp} \geq g\left(n_{Hr} - n_L \frac{L}{H-L}\right)$: $\delta T = H$

B. *else:* $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

ii. *if* $\delta T \leq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

A. *if* $n_L \frac{H-2L}{H-L} \geq g\left(n_{Hr} - n_L \frac{L}{H-L}\right)$: $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

B. *else:* $\delta T = L$

3. *Case 3:* $y_p < y_r < L < H$: $\delta T \in [L, H]$

Proof. See Appendix. □

If the government chooses a high level of official regulation, $\delta T \in [L, H]$, in some cases social welfare is independent of the chosen level of regulation. This occurs whenever the bureaucrat uses a mechanism that only depends on the underlying parameter values but not on the official regulation. Whenever the level of official regulation restricts the bribes, the bureaucrat can demand, social welfare also depends on the official level of regulation. The government considers that an increase in official waiting time increases bribe payments and waiting time, which decreases social welfare, but also raises the probability that the poor high-valuation agents obtain the good, which improves the allocation of goods and increases social welfare. Whenever the cost of misallocation is very high, the government increases official regulation to reduce misallocation even if this results in higher waiting time for the poor agents and higher bribe payments for the rich agents. Because an increase in official regulation also increases bribe payments, a government liking bribe payments is more likely to raise regulation than a government disliking bribe payments. Even a government encouraging bribery, however, does not always raise regulation because of the associated higher waiting costs for the poor agents.

Proposition 6. *If regulation T is restricted to be low, $\delta T \in [0, L)$, the government chooses the following regulation:*

1. *Case 1: $y_p < L < y_r < H$:*
   (a) *if $y_p \geq \delta T$:*
      i. *if $-g\delta > 0$: $\delta T = y_p$*
      ii. *if $-g\delta = 0$: $\delta T \in [0, y_p]$*
      iii. *else: $\delta T = 0$*
   (b) *else:*
      i. *if $\delta T \geq y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$:*
         A. *if $-\delta(\delta n_{Hp} + g(n_{Hr} + n_L)) > 0$: $\delta T = L$*
         B. *else: $\delta T = y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$*
      ii. *else:*
         A. *if $\delta \left(\frac{n_{Hp} n_L}{n_H + n_L} \left(-\frac{H}{L} + 1 \right) - g \left(n_{Hr} + n_L \right) \right) > 0$: $\delta T = y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$*
         B. *else: $\delta T = y_p$*
2. *Case 2: $y_p < L < H < y_r$: same as Case 1*
3. *Case 3: $y_p < y_r < L < H$:*
   (a) *$\delta T \geq y_r$: $\delta T \in [y_r, L)$*
   (b) *else:*
      i. *if $y_p \geq \delta T$:*
         A. *if $-g\delta > 0$: $\delta T = y_p$*
         B. *if $g = 0$: $\delta T \in [0, y_p]$*
         C. *else: $\delta T = 0$*
      ii. *else:*
         A. *if $\delta T \geq y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$:*
            I. *if $-\delta(\delta n_{Hp} + g(n_{Hr} + n_L)) > 0$: $\delta T = y_r$*

II. *else:* $\delta T = y_p + (1 - \pi_{Hp})L$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$

B. *if* $y_p + (1 - \pi_{Hp})L > \delta T$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$*:*

I. *if* $\delta \left(\frac{n_L n_{Hp}(L-H)}{L(n_H + n_L)} - g\left(n_H + n_L\right) \right) > 0$*:* $\delta T = y_p + (1 - \pi_{Hp})L$ *with* $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$

II. *else:* $\delta T = y_p$

Proof. See Appendix. □

For some combinations of agents' income, agents' valuation for the good and official regulation, the bureaucrat chooses a mechanism that is independent of the official level of regulation provided that it is in a certain interval. Whenever this is the case, official regulation has no impact on social welfare such that the government is indifferent between all values of official regulation in this range. On the other hand, whenever the official regulation limits the bribes, the bureaucrat can ask for, the government usually has a clear stance on whether it is optimal to increase or decrease official regulation. An exception are very low levels of official regulation. In this case, a government unconcerned about bribing can be indifferent between all levels of regulation that are sufficiently small, i.e. $\delta T \in [0, y_p]$, because all of them lead to the same allocation but different bribe levels.

In proposition 6 we have derived the optimal regulatory choice for different intervals of the level of regulation. The next step is to find the optimal level of regulation conditional on regulation being low, $\delta T \in [0, L)$. As an example let us consider the case where the rich agents' income is higher than the valuation of the low type, $y_r > L$. Because it is customary in the literature (e.g. Guriev 2004), we will assume that the government perceives bribes as neutral monetary transfers ($g = 0$). Let us start at a level of regulation of $\delta T \geq y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$. Then we are in Case 1.(b)i. and it follows that the government wants to reduce official regulation to its lowest level. This would be exactly $\delta T = y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$. But now we are in Case 1.(b)ii. and the government again wants to reduce official regulation to its lowest level because this raises social welfare. Hence, the government reduces official regulation to $\delta T = y_p$. Now we are in Case 1.(a). A government that views bribes as neutral monetary transfers ($g = 0$) is indifferent between all possible values of regulation if $\delta T \in [0, y_p]$. If official regulation is that low decreasing it even further only impacts the bribes paid by the agents but not the allocation of goods. A government unconcerned about the total volume of bribes, however, only cares about the allocation reached. It is worth pointing out that no agent actually follows the official regulation set by the government. Even if official regulation is at the highest level in this interval, i.e. $\delta T = y_p$, goods are still randomly distributed and no agent waits. If we thus compare a level of regulation of $\delta T = y_p$ both in the presence and in the absence of corruption, we notice that the allocation of goods remains the same because all agents would wait for the good given that waiting time is lower than the valuation of the low type ($\delta T = y_p < L$). The cost of waiting, however, is only incurred if there is no corruption because the corrupt bureaucrat in charge of distributing the goods translates the waiting time into bribes. If these are perceived as neutral monetary transfers, the government can impose higher regulation if there is corruption without impacting social welfare. This implies that for official regulation $\delta T = y_p$, the allocation is the same with and without corruption but social welfare is higher if there is corruption than if there is no corruption. To see this, we can look of the level of social welfare for both cases. If there is no corruption, official regulation $\delta T = y_p$ leads to a surplus of $S_{\delta T = y_p} = \frac{n_H}{n_H + n_L}(n_H H + n_L L) - (n_H + n_L) y_p$. The same level of official regulation leads to a surplus of $S_{\delta T = y_p} = \frac{n_H}{n_H + n_L}(n_H H + n_L L)$ if there is corruption and the government views bribes neutrally as monetary transfers ($g = 0$).

We continue by deriving the level of official regulation a government viewing bribes neutrally ($g = 0$) chooses conditional on official regulation being low.

Proposition 7. *If $\delta T \in [0, L)$ and $g = 0$, the government chooses a regulation of $\delta T \in [0, y_p]$. The resulting social welfare is given by*

$$S_{\delta T \in [0, y_p]} = \frac{n_H}{n_H + n_L}(n_H H + n_L L).$$

Proof. See Appendix. □

We continue by deriving the level of official regulation chosen by a government that perceives bribes as neutral monetary transfers ($g = 0$) if regulation is restricted to be high. The first step is to evaluate the conditions derived in proposition 5 at $g = 0$. As an example, we use Case 2.(b) again. From 2.(b)i. follows that $\delta T = \frac{n_{Hp} L + n_L H}{n_{Hp} + n_L}$ and from 2.(b)ii. follows that $\delta T = \frac{n_{Hp} L + n_L H}{n_{Hp} + n_L}$ if $H \geq 2L$ and $\delta T = L$ if this condition on the valuations for the good is not satisfied. The government therefore chooses the high level of regulation if we are in Case 2.(b) and $H \geq 2L$. Conducting a similar analysis for every case, we compute the government's choice of regulation if regulation is high.

Proposition 8. *If $\delta T \in [L, H]$ and $g = 0$, the government chooses the following level of regulation:*

1. *Case 1: $y_p < L < y_r < H$:*

    (a) if $-n_{Hr}(H - L) + n_L L - \varepsilon n_L < 0$:

i. *if* $H \geq 2L$*:*

A. *if* $y_r \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$*:* $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

B. *else:* $\delta T \in [y_r, H]$

ii. *else:* $\delta T = L$

(b) *else:* $\delta T \in [L, H]$

2. *Case 2:* $y_p < L < H < y_r$*:*

(a) *if* $-n_{Hr}(H-L) + n_L L - \varepsilon n_L > 0$*:* $\delta T \in [L, H]$

(b) *else:*

i. *if* $H \geq 2L$*:* $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$

ii. *else:* $\delta T = L$

3. *Case 3:* $y_p < y_r < L < H$*:* $\delta T \in [L, H]$

Proof. See Appendix. □

We are now in the position to compute the government's regulatory choice over the entire space of all possible regulations. We use again Case 2.(b) as an example. We have already seen that for $\delta T \in [L, H]$, $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$ is the optimal level of regulation if $H \geq 2L$.[17] Social welfare is given by

$$S_{\delta T \in [L,H]} = n_H H - n_L(H - \delta T) - n_{Hp}\left(L - \frac{H-\delta T}{H-L}\frac{n_L}{n_{Hp}}L - y_p\right).$$

Inserting that $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$ gives

$$S_{\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}} = n_H H - n_L\left(\frac{(H-L)n_{Hp}}{n_{Hp}+n_L}\right) - n_{Hp}\left(L\frac{n_{Hp}}{n_{Hp}+n_L} - y_p\right).$$

The government compares this level of social welfare to that of implementing the optimal level of the low regulation, $S_{\delta T \in [0,y_p]} = \frac{n_H}{n_H+n_L}(n_H H + n_L L)$. The government chooses the higher level of regulation if $S_{\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}} \geq S_{\delta T \in [0,y_p]}$. This condition can be written as

$$(H-L)\frac{n_L n_L n_{Hr}}{(n_H+n_L)} \geq n_{Hp}\left(n_{Hp}L - y_p(n_{Hp}+n_L)\right). \tag{8}$$

[17] At this level of regulation δT, the two different values for δT_{Hp} are equalized, $\delta T_{Hp} = \pi_{Hp}L - y_p = \pi_{Hp}H - y_p - H + \delta T$.

A similar condition for Case 1.(a) is given by

$$\frac{n_L}{(n_H+n_L)(H-L)}(n_H L^2 - n_L H(H-2L)) \geq n_{Hp}(L-y_p) - \frac{n_L}{(H-L)} y_r(H-2L). \quad (9)$$

By comparing the welfare for the optimal level of high regulation, ($\delta T \in [L, H]$), and the optimal level of low regulation, ($\delta T \in [0, L)$), for the different combinations of income and valuation, we find the optimal level of regulation overall.

Proposition 9. *A government with $g = 0$ chooses the following level of regulation:*

1. *Case 1: $y_p < L < y_r < H$:*
   (a) *if $-n_{Hr}(H-L) + n_L L - \varepsilon n_L < 0$:*
      i. *if $y_r \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$:*
         A. *if $H \geq 2L$ and equation (8): $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$*
         B. *else: $\delta T \in [0, y_p]$*
      ii. *else:*
         A. *if $H \geq 2L$ and equation (9): $\delta T \in [y_r, H]$*
         B. *else: $\delta T \in [0, y_p]$*
   (b) *else: $\delta T \in [0, y_p]$*
2. *Case 2: $y_p < L < H < y_r$:*
   (a) *if $-n_{Hr}(H-L) + n_L L - \varepsilon n_L > 0$: $\delta T \in [0, y_p]$*
   (b) *else:*
      i. *if $H \geq 2L$ and equation (8): $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$*
      ii. *else: $\delta T \in [0, y_p]$*
3. *Case 3: $y_p < y_r < L < H$: $\delta T \in [0, y_p]$*

Proof. See Appendix. □

The government increases regulation if the costs resulting from misallocation are high. The following example shows that the government chooses a high level of regulation if the bureaucrat is corrupt and no regulation if the bureaucrat is honest.

Example 1

We assume the following parameter values: $H = 20$, $L = 4$, $n_{Hr} = 150$, $n_{Hp} = 200$, $n_L = 175$, $y_r = 25$, $y_p = 0.5$, $\tau = 0.1$, $\varepsilon = 0.01$ and $\delta = 1.5$.

Assumptions 1, 2, 3, 4 and 5 are satisfied. We are in Case 2.(b) because $y_p < L < H < y_r$ and $-n_{Hr}(H-L) + n_L L - \varepsilon n_L = -1701.75 < 0$. If there is corruption, the government chooses the high regulation because $H \geq 2L$ and $(H-L)\frac{n_L n_L n_{Hr}}{(n_H+n_L)} - n_{Hp}\left(n_{Hp}L - y_p(n_{Hp}+n_L)\right) = 17,500 > 0$. The corresponding level of regulation is given by $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L} = 7.64$ in this case. If there is no corruption, the government chooses no regulation, $T = 0$, because $\frac{n_L}{n_L+n_H}\frac{H-L}{L} - \delta = -0.166$. Because low-valuation agents only attach a value of 4 to the good, which would translate into the highest level of regulation ever chosen in the absence of corruption, the optimal level of regulation if there is corruption is almost twice as high.

Higher regulation results in a better allocation but also leads to higher waiting time. If the first effect dominates, the government increases regulation.

Proposition 10. *If $y_r \geq \max\left\{L, \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}\right\}$, $g = 0$ and $H \geq 2L$ (cases 1.(a) and 2.(b)), the probability of a government choosing the high regulation $\delta T = \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$ instead of the low regulation $\delta T \in [0, y_p]$ increases in n_L, n_{Hr}, y_p and H; and decreases in n_{Hp} and L.*

Proof. See Appendix. □

The government is more likely to increase regulation if the high valuation is very high and the low valuation is very low. Regulation is also more beneficial if there are many agents with a low-valuation, n_L, because an increase in the size of this group raises the possible extent of misallocation. Regulation is also more likely higher if there are many rich high-valuation agents. This follows from higher regulation having a positive impact on allocation while the higher bribes paid by the rich high-valuation agents are not considered by the government. Interestingly, the government is less likely to increase regulation if there are many poor agents. This results from the waiting time which the bureaucrat imposes on the poor agents to reach a certain allocation. If regulation is high and high bribes can be extracted from the rich high-valuation agents, waiting time for the poor agents is high to prevent the rich agents from deviating. But if the group of poor agents is very large, many agents have to wait for a long time thereby decreasing social welfare.

3 Conclusion

This paper has investigated which level of regulation a government should choose if the bureaucracy is corrupt. We have seen that it is generally not optimal to choose the same level of regulation in the absence and presence of corruption. A government treating bribes as neutral monetary transfers increases regulation if the cost of misallocation is very high. A high level of misallocation results if there are many rich agents with a low-valuation or if the low valuation for the good is very low. Regulation is also increased if the waiting time imposed by the bureaucrat on the poor agents is very small, for example because this group is small or the poor agents are not too poor. This implies that a government facing corruption increases official regulation if individuals become richer.

Another interesting observation is that a government can indeed choose a higher level of regulation if there is corruption without having a negative effect on social welfare. The reason for this, as was hypothesized in the 1960's, is that corruption undoes the official regulation at least partly. A high level of regulation in the presence of corruption can therefore lead to the same allocation as a lower level of regulation if there is no corruption because the associated waiting time is not applied if there is corruption. If regulation is low, the bureaucrat cannot price discriminate between different types of agents and asks the same bribe from all rich agents. Studying the application process of getting a driver's license in India, Bertrand et al. (2007) indeed report that more able drivers are more likely to get the license but do not necessarily pay higher bribes.

The model predicts that a decrease in regulation has different effects on allocation depending on how corrupt the bureaucracy is. Lanau and Topalova (2016) study the effects of Italy's deregulation efforts conducted 2003-2013. They find that deregulation generally improves firms' productivity and value added. When factoring in the efficiency of provincial governments, however, the find that this effect is much higher for more efficient governments. We would usually assume that less efficient regional governments are also more corrupt such that a state-wide deregulation should have different effects at the regional level. It is then possible that more permits were given to worse producers than before the deregulation.

A possible modification of the model would be to allow the bureaucrat only to take bribes but not to impose waiting time. This would potentially retain the main results of the model while simplifying the exposition. The mechanism is meant to highlight the importance of the outside option in an applicant's decision to pay a bribe. Therefore, an alternative modeling approach would be to assume that applicants undergo testing that reveals their true type with a certain probability, similar to Guriev (2004). This would replace the assumption of goods being in limited supply.

A general insight of the model is that price discrimination of the bureaucrat leads to very high bribes. If wealth and valuation influence payments, rich agents who value the good a lot pay. But this in turn implies that the good is allocated to the agents with the highest valuations. If no price discrimination takes place, everyone pays the same bribe and the allocation becomes worse. This observation could provide insights into whether high bribes or the incidence of bribery should be targeted by anti-corruption efforts.

References

Acemoglu, D., Verdier, T. (2000). The choice between market failures and corruption, *The American Economic Review 90*(1), 194-211.

Aghion, P., Algan, Y., Cahuc, P, Shleifer, A. (2010). Regulation and distrust, *The Quarterly Journal of Economics 125*(3), 1015-1049.

Ahlin, C., Bose, P. (2007). Bribery, inefficiency, and bureaucratic delay, *Journal of Development Economics 84*, 465-486.

Aidt, T. S. (2003). Economic analysis of corruption: A survey, *The Economic Journal 113*(491), F632-F652.

Banerjee, A. (1997). A theory of misgovernance, *The Quarterly Journal of Economics 112*(4), 1289-1332.

Banerjee, A., Mullainathan, S., Hanna, R. (2012). Corruption, NBER Working Paper 17968.

Beck, P. J., Maher, M., W. (1986). A comparison of bribery and bidding in thin markets, *Economic Letters 20*, 1-5.

Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S. (2007). Obtaining a driver's license in India: An experimental approach to studying corruption, *The Quarterly Journal of Economics 122*(4), 1639-1676.

Bose, G. (2004). Bureaucratic delays and bribe-taking, *Journal of Economic Behavior & Control 54*, 313-320.

Breen, M., Gillanders, R. (2012). Corruption, institutions and regulation, *Economics of Governance 13*(3), 263-285.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2002). The regulation of entry, *The Quarterly Journal of Economics 117*(1), 1-37.

Esteban, J., Ray, D. (2006). Inequality, lobbying and resource allocation, *The American Economic Review 96*(1), 257-279.

Faría, H. J., Montesinos-Yufa, H. M., Morales, D. R., Aviles B., C. G., Brito-Bigott, O. (2013). Does corruption cause encumbered business regulations? An IV approach, *Applied Economics 45*(1), 65-83.

Guriev, S. (2004). Red tape and corruption, *Journal of Development Economics 73*, 489-504.

Huntington, S. P. (1968). *Political order in changing societies*. New Haven and London: Yale University Press.

International Monetary Fund (IMF) (2017). The role of the Fund in governance issues - Review of the guidance note - Preliminary considerations, *IMF Policy Paper*, Washington.

Kulshreshtha, P. (2007). An efficiency and welfare classification of rationing by waiting in the presence of bribery, *Journal of Development Economics 83*, 530-548.

Lanau, S. and P. Topalova (2016). The impact of product market reforms on firm productivity in Italy, *IMF Working Paper WP/16/119*, 1-24.

Leff, N. (1964). Economic development through bureaucratic corruption, *The American Behavioral Scientist 8*(3), 8-14.

Lui, F. T. (1985). An equilibrium queuing model of bribery, *Journal of Political Economy 93*(4), 760-781.

Méndez, F., Sepúlveda, F. (2013). Optimal government regulations and red tape in an economy with corruption, *Hitotsubashi Journal of Economics 54*, 51-77.

Méon, P.-G., Sekkat, K. (2005). Does corruption grease or sand the wheels of growth?, *Public Choice 122*, 69-97.

Myrdal, G. (1972). *Asian drama: An inquiry into the poverty of nations*. London: Allen Lane.

Nye, J. S. (1967). Corruption and political development: A cost-benefit analysis, *The American Political Science Review 61*(2).

Saha, B. (2001). Red tape, incentive bribe and the provision of subsidy, *Journal of Development Economics 65*, 113-133.

A Proofs

A.1 Proof of Remark 2

Proof. $T_L = 0$. A reduction in T_L relaxes (IR_L), $(L \geq Hr)$ and $(L \geq Hp)$. The opposite is true for $(Hr \geq L)$ and $(Hp \geq L)$. Hence, T_L can be reduced until either $(Hr \geq L)$ or $(Hp \geq L)$ binds.

Two cases: (IR_L) can bind or not. If (IR_L) does not bind, $b_L = y_r$ and $\pi_L = \pi_{Hr}$ (pooling between Hr and L). If $b_L < y_r$ it would be possible to raise b_L to make (IR_L) binding. Observe that because it is possible to give all goods to the high-valuation agents, it is always possible to make (IR_L) binding by setting $\pi_L = 0$. The only exception where π_L is not reduced occurs in a pooling equilibrium with $\pi_L = \pi_{Hr}$ and $b_L = b_{Hr}$.

Case 1: (IR_L) binding: as argued before we cannot have that both $b_L = y_r$ and we are in a pooling equilibrium with $\pi_L = \pi_{Hr}$. As long as $b_L < y_r$, the bureaucrat increases b_L. If (IR_L) binds, there are two possibilities, (i) $\pi_L L - b_L - \delta T_L = 0$ and (ii) $\pi_L L - b_L - \delta T_L = p_L L - \delta T$. In the first case it follows from (IR_L) that $b_L = \pi_L L - \delta T_L$ which remains satisfied if the change in T_L is such that $db_L = -\delta dT_L$. In the second case it follows from (IR_L) that $b_L = \delta T - \delta T_L$ such that again $db_L = -\delta dT_L$. Inserting this in $(Hr \geq L)$ and $(Hp \geq L)$ gives $\pi_{Hr}(H - b_{Hr} - \delta T_{Hr}) \geq \pi_L(H - L)$ and $\pi_{Hp}(H - b_{Hp} - \delta T_{Hp}) \geq \pi_L(H - L)$. Hence, T_L cancels in these equations, but reduces b_L. The bureaucrat reduces T_L.

If we are not in a pooling equilibrium with $\pi_L = \pi_{Hr}$ but $b_L = y_r$, the bureaucrat can reduce π_L as he reduces T_L until $(Hr \geq L)$ or $(Hp \geq L)$ becomes binding. Because $b_L = y_r$, Hp can never mimic L because $y_p < y_r$. $(Hp \geq L)$ is therefore irrelevant. Assume $(Hr \geq L)$ binds, $\pi_{Hr} H - b_{Hr} - \delta T_{Hr} = \pi_L H - b_L - \delta T_L$. Since $b_L = y_r$ implies that $y_r < L$, it follows from $L < H$ that also $b_{Hr} = y_r$. If $(Hr \geq L)$ binds, $(L \geq Hr)$ does not bind. Because $b_{Hr} = y_r > y_p$, Hp cannot mimic Hr and $T_{Hr} = 0$. But then $(Hr \geq L)$ reduces to $\pi_{Hr} H = \pi_L H - \delta T_L$. But for $T_L > 0$, this implies that $\pi_L > \pi_{Hr}$. We can reduce T_L and π_L and increase π_{Hr} to keep the constraint satisfied:

$$H d\pi_{Hr} = L d\pi_L - \delta dT_L$$

There are two ways to increase π_{Hr}: by decreasing π_L or by decreasing π_{Hp}. Following from the different outside options, there are two cases for (IR_L). From (IR_L), $\pi_L L - b_L - \delta T_L = p_L L - \delta T$, follows that $-b_L - \delta T_L = -\delta T$ such that $b_L = -\delta T_L$. A change in π_L cancels out. From (IR_L), $\pi_L L - y_r - \delta T_L = 0$ with $b_L = y_r$ follows that (IR_L) keeps being satisfied if $L d\pi_L - \delta dT_L = 0$.

Let the reduction in π_L be accompanied by an increase in π_{Hr}. Because $y_r =$

$b_{Hr} = b_L$, reducing T_L increases B. The change in the objective function is given by:

$$dB = n_{Hr}db_{Hr} + n_{Hp}db_{Hp} + n_L db_L - \varepsilon n_L d\pi_L - \tau(n_L dT_L + n_{Hr}dT_{Hr} + n_{Hp}dT_{Hp}).$$

For $db_L = db_{Hr} = db_{Hp} = 0$, $dT_{Hr} = dT_{Hp} = 0$ and $d\pi_L = \frac{\delta}{L}dT_L$ we have $dB = -\varepsilon n_L d\pi_L - \tau n_L dT_L = -n_L \varepsilon \frac{\delta}{L} dT_L < 0$. We can reduce T_L and π_L until either $T_L = 0$ or $b_L < y_r$ (already shown above).

Case 2: (IR_L) does not bind: $b_L = y_r$. This implies that $L > y_r$. Because $H > L > y_r$, $b_{Hr} = y_r$. From $y_r > y_p$ follows that $(Hp \geq L)$ becomes irrelevant (Hp cannot mimic L). (IR_L) not binding implies that π_L is higher than necessary because π_L can always be reduced to make (IR_L) binding as there are more agents than goods. Hence, the bureaucrat must prefer to have (IR_L) not binding instead of reducing π_L and increasing π_{Hp}. The bureaucrat decrease T_L until $(Hr \geq L)$ binds:

$$\pi_{Hr}H - y_r - \delta T_{Hr} = \pi_L H - y_r - \delta T_L$$

This continues to be satisfied if $dT_{Hr} = dT_L$. T_L is reduced until either (i) $T_L = 0$ or (ii) $T_{Hr} = 0$. In this case $\pi_{Hr}H = \pi_L H - \delta T_L$. From this and $T_L > 0$ follows that $\pi_L > \pi_{Hr}$. $(Hr \geq L)$ continues to be satisfied if

$$H d\pi_{Hr} = H d\pi_L - \delta dT_L.$$

From $\pi_{Hr} = \frac{N - \pi_L n_L - \pi_{Hp} n_{Hp}}{n_{Hr}}$ follows that $d\pi_{Hr} = -\frac{n_L}{n_{Hr}} d\pi_L$. Inserting this in $(Hr \geq L)$ above gives

$$H\left(-\frac{n_L}{n_{Hr}} d\pi_L\right) = H d\pi_L - \delta dT_L$$

such that $\delta dT_L = H(1 + \frac{n_L}{n_{Hr}}) d\pi_L$. Inserting this in dB gives

$$dB = -n_L dT_L \left(\varepsilon \frac{\delta}{H} \frac{n_{Hr}}{n_{Hr} + n_L} + 1\right) < 0$$

The bureaucrat reduces T_L.

$T_{Hr} = 0$.

A reduction in T_{Hr} makes $(Hr \geq Hp)$, $(Hr \geq L)$ and (IR_{Hr}) slacker. Reduce T_{Hr} until $(L \geq Hr)$ binds, $\pi_L L - b_L = \pi_{Hr} L - b_{Hr} - \delta T_{Hr}$, where we have used that $T_L = 0$. This condition continues to hold if the variables are changed such that $0 = d\pi_{Hr} L - db_{Hr} - \delta dT_{Hr}$. Inserting in objective function gives that T_{Hr} is decreased and b_{Hr} increased as long as possible. Two cases: (IR_{Hr}) binds or not.

Case 1: (IR_{Hr}) binds: (i) $b_{Hr} = \pi_{Hr}H - \delta T_{Hr} - p_{Hr}H + \delta T$ or (ii) $b_{Hr} = \pi_{Hr}H - \delta T_{Hr}$. Insert in $(L \geq Hr)$: $\pi_L L - b_L \geq \pi_{Hr}L - \pi_{Hr}H + p_{Hr}H - \delta T$ and (ii) $\pi_L L - b_L \geq \pi_{Hr}(L - H)$. In both cases T_{Hr} cancels but reduces b_{Hr}. We can make a similar calculation if $(Hp \geq Hr)$ binds. Reduce T_{Hr}.

Case 2: (IR_{Hr}) does not bind. Then we must have that $b_{Hr} = y_r$ as otherwise we could increase b_{Hr} to make (IR_{Hr}) binding. Hp can never mimic Hr because $y_p < y_r = b_{Hr}$. Can distinguish between two different cases: (i) $y_r > L$, then L never wants to mimic Hr because $\pi_L L - b_{Hr} = \pi_L(L - y_r) < 0$ and no additional T_{Hr} is needed. If $y_r < L$, L can mimic Hr. Decrease T_{Hr} until $(L \geq Hr)$ binds, $\pi_L L - b_L = \pi_{Hr}L - y_r - \delta T_{Hr}$. The change in the objective function is given by

$$dB = n_{Hr}db_{Hr} + n_{Hp}db_{Hp} + n_L db_L - \varepsilon n_L d\pi_L - \tau(n_L dT_L + n_{Hr}dT_{Hr} + n_{Hp}dT_{Hp}).$$

Insert $db_{Hr} = db_L = db_{Hp} = 0$ and $dT_L = dT_{Hp} = 0$ to get

$$dB = -\varepsilon n_L d\pi_L - \tau n_{Hr} dT_{Hr}.$$

The constraint $(L \geq Hr)$ continues to be satisfied if $d\pi_L = d\pi_{Hr} - \frac{\delta}{L}dT_{Hr}$ (with $db_L = 0$). From $\pi_L = \frac{N - \pi_{Hr}n_{Hr} - \pi_{Hp}n_{Hp}}{n_L}$, $d\pi_L = -\frac{n_{Hr}}{n_L}d\pi_{Hr}$. Insert:

$$-\frac{n_{Hr}}{n_L}d\pi_{Hr} = d\pi_{Hr} - \frac{\delta}{L}dT_{Hr}$$

Hence, $d\pi_{Hr}(-1 - \frac{n_{Hr}}{n_L}) = -\frac{\delta}{L}dT_{Hr}$ and therefore

$$d\pi_L = -\frac{\delta}{L}\left(-1 - \frac{n_{Hr}}{n_L}\right)^{-1} dT_{Hr} - \frac{\delta}{L}dT_{Hr} = -\frac{\delta}{L}\frac{n_{Hr}}{n_{Hr} + n_L}dT_{Hr}$$

Insert $d\pi_L$ in dB to get:

$$dB = n_{Hr}dT_{Hr}\left(\varepsilon\frac{\delta}{L}\frac{n_L}{n_{Hr} + n_L} - \tau\right)$$

This is negative by assumption. An increase in π_L does not violate other incentive constraints: $(Hr \geq L)$ is slack because $(L \geq Hr)$ binds. $(L \geq Hp)$ becomes slacker. $(Hp \geq Lr)$ not relevant. □

A.2 Proof of Remark 3

Proof. An increase in π_{Hr} relaxes (IR_{Hr}), $(Hr \geq Hp)$ and $(Hr \geq L)$. The opposite is true for $(Hp \geq Hr)$ and $(L \geq Hr)$. If we increase π_{Hr}, we need to decrease either

π_L or π_{Hp} to keep feasibility, $\pi_{Hr} + \pi_L + \pi_{Hp} = 1$. Because of the assumption that $\frac{N-n_{Hr}-n_L}{n_{Hp}} L \geq y_p$, increasing π_{Hr} and decreasing π_{Hp} does not affect b_{Hp} and can only reduce T_{Hp} which increases the bureaucrat's utility. If an increase in π_{Hr} leads to $(Hp \geq Hr)$ binding, this is not a concern because if Hp can pay b_{Hr} there is no need to differentiate between Hp and Hr. Suppose instead that we increase in π_{Hr} until $(L \geq Hr)$ binds, $\pi_L L - b_L = \pi_{Hr} L - b_{Hr}$. The condition continues to be satisfied if a change in π_{Hr} is such that $0 = L d\pi_{Hr} - db_{Hr}$. Increase b_{Hr} until $b_{Hr} = y_r$, $b_{Hr} = H$ or $b_{Hr} = \delta T$. If $y_r > L$ or $\delta T > L$, L never wants to mimic Hr and $(L \geq Hr)$ does not bind. Because an increase in π_{Hr} reduces waiting time T_{Hp}, it is optimal to increase π_{Hr} to $\pi_{Hr} = 1$. If $b_{Hr} = min\{\delta T, y_r\} \leq L$, L can mimic Hr. Increase π_{Hr} until $(L \geq Hr)$ binds, $\pi_L L - b_L = \pi_{Hr} L - b_{Hr}$. This condition continues to be satisfied if a change in π_L is such that $d\pi_L L = db_L$. Evaluate the change in the objective function.

$$dB = n_{Hr}\frac{db_{Hr}}{d\pi_{Hr}} + n_{Hp}\frac{db_{Hp}}{d\pi_{Hp}} + n_L\frac{db_L}{d\pi_L} - n_L \varepsilon d\pi_L - \tau(n_{Hp} dT_{Hp} + n_{Hr} dT_{Hr} + n_L dT_L)$$

If only b_L, π_L and T_{Hp} change:

$$dB = n_L db_L - \varepsilon n_L d\pi_L - \tau n_{Hp} dT_{Hp}$$

If $d\pi_L > 0$, $d\pi_{Hp} < 0$ such that $dT_{Hp} < 0$. Then, $dB > 0$ if $n_L > \varepsilon$ which is assumed to hold. Hence, increase π_L until $b_L = b_{Hr} = b$. Next, evaluate whether $\pi_L = \pi_{Hr}$ should be increased even further. An increase in π_L reduces T_{Hp} leading to an increase in B but increases misallocation leading to a decrease in B. Evaluating the change in the objective function:

$$dB = -\varepsilon n_L d\pi_L - \tau n_{Hp} dT_{Hp}$$

In order to assess the change in T_{Hp}, we need to derive T_{Hp} first. From $(Hr \geq Hp)$, $\pi_L H - b \geq \pi_{Hp} H - y_p - \delta T_{Hp}$, follows that $\delta T_{Hp} \geq H(\pi_{Hp} - \pi_L) - y_p + b$ and from $(L \geq Hp)$ follows that $\delta T_{Hp} \geq L(\pi_{Hp} - \pi_L) - y_p + b$. Then, $\delta T_{Hp} = max\{0, L(\pi_{Hp} - \pi_L) - y_p + b, H(\pi_{Hp} - \pi_L) - y_p + b\}$. For $V = L, H$, $V(\pi_{Hp} - \pi_L) - y_p + b \geq V(\pi_{Hp} - 1) - y_p + b$ because $1 \geq \pi_L$. Then, $max\{L(\pi_{Hp} - 1) - y_p + b, H(\pi_{Hp} - 1) - y_p + b\} = L(\pi_{Hp} - 1) - y_p + b$. From $b_L \geq L$, $L(\pi_{Hp} - 1) - y_p + b \geq L(\pi_{Hp} - 1) - y_p + L = \pi_{Hp} L - y_p > 0$ by assumption. Waiting time is always positive and therefore, $dT_{Hp} \neq 0$. From $\delta T_{Hp} = V(\pi_{Hp} - \pi_L) - y_p + b$ with $\pi_{Hp} = \frac{N - \pi_L(n_L + n_{Hr})}{n_{Hp}}$ follows that

$d\delta T_{Hp} = -V\frac{n_{Hr}+n_{Hp}+n_L}{n_{Hp}}d\pi_L$. Insert:

$$dB = -\varepsilon n_L d\pi_L + \tau V(n_{Hr} + n_{Hp} + n_L)d\pi_L$$

Then, $dB > 0$ if $\tau V(n_H + n_L) > \varepsilon n_L$ for both $V = H, L$. □

A.3 Proof of Proposition 1

Proof. Case 1: $b_{Hr} = y_r$, $b_L = L$ and $b_{Hp} = y_p$ are the highest bribes possible. From (IR_L) it follows that $b_L = \pi_L L$. Either $(Hr \geq L)$ or $(L \geq Hr)$ bind. First, suppose $b_{Hr} = y_r$ and $(Hr \geq L)$ binds: $H - y_r = \pi_L H - \pi_L L$. Then, $\pi_{Lr} = \frac{H-y_r}{H-L}$. Second, suppose $(L \geq Hr)$ binds: $\pi_L L - \pi_L L = L - b_{Hr}$. Then, $b_{Hr} = L$. From $(Hr \geq L)$ $H - L \geq \pi_H H - L$, it follows that $\pi_L = 1$. This gives rise to two potential solutions: (i) $b_{Hr} = y_r$, $\pi_L = \frac{H-y_r}{H-L}$; (ii) $b_{Hr} = L$, $\pi_L = 1$. The choice of the bureaucrat follows from the change in the objective function:

$$dB = n_{Hr}db_{Hr} + n_L db_L + n_{Hp}db_{Hp} - n_L\varepsilon d\pi_L - \tau(n_{Hr}dT_{Hr} + n_L dT_L + n_{Hp}dT_{Hp})$$

We want to evaluate the effect of a change in π_L. From (IR_L) binding follows that (IR_L) continues to hold if $Ld\pi_L = db_L$. From $(Hr \geq L)$, $H - b_{Hr} = \pi_L(H - L)$, follows that $(Hr \geq L)$ continues to hold if $-db_{Hr} = (H - L)d\pi_L$. We decrease π_{Hp} when increasing π_L to satisfy feasibility. Because y_p is small, this implies a change in T_{Hp}. This gives

$$dB = d\pi_L(-n_{Hr}(H - L) + n_L L - \varepsilon n_L) - \tau n_{Hp}dT_{Hp}.$$

Waiting time T_{Hp} is either determined by a binding $(Hr \geq Hp)$, $0 = Hd\pi_{Hp} - \delta T_{Hp}$, such that $dT_{Hp} = \frac{H}{\delta}d\pi_{Hp}$ or by a binding $(L \geq Hp)$, $0 = Ld\pi_{Hp} - \delta T_{Hp}$, such that $dT_{Hp} = \frac{L}{\delta}d\pi_{Hp}$. The change in T_{Hp} is given by $dT_{Hp} = \frac{V}{\delta}d\pi_{Hp}$ with $V = L, H$. From $\pi_{Hp} = \frac{N-\pi_L n_L-\pi_{Hr}n_{Hr}}{n_{Hp}}$ follows that $d\pi_{Hp} = -\frac{n_L}{n_{Hp}}d\pi_L$. Hence, $dT_{Hp} = -\frac{V}{\delta}\frac{n_L}{n_{Hp}}d\pi_L$. Inserting this in dB gives:

$$dB = d\pi_L(-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau\frac{V}{\delta}n_L)$$

This gives rise to two solutions. (i) If $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau\frac{V}{\delta}n_L < 0$, B falls in π_L, such that π_L is reduced to the lowest level $\pi_L = \frac{H-y_r}{H-L}$.

T_{Hp} follows from that $(Hr \geq Hp)$, $(Hp \geq L)$ and $(L \geq Hp)$ need to be satisfied. Note that $(Hp \geq L)$ and $(L \geq Hp)$ never bind simultaneously. From $(Hr \geq Hp)$, $H - y_r \geq \pi_{Hp}H - y_p - \delta T_{Hp}$, follows that $\delta T_{Hp} \geq \pi_{Hp}H - y_p - H + y_r$. From

$(L \geq Hp)$, $\pi_L L - \pi_L L = 0 \geq \pi_{Hp} L - y_p - \delta T_{Hp}$, follows that $\delta T_{Hp} \geq \pi_{Hp} L - y_p$.

$(Hp \geq Hr)$ is irrelevant because $b_{Hr} = y_r > y_p$ such that Hp cannot mimic Hr. We need to check whether $(Hp \geq L)$ is satisfied when $y_p \geq \pi_L L$: (i) $\delta T_{Hp} = \pi_{Hp} H - y_p - H + y_r$. $(Hp \geq L)$ becomes $H - y_r \geq \pi_L(H - L)$ which is satisfied. (ii) $\delta T_{Hp} = \pi_{Hp} L - y_p$. In this case $(L \geq Hp)$ binds which implies that $(Hp \geq L)$ is slack. Because $\pi_{Hp} L - y_p > 0$, waiting time is always positive.

(ii) If $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau \frac{V}{\delta} n_L > 0$, B increases in π_L, such that π_L is raised to the highest level $\pi_L = 1$. Hp can mimic neither Hr nor L because $b_L = b_{Hr} = L > y_p$. T_{Hp} again follows from that $(Hr \geq Hp)$ and $(L \geq Hp)$ need to be satisfied. $(Hr \geq Hp)$: $H - L \geq \pi_{Hp} H - y_p - \delta T_{Hp}$ gives $\delta T_{Hp} \geq \pi_{Hp} H - y_p - H + L$. $(L \geq Hp)$: $L - L = 0 \geq \pi_{Hp} L - y_p - \delta T_{Hp}$ gives $\delta T_{Hp} \geq \pi_{Hp} L - y_p$.

Note that it is possible that $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau \frac{H}{\delta} n_L > 0$ and $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau \frac{L}{\delta} n_L < 0$. We have excluded this case by assumption.

Case 2: The highest bribes possible are $b_{Hp} = y_p$, $b_{Hr} = H$ and, from (IR_L), $b_L = \pi_L L$. Either $(Hr \geq L)$ or $(L \geq Hr)$ binds. (i) Suppose $(Hr \geq L)$ binds $(H - b_{Hr} = \pi_L H - \pi_L L)$: $b_{Hr} = H - \pi_L(H - L)$ and $\pi_L = \frac{H - b_{Hr}}{H - L}$. (ii) Suppose $(L \geq Hr)$ binds $(\pi_L L - \pi_L L = L - b_{Hr})$: $b_{Hr} = L$.

For (i), evaluate the effect of a change in π_L: $db_{Hr} = -(H - L)d\pi_L$, $db_L = Ld\pi_L$ and $d\pi_{Hp} \neq 0$ to match the change in π_L.

$$dB = n_{Hr} db_{Hr} + n_L db_L - \varepsilon n_L d\pi_L - \tau n_{Hp} dT_{Hp}$$

The value of dT_{Hp} is such that $(Hr \geq Hp)$ and $(L \geq Hp)$ are both satisfied: $(Hr \geq Hp)$, $H - b_{Hr} \geq \pi_{Hp} H - b_{Hp} - \delta T_{Hp}$, and $(L \geq Hp)$, $L - b_L \geq \pi_{Hp} L - b_{Hp} - \delta T_{Hp}$. $(Hr \geq Hp)$ still holds if a change in the variables is such that $0 = H d\pi_{Hp} - \delta dT_{Hp}$. Similarly, $(L \geq Hp)$ still holds if $0 = L d\pi_{Hp} - \delta dT_{Hp}$. This implies that the change is $dT_{Hp} = \frac{V}{\delta} d\pi_{Hp}$ with $V = L, H$. From $\pi_{Hp} = \frac{N - n_L \pi_L - n_{Hr}\pi_{Hr}}{n_{Hp}}$, $d\pi_{Hp} = -\frac{n_L}{n_{Hp}} d\pi_L$. Inserting gives

$$dB = d\pi_L \left(-(H - L) n_{Hr} + n_L L - \varepsilon n_L + \tau \frac{H}{\delta} n_L \right)$$

This gives rise to two solutions. If $-(H - L)n_{Hr} + n_L L - \varepsilon n_L + \tau \frac{V}{\delta} n_L > 0$, increase π_L to $\pi_L = 1$; if $-(H - L)n_{Hr} + n_L L - \varepsilon n_L + \tau \frac{V}{\delta} n_L < 0$, decrease π_L to $\pi_L = 0$.

Note that it is possible that $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau \frac{H}{\delta} n_L > 0$ and $-n_{Hr}(H - L) + n_L L - \varepsilon n_L + \tau \frac{L}{\delta} n_L < 0$. We have excluded this case by assumption.

Solution (i), $\pi_L = 1$. T_{Hp} is such that $(Hr \geq Hp)$ and $(L \geq Hp)$ are satisfied. Because $\pi_L = 1$ implies $b_L = L$, $(Hp \geq L)$ is irrelevant as Hp can never pay b_L. If $(Hr \geq Hp)$ binds, $\delta T_{Hp} = \pi_{Hp} H - y_p - H + L$, if $(L \geq Hp)$ binds, $\delta T_{Hp} = \pi_{Hp} L - y_p$.

We can show that $\pi_{Hp}L - y_p > \pi_{Hp}H - y_p - H + L$.

Solution (ii), $\pi_L = 0$, $b_L = 0$, $b_{Hr} = H$, $\pi_{Hp} = 1$. T_{Hp} such that $(Hr \geq Hp)$: $0 \geq H - y_p - \delta T_{Hp}$ which requires $\delta T_{Hp} = H - y_p$ and $(L \geq Hp)$: $0 \geq L - y_p - \delta T_{Hp}$ which requires that $\delta T_{Hp} = L - y_p$. Then, $max\{L - y_p, H - y_p\} = H - y_p$.

Case 3: The highest bribe possible for Hr is $b_{Hr} = y_r$. (IR_L) either binds or does not bind. Suppose (IR_L) binds: $b_L = \pi_L L$. Increase π_L by decreasing π_{Hp}. Compute the change in π_{Hp} if π_L increases. From $\pi_{Hp} = \frac{N - n_{Hr} - \pi_L n_L}{n_{Hp}}$, $d\pi_{Hp} = -\frac{n_L}{n_{Hp}} d\pi_L$. Computing the change in the objective function:

$$dB = n_L db_L - \varepsilon n_L d\pi_L - \tau n_{Hp} dT_{Hp} = \left(n_L L - \varepsilon n_L\right) d\pi_L - \tau n_{Hp} dT_{Hp}$$

If $d\pi_L > 0$, then $dT_{Hp} < 0$. Because also $n_L > \varepsilon$, b_L is raised to the maximum value such that $b_L = y_r$. From $(L \geq Hr)$ and $(Hr \geq L)$ follows that $\pi_L = \pi_{Hr}$. We can now compute π_L and δT_{Hp}: From $(Hr \geq Hp)$, $\delta T_{Hp} \geq y_r - b_{Hp} - H(\pi_{Hr} - \pi_{Hp})$. From $(L \geq Hp)$, $\delta T_{Hp} \geq L(\pi_{Hp} - \pi_L) - b_{Hp} + y_r$. Lastly, we require that $\delta T_{Hp} \geq 0$. We do not need to check whether $(Hp \geq L)$ and $(Hp \geq Hr)$ are satisfied because Hp cannot pay the respective bribes. We can now derive $\pi_L = \pi_{Hr}$. The objective function is given by

$$B = (n_L + n_{Hr})y_r + n_{Hp}y_p - \varepsilon n_L \pi_L - \tau n_{Hp} T_{Hp}$$

Changing π_L induces the following change:

$$dB = -n_L \varepsilon d\pi_L - \tau n_{Hp} dT_{Hp}$$

Evaluate the change for the different values of T_{Hp}: (a) $\delta T_{Hp} = 0$. Then, $dT_{Hp} = 0$ such that $dB = -n_L \varepsilon d\pi_L < 0$ and the optimal choice is to reduce π_L to the smallest value possible. (b) $\delta T_{Hp} = y_r - y_p - L(\pi_L - \pi_{Hp})$. We can write $\pi_{Hp} = \frac{N - \pi_L(n_L + n_{Hr})}{n_{Hp}}$ and insert this in δT_{Hp} to get

$$\delta T_{Hp} = y_r - y_p + \frac{L}{n_{Hp}} \left(N - \pi_L(n_{Hr} + n_L + n_{Hp})\right)$$

Then, $\delta dT_{Hp} = -\frac{L}{n_{Hp}}(n_H + n_L) d\pi_L$ such that

$$dB = \left(-n_L \varepsilon + \tau \frac{L}{\delta}(n_H + n_L)\right) d\pi_L$$

Then, $dB > 0$ if $\tau(n_H + n_L)\frac{L}{\delta} > \varepsilon n_L$ or $\tau > \varepsilon \frac{\delta}{L} \frac{n_L}{n_H + n_L}$. Because by assumption, $\tau > \varepsilon \frac{\delta}{L} \frac{n_L}{n_{Hr} + n_L}$ and $\frac{n_L}{n_{Hr} + n_L} > \frac{n_L}{n_H + n_L}$, this is always satisfied such that $dB > 0$ if

$d\pi_L > 0$. Increase π_L to the highest value possible. (c) $\delta T_{Hp} = y_r - y_p - H(\pi_L - \pi_{Hp})$. Conducting a similar calculation as in (b) and remembering that $\frac{\delta}{L} > \frac{\delta}{H}$, we again arrive at $dB > 0$ if $d\pi_L > 0$ such that the optimal choice is to increase π_L to the highest value possible.

If π_L is at the highest value possible, $\pi_L = 1$, $\delta T_{Hp} = max\{y_r - y_p - H(1 - \pi_{Hp}), y_r - y_p - L(1 - \pi_{Hp})\} = y_r - y_p - L(1 - \pi_{Hp})$.

$(Hp \geq L)$ and $(Hp \geq Hr)$ are irrelevant because $y_p < y_r = b_L = b_{Hr}$. This is the solution if waiting time T_{Hp} is positive, $y_r - y_p - L(1 - \pi_{Hp}) > 0$. If $y_r - y_p - L(1 - \pi_{Hp}) < 0$, then $\delta T_{Hp} = 0$ and the bureaucrat wants to reduce π_L to the lowest level.

Because $(L \geq Hp)$ and $(Hr \geq Hp)$ still need to be satisfied, start by determining which value of waiting time is higher: $y_r - y_p - L(\pi_L - \pi_{Hp}) > y_r - y_p - H(\pi_L - \pi_{Hp})$ if $\pi_L > \pi_{Hp}$. Suppose this is true. The lowest level of π_L that still satisfies (IR_L), $\pi_L L - y_r = 0$, is $\frac{y_r}{L}$. We need to check whether the needed level of T_{Hp} is negative for $\pi_L = \frac{y_r}{L}$. In order to determine whether this level of π_L will be chosen, we insert this in T_{Hp}:

$$\delta T_{Hp} = y_r - b_{Hp} - L(\frac{y_r}{L} - \pi_{Hp}) = \pi_{Hp} L - y_p$$

But $y_p > \pi_L L$ by assumption. Even if $y_r - b_{Hp} - H(\frac{y_r}{L} - \pi_{Hp}) < 0$, the maximum and applied level of waiting time would still be positive. Therefore π_L has to be larger than $\frac{y_r}{L}$ to decrease T_{Hp} to 0. We can derive π_L from $\delta T_{Hp} = 0$, $(\pi_{Hp} - \pi_L)L - y_p + y_r = 0$. Inserting $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$ gives

$$\frac{N - \pi_L(n_{Hr} + n_{Hp} + n_L)}{n_{Hp}} L - y_p + y_r = 0$$

and finally $\pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$. We still need to check that $\pi_L \geq \pi_{Hp}$. Using that $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$ it follows that $\pi_L \geq \pi_{Hp}$ if $\pi_L \geq \frac{n_H}{n_H + n_L}$. We can check whether this is true for the solution $\pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$: $\pi_L = \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)} \geq \frac{n_H}{n_H + n_L}$ can be simplified to $n_{Hp}(y_r - y_p) \geq 0$ which is true. □

A.4 Proof of Proposition 2

Proof. Case 1: (a) (IR_{Hr}): $H - y_r \geq H - \delta T$ is satisfied if $\delta T \geq y_r$. $H - y_r \geq 0$ is always satisfied. Whether (IR_{Hp}): $\pi_{Hp} H - y_p - \delta T_{Hp} \geq \pi_{Hp} H - \delta T$ is satisfied depends on δT_{Hp}. There are different values for δT_{Hp}. (i) $\delta T_{Hp} = \pi_{Hp} H - y_p - H + y_r$. Inserting this in (IR_{Hp}) gives $\delta T \geq y_r - H(1 - \pi_{Hp})$. (ii) $\delta T_{Hp} = \pi_{Hp} L - y_p$. Inserting this in (IR_{Hp}) gives $\delta T \geq \pi_{Hp} L$. From $\delta T \geq L$, this is always satisfied.

Hp drop out if $\delta T_{Hp} = \pi_{Hp} H - y_p - H + y_r$ and $\delta T < y_r - H(1 - \pi_{Hp})$. Hr drop out if $\delta T < y_r$. For $\delta T \in (y_r - H(1 - \pi_{Hp}), y_r)$, Hr drop out but Hp do not. The bureaucrat adjusts the mechanism to keep Hr from dropping out.

Calculation of new optimal mechanism: From (IR_{Hr}) binding, $H-b_{Hr} = H-\delta T$, follows that $b_{Hr} = \delta T$. $(Hr \geq L)$ binding, $H - \delta T = \pi_L H - \pi_L L$, gives $\pi_L = \frac{H-\delta T}{H-L}$. Calculation of δT_{Hp}: (IR_{Hp}) is satisfied if $\delta T - y_p \geq \delta T_{Hp}$ and $\pi_{Hp}H - y_p - \delta T_{Hp} \geq 0$. From $(Hr \geq Hp)$, $\delta T_{Hp} \geq H(\pi_{Hp}-1)+\delta T - y_p$. From $(L \geq Hp)$, $\delta T_{Hp} \geq \pi_{Hp}L - y_p$. The level of δT_{Hp} needed to guarantee incentive compatibility should not violate (IR_{Hp}): $\delta T_{Hp} = H(\pi_{Hp} - 1) + \delta T - y_p < \delta T - y_p$ because $\pi_{Hp} < 1$; similarly, $\pi_{Hp}H - y_p - H(\pi_{Hp} - 1) - \delta T + y_p = H - \delta T \geq 0$. $\delta T_{Hp} = \pi_{Hp}L - y_p \leq \delta T - y_p$ because $\delta T \geq L$; similarly, $\pi_{Hp}H - y_p - \pi_{Hp}L + y_p \geq 0$. Which T_{Hp} is chosen depends on T.

(b) (IR_{Hr}): $H - L \geq H - \delta T$ is satisfied because $\delta T \geq L$. Similarly, $H - L \geq 0$. (IR_{Hp}): $\pi_{Hp}H - y_p - \delta T_{Hp} \geq \pi_{Hp}H - \delta T$ reduces to $\delta T \geq y_p + \delta T_{Hp}$. We also need to ensure that $\pi_{Hp}H - y_p - \delta T_{Hp} \geq 0$. Check whether (IR_{Hp}) holds for the different values of δT_{Hp}. (i) $\delta T_{Hp} = \pi_{Hp}H - y_p - H + L$. (IR_{Hp}) holds if $\delta T \geq \pi_{Hp}H - H + L$. This can be rewritten as $\delta T - L \geq H(\pi_{Hp} - 1)$ and is always satisfied because $\delta T \geq L$ and $1 > \pi_{Hp}$. Similarly, $\pi_{Hp}H - y_p - (\pi_{Hp}H - y_p - H + L) = H - L \geq 0$. (ii) $\delta T_{Hp} = \pi_{Hp}L - y_p$. (IR_{Hp}) holds because $\delta T \geq y_p + \pi_{Hp}L - y_p = \pi_{Hp}L$. Similarly, $\pi_{Hp}H - y_p - (\pi_{Hp}L - y_p) = \pi_{Hp}(H - L) \geq 0$.

Case 2:

(a) (IR_{Hr}): $H - L \geq H - \delta T$ is satisfied because $\delta T \geq L$. Also $H - L \geq 0$ is satisfied. (IR_{Hp}): $\pi_{Hp}H - y_p - \pi_{Hp}L + y_p \geq \pi_{Hp}H - \delta T$ reduces to $\delta T - \pi_{Hp}L \geq 0$. This is satisfied because $\delta T \geq L$. Also $\pi_{Hp}H - y_p - \pi_{Hp}L + y_p = \pi_{Hp}(H - L) \geq 0$ is satisfied.

(b)(IR_{Hr}): $H - H \geq H - \delta T$. This is not satisfied because $H > \delta T$. Compute the new mechanism in this case: (IR_{Hr}) binds, $H - b_{Hr} = H - \delta T$, such that $b_{Hr} = \delta T$. $(Hr \geq L)$ binds: $H - \delta T = \pi_L H - b_L$. From (IR_L) follows that $\pi_L L = b_L$. Insert $b_L = \pi_L L$ in $(Hr \geq L)$ to get: $H - \delta T = \pi_L H - \pi_L L = \pi_L(H - L)$. This gives $\pi_L = \frac{H-\delta T}{H-L}$. From $(Hr \geq Hp)$, $H - \delta T \geq \pi_{Hp}H - y_p - \delta T_{Hp}$, follows that $\delta T_{Hp} \geq \pi_{Hp}H - y_p - H + \delta T$. From $(L \geq Hp)$: $\pi_L L - \pi_L L \geq \pi_{Hp}L - y_p - \delta T_{Hp}$ follows that $\delta T_{Hp} \geq \pi_{Hp}L - y_p$. Need to take the maximum value of these two. Compare the different values for δT_{Hp}: $\pi_{Hp}H - y_p - H + \delta T \geq \pi_{Hp}L - y_p$ reduces to $\pi_{Hp}(H - L) \geq H - \delta T$. If $\delta T > \pi_{Hp}L + (1 - \pi_{Hp}H)$, $\pi_{Hp}H - y_p - H + \delta T$ is chosen. Waiting time could be zero if both $\pi_{Hp}H - y_p - H + \delta T < 0$ and $\pi_{Hp}L - y_p < 0$. Because $\pi_{Hp}L - y_p > 0$, waiting time is always positive.

(IR_{Hp}) is satisfied if $\pi_{Hp}H - y_p - \delta T_{Hp} \geq p_{Hp}H - \delta T$ which reduces to $\delta T - y_p \geq \delta T_{Hp}$. We need to check whether (IR_{Hp}) is satisfied for the different possible values of δT_{Hp}. If $\delta T_{Hp} = \pi_{Hp}H - y_p - H + \delta T$, (IR_{Hp}) is satisfied if $\delta T - y_p >$

$\pi_{Hp}H - y_p - H + \delta T$ which reduces to $(\pi_{Hp} - 1)H < 0$ and therefore holds. Similarly, $\pi_{Hp}H - y_p - \pi_{Hp}H + y_p + H - \delta T = H - \delta T \geq 0$. If $\delta T_{Hp} = \pi_{Hp}L - y_p$, (IR_{Hp}) is satisfied if $\delta T - y_p > \pi_{Hp}L - y_p$ which reduces to $\delta T > \pi_{Hp}L$ and holds because $\delta T \geq L$. Similarly, $\pi_{Hp}H - y_p - \pi_{Hp}L + y_p = \pi_{Hp}(H - L) \geq 0$.

Case 3:
(a) (IR_{Hr}): $H - y_r \geq H - \delta T$ is satisfied because $\delta T \geq L > y_r$. Same for $H - y_r \geq 0$. (IR_{Hp}): $\pi_{Hp}H - y_p - y_r + y_p + L(1 - \pi_{Hp}) \geq \pi_{Hp}H - \delta T$ reduces to $\delta T - y_r + L(1 - \pi_{Hp}) \geq 0$. This is satisfied because $\delta T \geq L > y_r$. $\pi_{Hp}H - y_p - y_r + y_p + L(1 - \pi_{Hp}) \geq 0$ reduces to $\pi_{Hp}(H - L) + L - y_r > 0$ and is also satisfied. (b) (IR_{Hr}): $\pi_{Hr}H - y_r \geq H - \pi_{Hr}\delta T$ is satisfied because $\delta T \geq L > y_r$. $\pi_{Hr}H - y_r \geq 0$ holds by construction but we can also show that $\pi_{Hr}H - y_r > \pi_{Hr}L - y_r \geq \frac{n_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}L - y_r = (n_{Hr} + n_L)(L - y_r) + n_{Hp}(L - y_p) > 0$. (IR_{Hp}): $\pi_{Hp}H - y_p \geq \pi_{Hp}H - \delta T$ because $\delta T > y_p$. $\pi_{Hp}H - y_p \geq 0$ because $\frac{N - \pi_L(n_L + n_{Hr})}{n_{Hp}}H - y_p \geq \frac{N - (n_L + n_{Hr})}{n_{Hp}}H - y_p > 0$ by assumption. □

A.5 Proof of Proposition 3

Proof. Case 1: From (IR_L): $b_L \leq \delta T$ follows $b_L = \delta T$. From (IR_{Hr}): $b_{Hr} \leq \delta T$ follows $b_{Hr} = \delta T$. From $(Hr \geq L)$: $\pi_{Hr}H - \delta T \geq \pi_L H - \delta T$ and $(L \geq Hr)$: $\pi_L L - \delta T \geq \pi_{Hr}H - \delta T$ follows that $\pi_L = \pi_{Hr}$. There can be two cases: (i) $y_p \geq \delta T$ and (ii) $\delta T > y_p$. If (i) $y_p \geq \delta T$, then $b_{Hp} = \delta T$. From $(Hr \geq Hp)$ and $(Hp \geq Hr)$ follows that $\pi_{Hr} = \pi_{Hp} = \pi_L$ and $\delta T_{Hp} = 0$.

If (ii) $\delta T > y_p$, then $b_{Hp} = y_p$. We need to ensure that all agents prefer to participate in the mechanism compared to not getting the good: Then, (IR_L) becomes $\pi_L L - \delta T \geq 0$ such that we need that $\pi_L \geq \frac{\delta T}{L}$. Similarly, it follows from (IR_{Hr}) that $\pi_{Hr} \geq \frac{\delta T}{H}$. The smallest value of $\pi_L = \pi_{Hr}$ that satisfied both of these conditions is therefore $\pi_L = \frac{\delta T}{L}$.

We now derive T_{Hp}. (IR_{Hp}) is satisfied if $\delta T - b_{Hp} \geq \delta T_{Hp}$. From $(Hr \geq Hp)$, $\delta T_{Hp} \geq \delta T - b_{Hp} - H(\pi_{Hr} - \pi_{Hp})$. Insert this in (IR_{Hp}) to get $\delta T - b_{Hp} > \delta T - b_{Hp} - H(\pi_{Hr} - \pi_{Hp})$ such that (IR_{Hp}) holds if $H(\pi_{Hr} - \pi_{Hp}) \geq 0$. From $(L \geq Hp)$, $\delta T_{Hp} \geq L(\pi_{Hp} - \pi_L) - b_{Hp} + \delta T$. Insert this in (IR_{Hp}) to get $\delta T - b_{Hp} > L(\pi_{Hp} - \pi_L) - b_{Hp} + \delta T$ such that (IR_{Hp}) holds if $L(\pi_L - \pi_{Hp}) \geq 0$. Lastly, we require that $\delta T_{Hp} \geq 0$. In this case (IR_{Hp}) holds if $\delta T \geq b_{Hp}$.

We do not need to check whether $(Hp \geq L)$ and $(Hp \geq Hr)$ are satisfied because Hp cannot pay the respective bribes because $y_p < \delta T = b_L = b_{Hr}$.

We can now derive $\pi_L = \pi_{Hr}$. The objective function is given by

$$B = (n_L + n_{Hr})\delta T + n_{Hp}y_p - \varepsilon n_L \pi_L - \tau n_{Hp}T_{Hp}$$

Changing π_L induces the following change:

$$dB = -n_L \varepsilon d\pi_L - \tau n_{Hp} dT_{Hp}$$

Evaluate the change for the different values of T_{Hp}: (a) $\delta T_{Hp} = 0$. Then, $dT_{Hp} = 0$ such that $dB = -n_L \varepsilon d\pi_L < 0$ and the optimal choice is to reduce π_L to the smallest value possible. (b) $\delta T_{Hp} = \delta T - b_{Hp} - L(\pi_L - \pi_{Hp})$. We can write $\pi_{Hp} = \frac{N - \pi_L(n_L + n_{Hr})}{n_{Hp}}$ and insert this in δT_{Hp} to get

$$\delta T_{Hp} = \delta T - b_{Hp} + \frac{L}{n_{Hp}} \left(N - \pi_L(n_{Hr} + n_L + n_{Hp})\right)$$

Then, $\delta dT_{Hp} = -\frac{L}{n_{Hp}}(n_H + n_L) d\pi_L$ such that

$$dB = \left(-n_L \varepsilon + \tau \frac{L}{\delta} (n_H + n_L)\right) d\pi_L$$

Then, $dB > 0$ if $\tau(n_H + n_L)\frac{L}{\delta} > \varepsilon n_L$ or $\tau > \varepsilon \frac{\delta}{L} \frac{n_L}{n_H + n_L}$. Because by assumption, $\tau > \varepsilon \frac{\delta}{L} \frac{n_L}{n_{Hr} + n_L}$ and $\frac{n_L}{n_{Hr} + n_L} > \frac{n_L}{n_H + n_L}$, this is always satisfied such that $dB > 0$ and the optimal choice is to increase π_L to the highest value possible. (c) $\delta T_{Hp} = \delta T - b_{Hp} - H(\pi_L - \pi_{Hp})$. Conducting a similar calculation as in (b) and remembering that $\frac{\delta}{L} > \frac{\delta}{H}$, we again arrive at $dB > 0$ and increasing π_L to the highest value possible.

If π_L is at the highest value possible, $\pi_L = 1$, $max\{\delta T - b_{Hp} - H(1 - \pi_{Hp}), \delta T - b_{Hp} - L(1 - \pi_{Hp})\} = \delta T - b_{Hp} - L(1 - \pi_{Hp})$ such that waiting time is $\delta T_{Hp} = max\{0, \delta T - b_{Hp} - L(1 - \pi_{Hp})\}$. If $\delta T_{Hp} = \delta T - b_{Hp} - L(1 - \pi_{Hp})$, we are in case (b) and $\pi_L = 1$. (IR_{Hp}) is satisfied because $L(\pi_L - \pi_{Hp}) = L(1 - \pi_{Hp}) \geq 0$.

If $\delta T_{Hp} = 0$, (IR_{Hp}) is satisfied because $\delta T \geq y_p$. We want to reduce π_L to the lowest level. $\pi_L = \frac{\delta T}{L}$ is the lowest level that still satisfies (IR_L). Insert this in δT_{Hp} and check whether the resulting level of δT_{Hp} is negative.

$$\delta T_{Hp} = \left(\pi_{Hp} - \frac{\delta T}{L}\right) L - y_p + \delta T = \pi_{Hp} L - y_p$$

Because $\pi_{Hp} L - y_p > 0$ by assumption, waiting time δT_{Hp} is positive if π_L is at the lowest level possible. We can derive π_L from equating δT_{Hp} to 0. This gives $(\pi_{Hp} - \pi_L)L - y_p + \delta T = 0$. Inserting $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$ gives

$$\frac{N - \pi_L(n_{Hr} + n_{Hp} + n_L)}{n_{Hp}} L - y_p + \delta T = 0$$

and finally $\pi_L = \frac{\pi_H L + n_{Hp}(\delta T - y_p)}{L(n_H + n_L)}$.

Case 2: same as Case 1.

Case 3: There are three possibilities: (a) $y_p < y_r < \delta T$ such that the maximum bribes possible are $b_{Hr} = b_L = y_r$ and $b_{Hp} = y_p$; (b) $y_p < \delta T < y_r$ such that the maximum bribes possible are $b_{Hr} = b_L = \delta T$ and $b_{Hp} = y_p$; (c) $\delta T < y_p < y_r$ such that the maximum bribes possible are $b_{Hr} = b_L = b_{Hp} = \delta T$. The proof of the first case is the same as for Case 3 with $\delta T \in [L, H]$. The last two cases are the same as Case 1. □

A.6 Proof of Proposition 5

Proof. Case 1

1.(a)i. Social welfare is given by

$$S_{\delta T \in [L,H]} = n_{Hr} H + \pi_{Hp} n_{Hp} H + \pi_L n_L L - \delta T_{Hp} n_{Hp} - g(n_{Hr} y_r + n_{Hp} y_p + n_L \pi_L L)$$

with $\pi_L = \frac{H - y_r}{H - L}$, $\pi_{Hp} = \frac{N - n_{Hr} - \pi_L n_L}{n_{Hp}}$ and $\delta T_{Hp} = max\{\pi_{Hp} H - y_p - H + y_r, \pi_{Hp} L - y_p\}$. Welfare is independent of T if $\delta T \in [L, H]$.

1.(a)ii. Social welfare is given by

$$S_{\delta T \in [L,H]} = n_{Hr} H + \pi_{Hp} n_{Hp} H + \pi_L n_L L - \delta T_{Hp} n_{Hp} - g(n_{Hr} \delta T + n_{Hp} y_p + n_L \pi_L L)$$

with $\pi_{Hp} = \frac{N - n_{Hr} - \pi_L n_L}{n_{Hp}}$, $\pi_L = \frac{H - \delta T}{H - L}$ and $\delta T_{Hp} = max\{H(\pi_{Hp} - 1) + \delta T - y_p, \pi_{Hp} L - y_p\}$. Then,

$$\frac{\partial S_{\delta T \in [L,H]}}{\partial T} = n_{Hp} H \frac{\partial \pi_{Hp}}{\partial T} + n_L \frac{\partial \pi_L}{\partial T} L - n_{Hp} \frac{\partial \delta T_{Hp}}{\partial T} - g\left(n_{Hr} \delta + n_L L \frac{\partial \pi_L}{\partial T}\right)$$

Inserting $\frac{\partial \pi_L}{\partial T} = -\frac{\delta}{H - L}$ and $\frac{\partial \pi_{Hp}}{\partial \pi_L} = -\frac{n_L}{n_{Hp}}$ gives

$$\frac{\partial S_{\delta T \in [L,H]}}{\partial T} = \delta \left(n_L - n_{Hp} \frac{\partial \delta T_{Hp}}{\partial T} - g\left(n_{Hr} - n_L \frac{L}{H - L}\right)\right)$$

This is the same expression as in Case 2.(b). The derivation is similar to that case.

1.(b) Social welfare is given by

$$S_{\delta T \in [L,H]} = n_{Hr} H + \pi_{Hp} n_{Hp} H + n_L L - \delta T_{Hp} n_{Hp} - g(n_{Hr} L + n_{Hp} y_p + n_L L)$$

with $\pi_{Hp} = \frac{N - n_L - n_{Hr}}{n_{Hp}}$ and $\delta T_{Hp} = max\{\pi_{Hp} H - y_p - H + y_r, \pi_{Hp} L - y_p\}$. Surplus is independent of T if $\delta T \in [L, H]$.

Case 2:

2.(a) Social welfare is given by

$$S_{\delta T\in[L,H]} = n_{Hr}H + \pi_{Hp}n_{Hp}H + n_L L - (\pi_{Hp}L - y_p)n_{Hp} - g(n_{Hr}L + n_{Hp}y_p + n_L L)$$

with $\pi_{Hp} = \frac{N-n_L-n_{Hr}}{n_{Hp}}$. Surplus is independent of T conditional on $\delta T \in [L,H]$.

2.(b) Social welfare is given by

$$S_{\delta T\in[L,H]} = n_{Hr}H + \pi_{Hp}n_{Hp}H + \pi_L n_L L - \delta T_{Hp}n_{Hp} - g(n_{Hr}\delta T + n_{Hp}y_p + n_L\pi_L L)$$

with $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, $\pi_L = \frac{H-\delta T}{H-L}$ and $\delta T_{Hp} = max\{H(\pi_{Hp}-1)-y_p+\delta T, \pi_{Hp}L-y_p\}$. Then,

$$\frac{\partial S_{\delta T\in[L,H]}}{\partial T} = n_{Hp}\frac{\partial \pi_{Hp}}{\partial \pi_L}\frac{\partial \pi_L}{\partial T}H + n_L\frac{\partial \pi_L}{\partial T}L - n_{Hp}\frac{\partial \delta T_{Hp}}{\partial T} - g\left(\delta n_{Hr} + n_L\frac{\partial \pi_L}{\partial T}L\right)$$

Inserting $\frac{\partial \pi_L}{\partial T} = -\frac{\delta}{H-L}$, from $\pi_L = \frac{H-\delta T}{H-L}$, and $\frac{\partial \pi_{Hp}}{\partial \pi_L} = -\frac{n_L}{n_{Hp}}$, from $\pi_{Hp} = \frac{N-n_{Hr}-\pi_L n_L}{n_{Hp}}$, gives

$$\frac{\partial S_{\delta T\in[L,H]}}{\partial T} = \delta n_L - n_{Hp}\frac{\partial \delta T_{Hp}}{\partial T} - g\left(\delta n_{Hr} - \frac{\delta n_L L}{H-L}\right)$$

In order to assess the derivative, we need to compute $\frac{\partial \delta T_{Hp}}{\partial T}$:

2.(b)i. If $\delta T \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$, $\frac{\partial \delta T_{Hp}}{\partial T} = \frac{\partial \pi_{Hp}}{\partial T}H + \delta = \delta\left(\frac{n_L}{n_{Hp}}\frac{H}{H-L}+1\right)$. Inserting gives

$$\begin{aligned}\frac{\partial S_{\delta T\in[L,H]}}{\partial T} =& \delta n_L - \delta n_{Hp}\left(\frac{n_L}{n_{Hp}}\frac{H}{H-L}+1\right) - g\left(\delta n_{Hr} - n_L\frac{\delta n_L L}{H-L}\right)\\ =& \delta\left(-n_L\frac{L}{H-L} - n_{Hp} - g\left(n_{Hr} - n_L\frac{L}{H-L}\right)\right)\end{aligned}$$

Hence, $\frac{\partial S}{\partial T} > 0$ if $-n_L\frac{L}{H-L} - n_{Hp} > g\left(n_{Hr} - n_L\frac{L}{H-L}\right)$. Because $-n_{Hr}(H-L) + n_L L - \varepsilon n_L < 0$ needs to hold to be in Case 2.(b), $n_{Hr} - n_L\frac{L}{H-L} + \frac{\varepsilon n_L}{H-L} > 0$. Because ε is close to zero, we will usually have that $n_{Hr} - n_L\frac{L}{H-L} > 0$ such that this can only be satisfied if $g < 0$.

2.(b)ii. in text.

Case 3:

If $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, welfare is given by

$$S_{\delta T\in[L,H]} = n_H H + n_L L - n_L H - \left(y_r - y_p - L\left(1 - \frac{n_{Hp}-n_L}{n_{Hp}}\right)\right) - g(n_{Hr}y_r + n_{Hp}y_p + n_L y_r)$$

Surplus is independent of T conditional on $\delta T \in [L, H]$.

If $y_r \leq y_p + L(1 - \pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp} - n_L}{n_{Hp}}$, welfare is given by

$$S_{\delta T \in [L,H]} = n_{Hp}H + \pi_L n_L(L - H) - g(n_{Hr}y_r + n_{Hp}y_p + n_L y_r)$$

with $\pi_L = \frac{\pi_H L + n_{Hp}(y_r - y_p)}{L(n_H + n_L)}$. Surplus is independent of T conditional on $\delta T \in [L, H]$.

□

A.7 Proof of Proposition 6

Proof. Case 1:

1.(a) Surplus is given by

$$S_{\delta T \in [0,L)} = \frac{N}{n_H + n_L}(n_{Hr}H + n_{Hp}H + n_L L) - g(n_{Hr}\delta T + n_{Hp}\delta T + n_L \delta T)$$

Then,

$$\frac{\partial S_{\delta T \in [0,L)}}{\partial T} = -g\delta(n_{Hr} + n_{Hp} + n_L)$$

1.(b) Surplus is given by

$$S_{\delta T \in [0,L)} = \pi_{Hr} n_{Hr} H + \pi_{Hp} n_{Hp} H + \pi_L n_L L - \delta T_{Hp} n_{Hp} - g(n_{Hr}\delta T + n_{Hp} y_p + n_L \delta T)$$

with $\pi_{Hr} = \pi_L$ and the other values depending on whether $\delta T > y_p + (1 - \pi_{Hp})L$ with $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$. Then,

$$\frac{\partial S_{\delta T \in [0,L)}}{\partial T} = (n_{Hr}H + n_L L)\frac{\partial \pi_L}{\partial T} + n_{Hp}H\frac{\partial \pi_{Hp}}{\partial T} - n_{Hp}\frac{\partial \delta T_{Hp}}{\partial T} - g\delta(n_{Hr} + n_L)$$

1.(b)i. $\pi_L = 1$, $\pi_{Hp} = \frac{N - n_{Hr} - n_L}{n_{Hp}}$, $\delta T_{Hp} = (\pi_{Hp} - 1)L - y_p + \delta T$. Insert $\frac{\partial \pi_L}{\partial T} = 0$, $\frac{\partial \pi_{Hp}}{\partial T} = 0$, $\frac{\partial \delta T_{Hp}}{\partial T} = \delta$ to get that

$$\frac{\partial S_{\delta T \in [0,L)}}{\partial T} = -\delta(\delta n_{Hp} + g(n_{Hr} + n_L))$$

The government only increases T if $g(n_{Hr} + n_L) < -\delta n_{Hp}$.

1.(b)ii. $\pi_L = \frac{n_H L + n_{Hp}(\delta T - y_p)}{L(n_H + n_L)}$, $\pi_{Hp} = \frac{N - \pi_L(n_{Hr} + n_L)}{n_{Hp}}$, $\delta T_{Hp} = 0$. Insert $\frac{\partial \pi_L}{\partial T} = \frac{\delta}{L}\frac{n_{Hp}}{n_H + n_L}$, $\frac{\partial \pi_{Hp}}{\partial \pi_L} = -\frac{n_{Hr} + n_L}{n_{Hp}}$, $\frac{\partial T_{Hp}}{\partial T} = 0$ to get

$$\frac{\partial S_{\delta T \in [0,L)}}{\partial T} = \delta\left(\frac{1}{n_H + n_L}\left(-\frac{H}{L}n_{Hp}n_L + n_L n_{Hp}\right) - g\left(n_{Hr} + n_L\right)\right)$$

Case 2: same as Case 1.

Case 3:

3.(a) If $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{N-n_{Hr}-n_L}{n_{Hp}}$ (case 3(a)i. in proposition 3), social welfare is given by

$$S_{\delta T\in[0,L)} = n_{Hr}H + \pi_{Hp}n_{Hp}H + n_L L - \delta n_{Hp}T_{Hp} - g(n_{Hr}y_r + n_L y_r + n_{Hp}y_p)$$

with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$ and $\delta T_{Hp} = y_r - y_p - L(1-\pi_{Hp})$. Surplus does not depend on T conditional on $\delta T \in [y_r, L)$.

If $y_r \le y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{N-n_{Hr}-n_L}{n_{Hp}}$ (case 3(a)i. in proposition 3), surplus is given by

$$S_{\delta T\in[0,L)} = \pi_{Hr}n_{Hr}H + \pi_{Hp}n_{Hp}H + \pi_L n_L L - g(n_{Hr}y_r + n_L y_r + n_{Hp}y_p)$$

with $\pi_{Hr} = \pi_L = \frac{n_H L + n_{Hp}(y_r-y_p)}{L(n_H+n_L)}$ and $\pi_{Hp} = \frac{N-\pi_L(n_L-n_{Hr})}{n_{Hp}}$. Surplus does not depend on T conditional on $\delta T \in [y_r, L)$.

3.(b)i. Social welfare is given by

$$S_{\delta T\in[0,L)} = \pi(n_{Hr}H + n_{Hp}H + n_L L) - g(n_H + n_L)\delta T$$

with $\pi = \frac{n_H}{n_H+n_L}$. Then,

$$\frac{\partial S_{\delta T\in[0,L)}}{\partial T} = -g(n_H + n_L)\delta$$

3.(b)ii.A. Surplus is given by

$$S_{\delta T\in[0,L)} = n_{Hr}H + \pi_{Hp}n_{Hp}H + n_L L - \delta n_{Hp}T_{Hp} - g(n_{Hr}\delta T + n_L\delta T + n_{Hp}y_p)$$

Then,

$$\frac{\partial S_{\delta T\in[0,L)}}{\partial T} = -\delta(\delta n_{Hp} + g(n_{Hr} + n_L))$$

3.(b)ii.B. Surplus is given by

$$S_{\delta T\in[0,L)} = \pi_{Hr}n_{Hr}H + \pi_{Hp}n_{Hp}H + \pi_L n_L L - g(n_{Hr}\delta T + n_{Hp}y_p + n_L\delta T)$$

with $\pi_L = \frac{n_H L + n_{Hp}(\delta T - y_p)}{L(n_H+n_L)}$ and $\pi_{Hp} = \frac{N-\pi_L(n_L+n_{Hr})}{n_{Hp}}$. Then,

$$\frac{\partial S_{\delta T\in[0,L)}}{\partial T} = \frac{\partial \pi_L}{\partial T}(n_{Hr}H + n_L L) + n_{Hp}H\frac{\partial \pi_{Hp}}{\partial T} - g\delta(n_{Hr} + n_L)$$

Insert $\frac{\partial \pi_L}{\partial T} = \frac{\delta n_{Hp}}{L(n_H+n_L)}$ and $\frac{\partial \pi_{Hp}}{\partial \pi_L} = -\frac{n_L+n_{Hr}}{n_{Hp}}$ to get

$$\frac{\partial S}{\partial T} = \delta \left(\frac{n_L n_{Hp}(L-H)}{L(n_H+n_L)} - g(n_H+n_L) \right)$$

$\square$

A.8 Proof of Proposition 7

Proof. We show the optimal regulation case by case following from proposition 6 if $g = 0$:

Case 1:

1.(a) $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} = 0$ leading to $\delta T \in [0, y_p]$. 1.(b)i. $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} = -\delta \delta n_{Hp} < 0$ leading to $\delta T = y_p + (1-\pi_{Hp})L$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$. 1.(b)ii. $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} = \frac{\delta}{n_H+n_L} n_{Hp} n_L \left(-\frac{H}{L}+1\right) < 0$ leading to $\delta T = y_p$.

Case 2: same as Case 1.

Case 3:

3.(a) $\delta T \in [y_r, L)$. 3.(b)i. $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} = 0$ such that $\delta T \in [0, y_p]$. 3.(b)ii.A. $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} < 0$ such that $\delta T = y_p + (1-\pi_{Hp})L$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$. 3.(b)ii.B. $\frac{\partial S_{\delta T \in [0,y_p]}}{\partial T} < 0$ such that $\delta T = y_p$ (condition for 3.(b)ii.).

We derive the regulation over all cases. Proof for $\delta T \in [0, y_p]$ for Case 1 in text. Proof for $\delta T \in [0, y_p]$ for Case 3. Start with $y_r > \delta T$ (Case 3.(b)). From 3.(b)ii.A. it is optimal to reduce δT to $\delta T = y_p + (1-\pi_{Hp})L$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$ and from 3.(b)ii.B. it follows that it is optimal to reduce δT to y_p. For $\delta T \in [0, y_p]$ all values of δT are equally good (3.(b)i.). But if we start with $\delta T \geq y_r$ (Case 3.(a)), the government is indifferent between all values of $\delta T \in [y_r, L)$ because the allocation does not depend on δT in this case. There are therefore two potential solutions: $\delta T \in [0, y_p]$ and $\delta T \in [y_r, L)$. Social welfare for $\delta T \in [0, y_p]$ is given by

$$S_{\delta T \in [0,y_p]} = \frac{n_H}{n_H+n_L}(n_H H + n_L L).$$

If $\delta T \in [y_r, L)$ social welfare can take two values depending on whether $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$.

(i) If $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, surplus is given by

$$S_{\delta T \in [y_r,L)} = n_{Hr} H + n_L L + \pi_{Hp} n_{Hp} H - n_{Hp}(y_r - y_p - L(1-\pi_{Hp})).$$

Inserting $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, this can be written as $S_{\delta T \in [y_r,L)} = n_H H + n_L(L-H) -$

$n_{Hp}(y_r - y_p - L(1-\pi_{Hp}))$. We can show that

$$\frac{n_H}{n_H+n_L}(n_H H + n_L L) \geq n_H H + n_L(L-H)$$

can be simplified to $0 > -n_L n_L$. This implies that $S_{\delta T\in[0,y_p]} \geq S_{\delta T\in[y_r,L)}$ and therefore $\delta T \in [0, y_p]$ is chosen if $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$.

(ii) If $y_r \leq y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, social surplus is given by:

$$S_{\delta T\in[y_r,L)} = (n_{Hr}H + n_L L)\pi_L + \pi_{Hp} n_{Hp} H$$

with $\pi_L = \frac{n_H L + n_{Hp}(y_r-y_p)}{L(n_H+n_L)}$ and $\pi_{Hp} = \frac{N-\pi_L(n_L-n_{Hr})}{n_{Hp}}$. This can be written as $S_{\delta T\in[y_r,L)} = n_H H + \pi_L(L-H)n_L$. Then $S_{\delta T\in[0,y_p]} \geq S_{\delta T\in[y_r,L)}$ if

$$\frac{n_H}{n_H+n_L}(n_H H + n_L L) \geq n_H H + \pi_L(L-H)n_L$$

which can be simplified to $n_{Hp}(y_r - y_p) \geq 0$. Therefore $\delta T \in [0, y_p]$ is also chosen if $y_r < y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$. □

A.9 Proof of Proposition 8

Proof. We show the optimal regulation following from proposition 5 if $g = 0$.
Case 1:
1.(a)i. $\delta T \in [y_r, H]$. 1.(a)ii. From proposition 5, 1.(a)ii.A. and 1.(a)ii.B. follows that $\delta T = \min\left\{y_r, \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}\right\}$ if $H \geq 2L$ and $\delta T = L$ else.
It remains to compare 1.(a)i. and 1.(a)ii. Social surplus if $\delta T \in [y_r, H]$ is given by

$$S_{\delta T\in[y_r,H]} = n_H H - (H-y_r)n_L - n_{Hp}\delta T_{Hp}.$$

There are two different values for δT_{Hp} if $\delta T \in [y_r, H]$ (Case 1.(a)i.): $\pi_{Hp}H - y_p - H + y_r$ and $\pi_{Hp}L - y_p$ both with $\pi_{Hp} = \frac{n_{Hp}-\pi_L n_L}{n_{Hp}}$ and $\pi_L = \frac{H-y_r}{H-L}$. We compare these two values and $\pi_{Hp}H - y_p - H + y_r \geq \pi_{Hp}L - y_p$ if

$$y_r \geq \frac{n_L H + n_{Hp}L}{n_{Hp}+n_L}.$$

Note that from Case 1.(a)ii. follows that the highest level of regulation the government ever chooses is $\delta T = \frac{n_L H + n_{Hp}L}{n_{Hp}+n_L}$ if $H \geq 2L$. For $y_r > \frac{n_L H + n_{Hp}L}{n_{Hp}+n_L}$, $\delta T = y_r$ is therefore never chosen. We only need to look at the case that $y_r < \frac{n_L H + n_{Hp}L}{n_{Hp}+n_L}$ which implies that $\delta T_{Hp} = \pi_{Hp}L - y_p$. The alternative level of regulation is $\delta T = L$ with $\pi_L = \frac{H-L}{H-L} = 1$ and $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$. For $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, we can show that

$\pi_{Hp}L - y_p > \pi_{Hp}H - y_p - H + L$ to find the level of waiting time in this case. Insert this to get social surplus:

$$S_{\delta T=L} = n_H H + n_L(L-H) - n_{Hp}(\frac{n_{Hp}-n_L}{n_{Hp}}L - y_p)$$

Then, $S_{\delta T\in[y_r,H]} \geq S_{\delta T=L}$ if

$$y_r n_L - n_{Hp}\left(\left(1 - \frac{H-y_r}{H-L}\frac{n_L}{n_{Hp}}\right)L - y_p\right) \geq n_L L - ((n_{Hp}-n_L)L - n_{Hp}y_p)$$

rearranging gives that this holds if $H \geq 2L$. The government chooses $\delta T \in [y_r, H]$ if $H \geq 2L$.

1.(b) $\delta T \in [L, H]$

Case 2:

2.(a) $\delta T \in [L, H]$. 2.(b) in text.

Case 3: $\delta T \in [L, H]$ □

A.10 Proof of Proposition 9

Proof. Case 1:

1.(a) If $y_r \geq \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$, the analysis is the same as for Case 2.(b) because the government never chooses a level of regulation higher than that. It remains to investigate $y_r < \frac{n_{Hp}L+n_L H}{n_{Hp}+n_L}$.

$$S_{\delta T\in[y_r,H]} = n_H H - (H-y_r)n_L - n_{Hp}\delta T_{Hp}$$

Insert $\delta T_{Hp} = \pi_{Hp}L - y_p$ with $\pi_{Hp} = 1 - \frac{n_L}{n_{Hp}}\frac{H-y_r}{H-L}$ and rewrite:

$$S_{\delta T\in[y_r,H]} = n_H H + \frac{n_L}{H-L}(H-y_r)(2L-H) + n_{Hp}(y_p - L)$$

We then have $S_{\delta T\in[y_r,H]} \geq S_{\delta T\in[0,y_p]}$ if

$$n_H H + \frac{n_L}{H-L}(H-y_r)(2L-H) + n_{Hp}(y_p-L) \geq \frac{n_H}{n_H+n_L}(n_H H + n_L L)$$

this can be rewritten as

$$\frac{n_L}{(n_H+n_L)(H-L)}(n_H L^2 - n_L H(H-2L)) \geq n_{Hp}(L-y_p) - \frac{n_L}{(H-L)}y_r(H-2L)$$

1.(b) Social welfare is given by

$$S_{\delta T\in[L,H]} = n_{Hr}H + n_L L + \pi_{Hp}n_{Hp}H - \delta T_{Hp}n_{Hp} = n_H H + n_L(L-H) - \delta T_{Hp}n_{Hp}$$

(i) $\delta T_{Hp} = \pi_{Hp}H - y_p - H + y_r$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$
Insert in $S_{\delta T\in[L,H]}$. Then we get

$$\begin{aligned} S_{\delta T\in[L,H]} &= n_H H + n_L(L-H) - \left(\frac{n_{Hp}-n_L}{n_{Hp}}H - y_p - H + y_r\right) \\ &= n_H H - n_L(H-L) - n_{Hp}\left(-\frac{n_L}{n_{Hp}}H - y_r\right) \end{aligned}$$

$S_{\delta T\in[0,y_p]} \geq S_{\delta T\in[L,H]}$ because we have shown in proposition 7, Case 3 that

$$\frac{n_H}{n_H+n_L}(n_H H + n_L L) > n_H H + n_L(L-H)$$

(ii) $\delta T_{Hp} = \pi_{Hp}L - y_p$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$ (same as Case 2.(a)).

Case 2:

2.(a) Social surplus is given by

$$\begin{aligned} S_{\delta T=L} =& n_{Hr}H + n_{Hp}\frac{n_{Hp}-n_L}{n_{Hp}}H + n_L L - n_{Hp}(\frac{n_{Hp}-n_L}{n_{Hp}} - y_p) \\ =& n_H H + n_L(L-H) - \left(n_{Hp} - n_L - \frac{y_p}{n_{Hp}}\right) \end{aligned}$$

We have shown in proposition 7, Case 3 that $\frac{n_H}{n_H+n_L}(n_H H+n_L L) > n_H H+n_L(L-H)$ such that $\delta T \in [0, y_p]$.

2.(b) in text for $H \geq 2L$. If $2L > H$, $\delta T = L$ is the optimal level of regulation for $\delta T \in [L, H]$ and we compare this to $\delta T \in [0, y_p]$. For $\delta T = L$: $\pi_L = \frac{H-L}{H-L} = 1$ such that $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$. In order to determine the relevant amount of waiting time for poor agents, we compare both levels that are possible: $H(\pi_{Hp} - 1) - y_p + L \geq (\pi_{Hp})L - y_p$ can be simplified to $0 \geq \frac{n_L}{n_{Hp}}(H - L)$ which is not true. Therefore, $\delta T_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}L - y_p$. Insert this in surplus:

$$\begin{aligned} S_{\delta T=L} =& n_{Hr}H + n_{Hp}\frac{n_{Hp}-n_L}{n_{Hp}}H + n_L L - n_{Hp}(\frac{n_{Hp}-n_L}{n_{Hp}} - y_p) \\ =& n_H H + n_L(L-H) - \left(n_{Hp} - n_L - \frac{y_p}{n_{Hp}}\right) \end{aligned}$$

We have already shown that $\delta T \in [0, y_p]$ is better.

Case 3:

If $y_r > y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, welfare is given by

$$S_{\delta T\in[L,H]} = n_H H + n_L L - n_L H - \left(y_r - y_p - L\left(1 - \frac{n_{Hp}-n_L}{n_{Hp}}\right)\right)$$

We have already shown in 7, Case 3 that $\frac{n_H}{n_H+n_L}(n_H H + n_L L) > n_H H + n_L(L-H)$ such that $\delta T \in [0, y_p]$ is chosen by the government.

If $y_r \leq y_p + L(1-\pi_{Hp})$ with $\pi_{Hp} = \frac{n_{Hp}-n_L}{n_{Hp}}$, welfare is given by

$$S_{\delta T\in[L,H]} = \pi_L(n_{Hr}H + n_L L) + (n_{Hp} - \pi_L(n_{Hr}+n_L))H = n_{Hp}H + \pi_L n_L(L-H)$$

with $\pi_L = \frac{\pi_H L + n_{Hp}(y_r-y_p)}{L(n_H+n_L)}$ Then, $S_{\delta T\in[L,H]} > S_{\delta T\in[0,y_p]}$ if

$$n_{Hp}H + n_L(L-H)\frac{n_H L + n_{Hp}(y_r-y_p)}{L(n_H+n_L)} \geq \frac{n_H}{n_H+n_L}(n_H H + n_L L)$$

We have already shown that this does not hold (proposition 7, Case 3). □

A.11 Proof of Proposition 10

Proof.

$$C = (H-L)\frac{n_L n_L n_{Hr}}{(n_H+n_L)} - n_{Hp}\left(n_{Hp}L - y_p(n_{Hp}+n_L)\right)$$

$$\frac{\partial C}{\partial n_{Hp}} = -(H-L)\frac{n_L n_L n_{Hr}}{(n_{Hr}+n_{Hp}+n_L)^2} - (2n_{Hp}L - 2n_{Hp}y_p - y_p n_L)$$

We can show that from $\frac{n_{Hp}-n_L}{n_{Hp}}L \geq y_p$

$$\begin{aligned} y_p n_L &- (H-L)\frac{n_L n_L n_{Hr}}{(n_{Hr}+n_{Hp}+n_L)^2} \\ &\leq \frac{n_L}{(n_{Hr}+n_{Hp}+n_L)^2}\left(\frac{n_{Hp}-n_L}{n_{Hp}}L(n_{Hr}+n_{Hp}+n_L)^2 - (H-L)n_L n_{Hr}\right) \\ &= \frac{n_L}{(n_{Hr}+n_{Hp}+n_L)^2}\left(L(-n_L^2 + n_L n_{Hp} + (n_{Hr}+n_{Hp})^2 - \frac{n_L}{n_{Hp}}(n_{Hr}+n_L)^2) - Hn_L n_{Hr}\right) \end{aligned}$$

Inserting this in the equation above gives

$$\begin{aligned} -(H-L)\frac{n_L n_L n_{Hr}}{(n_{Hr}+n_{Hp}+n_L)^2} - (2n_{Hp}L - 2n_{Hp}y_p - y_p n_L) &\leq -2n_{Hp}(L-y_p) \\ + \frac{n_L}{(n_H+n_L)^2}\left(L\left(-n_L^2 + n_L n_{Hp} + n_H^2 - \frac{n_L}{n_{Hp}}(n_{Hr}+n_L)^2\right) - Hn_L n_{Hr}\right) \end{aligned}$$

We apply again that $\frac{n_{Hp}-n_L}{n_{Hp}}L \geq y_p$:

$$-2n_{Hp}(L-y_p) + \frac{n_L}{(n_H+n_L)^2}\left(L\left(-n_L^2 + n_L n_{Hp} + n_H^2 - \frac{n_L}{n_{Hp}}(n_{Hr}+n_L)^2\right) - Hn_L n_{Hr}\right)$$
$$\leq \frac{n_L}{(n_H+n_L)^2}\left(L\left(-3n_L^2 - n_H^2 - 4n_L n_{Hr} - 3n_L n_{Hp} - \frac{n_L}{n_{Hp}}(n_{Hr}+n_L)^2\right) - Hn_L n_{Hr}\right)$$

Therefore, $\frac{\partial C}{\partial n_{Hp}} < 0$.

$$\frac{\partial C}{\partial H} = \frac{n_L n_L n_{Hr}}{n_H + n_L} > 0$$
$$\frac{\partial C}{\partial L} = -\frac{n_L n_L n_{Hr}}{n_H + n_L} - n_{Hp} n_{Hp} < 0$$
$$\frac{\partial C}{\partial n_L} = (H-L)\frac{n_L n_{Hr}(2n_H + n_L)}{(n_H+n_L)^2} + n_{Hp} y_p > 0$$
$$\frac{\partial C}{\partial n_{Hr}} = (H-L)\frac{n_L n_L (n_{Hp} + n_L)}{(n_H+n_L)^2} > 0$$
$$\frac{\partial C}{\partial y_p} = n_{Hp}(n_{Hp} + n_L) > 0$$

□

Chapter 3

Voters and Interest Groups: Organizing Votes to Influence Policy

1 Introduction

This paper investigates which stance on policy interest groups formed by voters take and how these groups with a large membership base can influence policy. While the impact of special interest groups on policy has received a lot of attention, the existing literature usually focuses on monetary contributions by interest groups as means to influence policy. The question of how interest groups with a large membership base but little funds, like the Sierra Club or the American Association of Retired People, can influence policy has received less attention. This paper shows that interest groups can use their informational advantage to trade their members' votes for changes in policy.

Interest groups can use two channels to influence policy: votes and monetary contributions (Stigler 1971). The question of how monetary incentives, especially in the form of campaign contributions, influence policy has received a lot of attention in the literature. The question of how interest groups can use votes to influence policy has received much less attention. Voting strength of interest groups has mostly been analyzed in the context of transfers and redistribution. I combine these two strands of the literature and show how interest groups can use votes to influence policy choices. Additionally, I model the endogenous formation of these interest groups.

I build a simple model where voters first organize in interest groups and then a politician chooses a policy. There are two types of politicians. While one type is purely office-motivated, the other type is purely policy-motivated. Only some voters can observe the policy choice before reelection takes place with voters organized in interest groups being better informed about the policy choice of the politician than the average voter. Voters engage in retrospective voting and punish or reward the politician for policy choices in the past. Voting behavior depends on whether the voter is organized in a group. Voters organized in interest groups can bundle their votes and reward the politician for choosing a policy close to the group's position.

Voters not organized in a group reward the politician for choosing a policy close to their own position. The office-seeking politician implements an interest group's policy if this group is sufficiently large to offset the votes lost from the informed but unorganized voters. I also endogenize the group formation process. Each voter can decide whether to found an interest group. I show that, in a symmetric equilibrium with two groups, the groups' positions are more extreme the higher the cost of founding a group and the lower the share of office-seeking politicians.

A common assumption in the previous literature is that special interest groups try to influence politicians by campaign contributions. Additionally, it is often assumed that these campaign spendings only influence uninformed voters while informed voters base their voting decision on the policy platform chosen by the politician (e.g. Baron 1994, Grossman and Helpman 1996). Denzau and Munger (1986) assume that all voters are uninformed and can be influenced by campaign contributions paid by an interest group to the politician. Empirical evidence on the other hand suggests that campaign contributions cannot influence ideological votes (Potters and Sloof 1996). Snyder and Ting (2008) model the strategic voting decision of an informed median voter when an interest group can influence the politician's policy choice with monetary transfers. The voter has preferences over the policy but prefers high types of the politician, who only cares about office-holding. The present analysis also assumes that some voters are better informed than others, but departs from the assumption that the politician trades off money against votes and instead looks at the tradeoff between votes from different voters.

The previous literature which assumed that interest groups have a certain voting strength has focused on the effects of voting on redistribution and taxation. Peltzman (1976) studies which wealth transfers are chosen by a regulator who wants to maximize votes where the probability of receiving a vote depends on whether the voter is subsidized or taxed. A similar analysis is conducted by Plotnick (1986) who also models the choice of redistributive programs when voting behavior depends on the transfer. Becker (1983) also studies the determinants of a group's influence on transfers and finds that groups can reduce taxes or raise their subsidies when they become more efficient at exerting political power relative to other groups.

The present paper is also related to the branch of the literature studying the influence of the politician's characteristics on policy choices. Alesina and Cukierman (1990) assume that a politician cares both about policy and about being in office, but study a politician's choice of ambiguity. Maskin and Tirole (2004) assume that politicians derive ego rents from being in office and are better informed about the best policy than the voter. In Kartik and McAfee (2007) politicians with character announce the best policy while politicians without character do not. Bernhardt et al. (2011) assume that candidates are defined by ideology and valence. Because the median voter values both valence and policy, higher-valence candidates are also reelected when they display more extreme positions. Anderson and Glomm (1992) study the advantage of being the incumbent and assume that politicians differ in character. The present paper assumes that there are different types of politicians and that incumbents can choose a policy in order to maximize the probability of reelection.

There are some papers studying the size and endogenous formation of interest groups. Hamlin and Jennings (2004) use the citizen-candidate model to approach the endogenous formation of political pressure groups. While they assume that the group with the highest activity rate can implement its preferred policy, I introduce a politician who chooses the policy thereby adding an additional dimension of uncertainty for the voters. Schneider (2014) shows that larger interest groups may lose when they grow past a certain threshold and groups use payments to reach their goals. Lastly, Brekke et al. (2007) assume that agents can produce a local and a public good and form groups endogenously. They then analyze the level of public good production in the different groups.

The next section introduces the model. Then the equilibrium is derived and analyzed. The last section concludes.

2 The model

An economy is populated by a continuum of voters and an incumbent politician. There are two periods. In the first period, voters can strategically form interest groups to influence the incumbent's policy choice by conditioning the votes of the entire group on the implemented policy. After the incumbent politician has set the policy in the first period, only some voters, including all voters organized in interest groups, can observe the policy choice. Then elections take place where the incumbent is challenged by a new politician. In the second period, the elected politician chooses a policy.

There are two types of politicians. One type cares only about being in office receiving ego rent h and solely wants to maximize the probability of reelection. If reelection is not a concern for the office-motivated politician, he chooses the welfare-maximizing policy, the median policy. The other type of politician is purely policy-motivated and always implements the policy of the median voter, μ.[1] Let $\beta \in [0, 1]$

[1] An alternative assumption would be that a politician is both office- and policy-motivated. In

denote the share of the office-seeking politicians among all politicians. Only the office-motivated politician behaves strategically. His utility function for both periods is given by

$$U(x) = h + qp(x)h \tag{1}$$

where $0 < q \leq 1$ is the discount factor and $p(x)$ is the reelection probability given the implemented policy x.

There is a continuum of voters. Voters, indexed by i, can be characterized by their policy preference, r_i, which is uniformly distributed between zero and one, $r_i \sim u[0,1]$. Voters have single-peaked preferences with their bliss policy placed at r_i.

$$u(x, r_i) = -(x - r_i)^2 \tag{2}$$

where x is the implemented policy. Each voter can first decide whether to found an interest group where this choice is denoted by $v^i \in \{0,1\}$.[2] The founder incurs a cost c that is constant and does not depend on group size. Examples include getting informed about legal issues, organizing a room or a mailing list. The groups' position equals that of the founder.[3] After observing all groups' positions, all voters who have not founded a group can choose to join a group where this choice is denoted by $w^i \in \{0,1\}$. Voter i enjoys benefit b from joining a group but incurs a cost $k(r_j - r_i)^2$ that increases in the distance from the group's position j. The benefit of being in a group could result from voters liking to be organized, being better informed or receiving a monetary benefit. Membership cost could be fees but also costs from supporting a position different from ones own. The voter's utility for two periods is given by

$$U^v(x, v^i, w^i) = u(\mathbb{E}x, r_i) - v^i c + w^i(b - k(r_j - r_i)^2) + qu(\mu, r_i) \tag{3}$$

where $0 < q \leq 1$ is the discount factor, $\mathbb{E}x$ is the expectation of the implemented policy and r_j is the position of the group joined as member.

There is a noisy signal which reveals the politician's policy decision to a fraction $\alpha \in [0,1]$ of the electorate. Then a fraction α of each interest group observes

this case, however, the analysis becomes more complicated. Additionally, the exact specification of the voting behavior becomes more important.

[2] One can also interpret this as the representative voter of a small group of voters amounting to a fraction ε of the overall population.

[3] One could assume that the founder chooses a position unequal from her own. Because potential members can always observe the policy position chosen, the only benefit of allowing diverging positions of founder and group would be that a different position is better in expectation. We assume that the founder chooses what she wants to see implemented which is her own policy preference.

the true policy. Assuming that group members communicate with each other, this information spreads to all members such that the entire group is informed. Hence, all members of interest groups can condition their voting behavior on the implemented policy. Also a fraction α of all unorganized voters observe the signal but do not communicate this information to other voters. Then, the informed citizens can base their voting decision on the policy while the remaining $1-\alpha$ unorganized and uninformed voters cannot.

Voters use their vote to punish the incumbent for his policy choice in the first period. Informed unorganized voters vote for the incumbent if their policy preference is closer to the chosen policy than to the median policy.[4] They randomize if the chosen policy is the median policy. Voters organized in groups bundle their votes and vote for the incumbent the closer the chosen policy to the group's policy. Uninformed voters randomize. The structure of the game is summarized in the following:

1. first period
   (a) first stage
      i. voters found interest groups
      ii. voters join existing interest groups
   (b) second stage
      i. politician sets policy
      ii. some voters observe the policy choice
      iii. election takes place
2. second period: elected politician sets policy

We restrict attention to subgame perfect Nash equilibria in pure strategies where voters first found and subsequently join groups, and then the politician chooses the policy.

Definition 1. An equilibrium is a policy choice x for the politician maximizing equation (1) and a vector of choices (v^i, w^i) for each voter i maximizing equation (3).

[4] An alternative assumption would be that informed unorganized voters reward the incumbent for choosing a policy closer to the median μ, but this implies that a voter with $r_i = 1$ could vote for the incumbent if $x < \mu$ and this seems unreasonable.

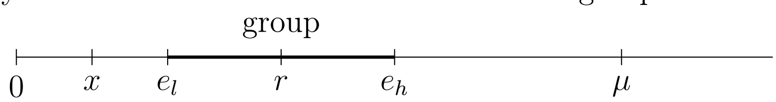
2.1 Reelection probabilities

We first derive the probability that the incumbent is reelected for different policy choices x. Because incumbent and challenger implement the same policy in the second period, voters are indifferent between both of them in terms of future policy. Voting is used to punish or reward the incumbent politician for the policy choice. The reelection probability of the incumbent therefore depends on the policy choice x. We start by assuming that there is one group to the left of the median policy with size n and position $r < \mu$. The interest group conditions the members' votes on the policy choice. The probability that the group votes for the incumbent increases the closer policy x is to the group's position. If the politician chooses a policy on the other side of the median, the group does not vote for the politician. The probability that the group votes for the incumbent is summarized below.

$$p_{group}(x) = \begin{cases} n\frac{\mu-r}{\mu-x} \text{ if } x \leq r \\ n\frac{\mu-x}{\mu-r} \text{ if } r \leq x \leq \mu \\ 0 \text{ if } \mu < x \end{cases}$$

Voters not organized in a group are either informed or uninformed about the policy choice. Informed voters condition their vote on the policy and vote for the incumbent if their policy preference is closer to policy x than to the median policy μ. If the incumbent chooses the median policy, $x = \mu$, informed voters randomize between voting for the incumbent and the challenger. Uninformed voters can not observe policy x and therefore always randomize between voting for the incumbent and the challenger, $\frac{1}{2}(1-n)(1-\alpha)$.

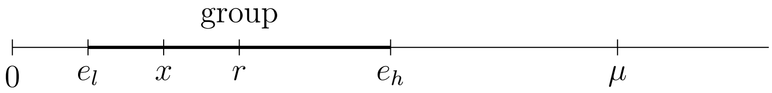
In order to determine the reelection probability, we need to consider the position of the chosen policy x. Depicted below is an interest group with position r where the most extreme group member is e_l and the most moderate member is e_h such that the group's size is given by $n = e_h - e_l$. We start with the case where the chosen policy is more extreme than the most extreme group member, $x < e_l$.



Each voter of the interest group votes for the politician with probability $\frac{\mu-r}{\mu-x}$ such that the entire expected votes from the group are $n\frac{\mu-r}{\mu-x}$. The vote of the informed voters depends on whether their policy preference is closer to policy x or to the median policy. The voter indifferent between voting for the incumbent or the challenger is given by $x+\frac{\mu-x}{2}$. This voter can be more or less moderate than the most moderate voter of the group, e_h. The expected vote share of the informed voters is therefore given by $\alpha(x + \frac{\mu-x}{2} - (e_h - e_l))$ if $x + \frac{\mu-x}{2} > e_h$ and αe_l if $e_h > x + \frac{\mu-x}{2}$. It is also possible that the voter indifferent between voting for the challenger and the incumbent is more extreme than the most extreme group member, $x + \frac{\mu-x}{2} < e_l$. The reelection probability is thus given by

$$p(x) = \begin{cases} n\frac{\mu-r}{\mu-x} + \alpha(x + \frac{\mu-x}{2} - n) + \frac{1}{2}(1-n)(1-\alpha) \text{ if } e_h < x + \frac{\mu-x}{2} \\ n\frac{\mu-r}{\mu-x} + \alpha e_l + \frac{1}{2}(1-n)(1-\alpha) \text{ if } e_l < x + \frac{\mu-x}{2} \leq e_h \\ n\frac{\mu-r}{\mu-x} + x + \frac{\mu-x}{2} \text{ if } x + \frac{\mu-x}{2} \leq e_l \end{cases}$$

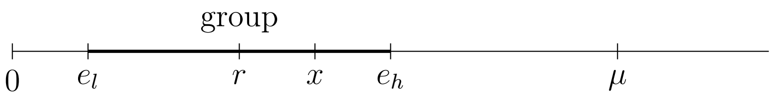
We continue by looking at the reelection probability if the policy is more moderate than the most extreme group member but more extreme than the group's position, $e_l < x < r$.



The vote share of the group is again given by $n\frac{\mu-r}{\mu-x}$. The votes of the informed unorganized voters depend on the whether the voter indifferent between voting for the challenger and the incumbent is more moderate than the most moderate group member. The reelection probability is given by

$$p(x) = \begin{cases} n\frac{\mu-r}{\mu-x} + \alpha(x + \frac{\mu-x}{2} - n) + \frac{1}{2}(1-n)(1-\alpha) \text{ if } x + \frac{\mu-x}{2} > e_h \\ n\frac{\mu-r}{\mu-x} + \alpha e_l + \frac{1}{2}(1-n)(1-\alpha) \text{ else} \end{cases}$$

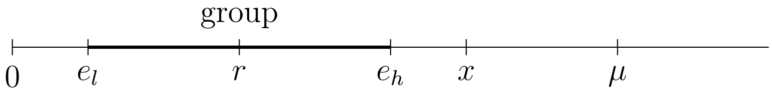
We compute the reelection probability if the policy is more moderate than the group's policy but more extreme than the most moderate group member, $r < x < e_h$.



The expected vote share of the group is given by $n\frac{\mu-x}{\mu-r}$. The vote share of the informed voters depends on whether the indifferent voter is more moderate than the most moderate group member.

$$p(x) = \begin{cases} n\frac{\mu-x}{\mu-r} + \alpha(\frac{\mu+x}{2} - n) + \frac{1}{2}(1-n)(1-\alpha) \text{ if } x + \frac{\mu-x}{2} > e_h \\ n\frac{\mu-x}{\mu-r} + \alpha e_l + \frac{1}{2}(1-n)(1-\alpha) \text{ else} \end{cases}$$

Lastly, we compute the reelection probability if the policy choice is more moderate than the most moderate group member, $e_h < x$.



The vote share of the group is given by $n\frac{\mu-r}{\mu-x}$ and the vote share of the informed voters is given by $\alpha(\frac{\mu+x}{2}-n)$. Then, the reelection probability is given by

$$p(x) = n\frac{\mu-x}{\mu-r} + \alpha\left(\frac{\mu+x}{2}-n\right) + \frac{1}{2}(1-n)(1-\alpha).$$

2.2 Exogenous groups

In this section we investigate the policy choice of an office-motivated incumbent and the welfare implications if there is one exogenously given interest group. We will see that the politician implements the group's policy position if the group is sufficiently large. Expected welfare decreases the higher the share of office-seeking politicians.

Suppose that there is an exogenously given interest group on the left hand side of the median, $r < \mu$. Even though the incumbent could implement any policy, he chooses to either implement the median policy or the group's policy. Which of these options is chosen depends on the size of the group.

If the politician chooses the median policy, all unorganized voters, both informed and uninformed, randomize between voting for the incumbent and the challenger.

$$p(\mu) = \frac{1}{2}\alpha(1-n) + \frac{1}{2}(1-\alpha)(1-n)$$

If the politician chooses the group's policy, the probability that he is reelected is given by

$$p(r) = \begin{cases} n + \alpha(\frac{\mu+r}{2} - n) + \frac{1}{2}(1-\alpha)(1-n) \text{ if } \frac{r+\mu}{2} \geq e_h \\ n + \alpha e_l + \frac{1}{2}(1-\alpha)(1-n) \text{ if } \frac{r+\mu}{2} < e_h \end{cases}$$

A politician maximizing the probability of reelection chooses the group's policy if the group is sufficiently large.

Proposition 1. *If one group with size n and position $r < \mu$ exists, the office-motived politician chooses the group's policy, $x = r$, if the group is sufficiently large, $n \geq \frac{\alpha}{2-\alpha}(\mu - r)$.*

Proof. See Appendix. □

The sole consideration of the office-seeking politician is to accumulate the largest number of votes possible. The many votes of a large interest group compensate for the lost votes from informed votes if the politician chooses the group's policy. If the group becomes more moderate, the politician can collect more votes from the informed unorganized votes if choosing the group's policy. Therefore, the size requirement for the group decreases if the group becomes more moderate.[5]

$$\frac{\partial n}{\partial r} = -\frac{\alpha}{2-\alpha} < 0$$

Lastly, we look at the effect on welfare induced by a change in the composition of politicians. One possible measure is aggregating the expected utility of each voter.

$$W = -\int_0^1 (\mathbb{E}x - r_i)^2 f(r) dr$$

where $\mathbb{E}x = \beta r + (1-\beta)\mu$. Intuitively, the higher the share of office-seeking politicians the more likely the group's policy is implemented. This, in turn, reduces welfare because implementing the median policy maximizes welfare. Similarly, the closer the group's position to the median policy, the higher is expected welfare.

Proposition 2. *The higher the share of the office-seeking politicians and the more extreme the existing group, the lower is expected welfare.*

Proof. Inserting expected policy $\mathbb{E}x$ gives $W = -\int_0^1 (\beta r + (1-\beta)\mu - r)^2 f(r) dr$. Re-arranging results in $W = -\left(\frac{1}{3} + r(r-1)\beta^2 + \mu(1-\beta)(\mu(1-\beta)-1)\right)$. Then, $\frac{\partial W}{\partial r} = -\beta^2(r-\mu) > 0$ and $\frac{\partial W}{\partial \beta} = -\beta(r-\mu)^2 < 0$. □

If there is one endogenously given group, an office-seeking politician chooses the group's policy instead of the median policy if the group is sufficiently large. This decreases expected welfare.

2.3 Endogenous group formation

In this section we analyze where groups form endogenously. We first show that in an equilibrium with group formation more than one group enters. We then derive the position of groups in an equilibrium with two groups.

A voter founds a group if the policy benefit from doing so is higher than the founding cost c. If there is no group, both types of politicians implement the policy preferred by the median voter, $\mathbb{E}x = \mu$. If the voter founds a group, the office-seeking politician implements the group's policy position provided that the group is sufficiently large. A group will therefore only be formed if the necessary size is reached. If the group is sufficiently large to influence the office-seeking politician's policy choice, the expected policy is given by $\mathbb{E}x = \beta r + (1-\beta)\mu$. If no group

[5] As an example consider $\alpha = 1/2$. The minimum required group size to implement the group's policy is $n = 1/3$ if $r = 0$ and $n = 1/8$ if $r = 1/8$.

already exists, voter i founds a group if

$$-(\mathbb{E}x - r_i)^2 - c \geq -(\mu - r_i)^2. \tag{4}$$

The benefit of starting a group is an expected policy closer to the founder's position. Because the median policy is chosen if there is no group, the group formation problem is symmetric around the median. Therefore, voters equidistant to the median have the same incentives to start a group.

Proposition 3. *If there exists an equilibrium with group formation, more than one group enters.*

Proof. Inserting $\mathbb{E}x = \beta r + (1-\beta)\mu$ in equation (4), we get that $(\mu - r)^2 - (\beta r + (1-\beta)\mu - r)^2 - c \geq 0$. Rewriting gives that $-\beta(\beta r^2 + 2r\mu - 2\beta r\mu - 2\mu^2 + \beta\mu^2 - 2r^2 + 2\mu r) \geq c$ and finally,

$$r = \mu \pm \sqrt{\frac{c}{\beta(2-\beta)}}.$$

□

After establishing that there will be at least two groups in an equilibrium with group formation, we investigate an equilibrium where exactly two groups form.

An equilibrium is characterized by three things. First, the founders need to enter and the members to join. Second, the groups need to reach the necessary size to influence the office-seeking politician. Third, no voter besides the founders wants to start a group neither in addition nor instead of the existing equilibrium groups.

We first investigate which voters start a group in equilibrium. Suppose there is one group with position $\tilde{r}$ that is sufficiently large to induce the office-seeking politician to implement the group's policy position. This results in an expected policy $\mathbb{E}x = \beta\tilde{r} + (1-\beta)\mu$. If voter i starts a group, expected policy changes to $\mathbb{E}x_i$. Voter i with $r_i < \mu$ starts a group if the expected policy gain is larger than the cost from founding.

$$-(\mathbb{E}x_i - r_i)^2 - c \geq -(\beta\tilde{r} + (1-\beta)\mu - r_i)^2 \tag{5}$$

Expected policy $\mathbb{E}x_i$ depends on which group's policy the office-seeking politician implements. If the politician prefers the group of voter $\tilde{r}$, $\mathbb{E}x_i = \beta\tilde{r} + (1-\beta)\mu$. If the politician is indifferent between both groups, $\mathbb{E}x_i = \beta(\frac{1}{2}\tilde{r} + \frac{1}{2}r_i) + (1-\beta)\mu$ and if the politician prefers the new group of voter r_i, $\mathbb{E}x_i = \beta r_i + (1-\beta)\mu$. Obviously, no voter founds a group if this has no effect on policy. In order to distinguish the second and the last case, we need to understand which group the politician prefers.

Remark 1. *For groups of the same size, the politician prefers the more moderate group. For groups with the same position, the politician prefers the larger group.*

Proof. If the politician chooses policy x, his utility is given by $U(r) = h + qhp(r)$. For $r < \mu$, if $\frac{\mu+r}{2} > e_h$, $p(r) = n + \alpha\left(\frac{\mu+r}{2} - n\right) + \frac{1}{2}(1-n)(1-\alpha)$ such that $U(r) = h + qh\left(n + \alpha\left(\frac{\mu+r}{2} - n\right)\right)$. Then, $\frac{\partial U(r)}{\partial r} = qh\alpha\frac{1}{2} > 0$ and $\frac{\partial U(r)}{\partial n} = qh(1-\alpha) > 0$.
If $\frac{\mu+r}{2} < e_h$, $p(r) = n + \alpha e_l + \frac{1}{2}(1-n)(1-\alpha)$: $\frac{\partial U(r)}{\partial r} = 0$ and $\frac{\partial U(r)}{\partial n} = qh > 0$. □

The more voters are organized in a group, the more votes the group can promise for choosing the group's policy. The politician therefore prefers the larger group if relative positions are the same. If the informed unorganized voter being indifferent between voting for the incumbent and the challenger is located outside of the group, implementing the policy of a more moderate group guarantees more votes from informed unorganized voters. If, however, the indifferent informed unorganized voter would be located inside of the group, a more moderate group policy does not immediately imply more votes from unorganized voters. The politician is therefore indifferent among group positions in this case. For simplicity, we assume that the politician prefers implementing a more moderate position also in this case.

The politician can pick the group that he prefers. We look for an equilibrium in pure strategies. Let us assume that there is one group on one side of the median. Then the voters on the other side of the median observe this position and the voter who is just a bit more moderate than the founder of the existing group starts a group. Because the new group is more moderate, this results in the new group chosen for sure by the politician.[6] But then the first founder would not want to start the first group. Therefore, in equilibrium the politician is indifferent between choosing either group and both groups have the same size and the same relative position, $|\mu - \bar{r}| = |\mu - \tilde{r}|$ and $\bar{n} = \tilde{n}$.

We compute the founders' positions. Because $\tilde{r} = 1 - r_i$, $\mathbb{E}x_i = \beta(\frac{1}{2}\tilde{r} + \frac{1}{2}r_i) + (1-\beta)\mu = \mu$ and the founding condition is given by

$$-(\mu - r_i)^2 - c \geq -(\beta\tilde{r} + (1-\beta)\mu - r_i)^2. \quad (6)$$

Because the policy benefit from moving the policy away from the median policy increases the more extreme the voter, also the benefit of founding a group is higher the more extreme the voter, $\frac{\partial(6)}{\partial r} = 2\beta(r - \mu) < 0$. Because the politician chooses the most moderate group for constant size, the most moderate voter still gaining

[6] The underlying assumption here is of course that groups do not decrease in size if the founder moves to the median.

from a group is the founder, $\bar{r} < \mu$. While voters with $r < \bar{r}$ profit more from a group, their group will never be chosen by the politician.

Remark 2. *If there is an equilibrium with group formation, founders are located at*

$$r_i = \mu \pm \sqrt{\frac{c}{\beta(2+\beta)}}.$$

Proof. Remembering that $\mu = \frac{1}{2}$ and that in equilibrium $\tilde{r} = 1 - r_i$ we can rewrite equation (6) as $\beta^2(1-r_i)^2 + r_i(-2\beta(1-r_i) - (1-\beta) - \beta(1-\beta)) + \beta(1-\beta) + \mu^2(-2\beta+\beta^2) + r_i = c$ and finally get the founders' positions. □

We next look at which voters join an existing group. Remembering that the cost of membership increases in the distance from the group's position, $k(r_i - \bar{r})^2$, while the benefits are constant, b, voter i joins the equilibrium group if

$$b \geq k(r_i - \bar{r})^2$$

Then, the voters indifferent between joining the group or staying outside are given by

$$r_i = \bar{r} \pm \sqrt{\frac{b}{k}}.$$

This implies that group size is more or less exogenously given. If the founder moves very close to the extreme, part of the size potential vanishes. The maximum attainable group size is $n_{max} = 2\sqrt{\frac{b}{k}}$. If the founder is very close to the extreme, the size of the group is $\sqrt{\frac{b}{k}} + \bar{r}$. Groups will only form if the attainable size is higher than the size necessary to influence the incumbent's policy choice.

We have found the positions of the equilibrium founders and the voters joining the resulting groups. The next step is to show that no other voters want to found a group. We do this by proving a series of claims.

If two groups are on one side of the median, with group 1 being the extreme group and group 2 being the moderate group, the politician chooses to implement the policy of the moderate group if this group is at least as large as the extreme group.

Claim 1. *If there are two groups on one side of the median and the politician chooses a group's policy, the politician chooses the moderate position if $n_2 \geq n_1$.*

Proof. See Appendix. □

We investigate whether voters with a more extreme policy position than the equilibrium founders want to found a third group, $r_i < \bar{r}$. If this is the case, there is no equilibrium with two groups in pure strategies. Because the equilibrium founder $\bar{r}$ is the most moderate voter still benefitting from a group and this benefit increases the more extreme the founder, voters with a more extreme policy position than the equilibrium founder would gain more from starting a group if their group's position was chosen. We have shown in claim 1 that, if there are two groups on the same side of the median, the politician chooses the position of the moderate group if this group is not smaller than the extreme group. Because in equilibrium the median voter is not part of any group, the extreme third group cannot be larger then the moderate equilibrium group. Therefore, an unorganized voter $r_i < \mu$ founds a third group if this induces the politician to choose the moderate group's policy, $\bar{r}$, on the same side of the median for sure instead of randomizing between the positions of the equilibrium groups.

$$-c - (\beta\bar{r} + (1-\beta)\mu - r_i)^2 \geq -(\mu - r_i)^2. \tag{7}$$

Intuitively, unorganized voters start a third group if founding costs are small.

Claim 2. *If the creation of a third extreme group ensures that the politician chooses the equilibrium group's position on the same side of the median, unorganized extreme voters do not start a third group if*

$$\frac{1}{4(1+\beta)^2} < \frac{c}{\beta(2+\beta)}.$$

Proof. Following from the derivative of equation (7) with respect to r_i, $\frac{\partial(7)}{\partial r_i} = 2\beta(\bar{r} - \mu) < 0$, the highest benefit of starting a group accrues to $r_i = 0$. It therefore suffices to check whether $r_i = 0$ wants to start group. Insert the equilibrium value $\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}}$ we rewrite equation (7) as $-c + \sqrt{\frac{\beta c}{2+\beta}} - \frac{c\beta}{2+\beta} \geq 0$. Further rewriting gives the condition in the claim. □

If the most extreme unorganized voter does not start a third group because the founding costs are too high, no other voter would want to found a group. This follows from the benefit of changing the policy away from the median being smaller the more moderate the voter. Additionally, voters joining the group in equilibrium would also forego the benefit of being a group member. A voter prefers to join a group instead of founding one if

$$b - k(r_i - \bar{r})^2 \geq -c + (\mu - r_i)^2 - (\beta\bar{r} + (1-\beta)\mu - r_i)^2.$$

Because a voter only joins a group if the benefit of doing so is positive, a group member has a smaller incentive than an unorganized voter to found a group. Excluding that the most extreme unorganized voter founds a group therefore excludes that any voter more extreme than the equilibrium founder starts a third group.

We next exclude that voters more moderate than the equilibrium founder start a group. If a voter more moderate than the equilibrium founder enters, the politician chooses that position for sure. If this voter does not enter, expected policy is μ and if she does, expected policy is $\mathbb{E}x = \beta r + (1-\beta)\mu$. This voter enters if

$$-(\mathbb{E}x - r_i)^2 - c \geq -(\mu - r_i)^2.$$

This is equation (4) and we have already computed the solution, $r = \mu - \sqrt{\frac{c}{\beta(2-\beta)}}$. Now we compare this to the equilibrium founder, $\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}}$. We find that $r < \bar{r}$. Therefore, a voter with $r_i > \bar{r}$ never wants to start a group.

In the derivation of the claim above we have assumed that the entry of a third extreme group results in the politician choosing the moderate equilibrium group's position on the same side. This does not have to be the case. The next step is therefore to investigate how the entry of a third group influences the reelection probabilities associated with the two equilibrium groups.

In the following we assume that the benefit of being a group member is large.

Assumption 1. $\min\left\{\sqrt{\frac{b}{k}}, \bar{r}\right\} = \bar{r}$

This implies that the most extreme voters are members of the equilibrium groups. This assumption also implies that entry of a third extreme group does not lead to more voters organized in groups and instead only diverts already existing organization. There are thus no unorganized voters between the extreme and moderate group.

Claim 3. *Under assumption 1, the entry of a third group that is more extreme than the equilibrium groups induces the politician to choose the position of the equilibrium group on the other side of the median.*

Proof. See Appendix. □

Because the benefit of being organized is high, the most extreme voters are members of the equilibrium groups. Therefore, founding an extreme group only diverts existing organization but does not create new organization of voters. Because some voters join the more extreme group instead of the moderate equilibrium group and vote accordingly, the politician receives fewer votes when choosing the moderate group's position. Because there is a second equilibrium group that still has the full size and therefore promises more votes, the politician chooses the equilibrium group on the other side of the median if an extreme group is founded on one side of the median.

If the politician chooses a group's policy instead of the median policy, he loses votes from the informed voters. Therefore, groups need to be sufficiently large where the size required by the politician depends on the share of informed voters, α. Because of assumption 1, we can express the required minimum group size in terms of the founders' positions.

Claim 4. *Minimum equilibrium group size is realized if*

$$\bar{r} \geq \begin{cases} \frac{2}{2+\alpha}\left(\frac{\alpha}{4} - \sqrt{\frac{b}{k}}\right) & \text{if } \sqrt{\frac{b}{k}} < \frac{1}{2}\sqrt{\frac{c}{\beta(2+\beta)}} \\ \frac{\alpha}{2+\alpha} - \sqrt{\frac{b}{k}} & \text{else} \end{cases}$$

Proof. See Appendix. □

If the median voter wants to join a group in equilibrium, then, because of the underlying symmetry, that voter wants to join both groups. Therefore, one group would be bigger than the other which cannot be an equilibrium. In equilibrium, the median voter does not join a group, $\bar{r} + \sqrt{\frac{b}{k}} < \mu$. The corresponding condition is given by

$$\sqrt{\frac{b}{k}} < \sqrt{\frac{c}{\beta(2+\beta)}}.$$

If the equilibrium value for the founder $\bar{r}$ is negative, there is no group formation because every voter is more moderate than the equilibrium founder. Thus, there can only be group formation if

$$\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}} \geq 0.$$

In a symmetric equilibrium with two groups, a number of conditions need to be met. The median voter does not join a group ($\sqrt{\frac{b}{k}} < \sqrt{\frac{c}{\beta(2+\beta)}}$), founders exist ($\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}} \geq 0$) and assumption 1 holds ($\mu - \sqrt{\frac{c}{\beta(2+\beta)}} < \sqrt{\frac{b}{k}}$).

Proposition 4. *In an equilibrium with two groups, the groups are positioned at*

$$\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}} \text{ and } \tilde{r} = \mu + \sqrt{\frac{c}{\beta(2+\beta)}}$$

and have size $\bar{n} = \tilde{n} = \bar{r} + \sqrt{\frac{b}{k}}$. *This equilibrium exists if* $\mu - \sqrt{\frac{c}{\beta(2+\beta)}} < \sqrt{\frac{b}{k}} <$

$\sqrt{\frac{c}{\beta(2+\beta)}} < \mu$ *and*

$$\bar{r} \geq \begin{cases} \frac{2}{2+\alpha}\left(\frac{\alpha}{4} - \sqrt{\frac{b}{k}}\right) & if \ \sqrt{\frac{b}{k}} < \frac{1}{2}\sqrt{\frac{c}{\beta(2+\beta)}} \\ \frac{\alpha}{2+\alpha} - \sqrt{\frac{b}{k}} & else \end{cases}$$

This is not necessarily the only equilibrium with two groups. As long as a third group leads to more diversion of organization than to new creation, the politician chooses the equilibrium group on the other sound. Similarly, if founding costs are high, extreme voters do not start a group even if the politician chooses the policy of the equilibrium group close to them.

Because there are two groups in equilibrium, expected policy continues to be the median policy if there is group formation. We can, however, look at the equilibrium group position, $\bar{r} = \mu - \sqrt{\frac{c}{\beta(2+\beta)}}$. An increase in the share of office-seeking politicians makes the groups' positions more moderate.[7] Conditional on a group's policy being chosen, welfare therefore increases in the share of office-seeking politicians.

The following example shows how the position and size of groups changes when the share of office-seeking politicians varies.

Example 1

The parameter values are $c = 0.3$, $b = 0.01$, $k = 0.25$ and $\alpha = 0.5$. If the share of office-seeking politicians is small, $\beta = 0.25$, there are no groups. If the share of office-seeking politicians is intermediate, $\beta = 0.5$, groups have extreme positions, $\bar{r} = 0.01$, and an intermediate size, $\bar{n} = 0.21$. If there are many office-seeking politicians, $\beta = 0.75$, groups have a more moderate position, $\bar{r} = 0.12$, and are large, $\bar{n} = 0.32$.

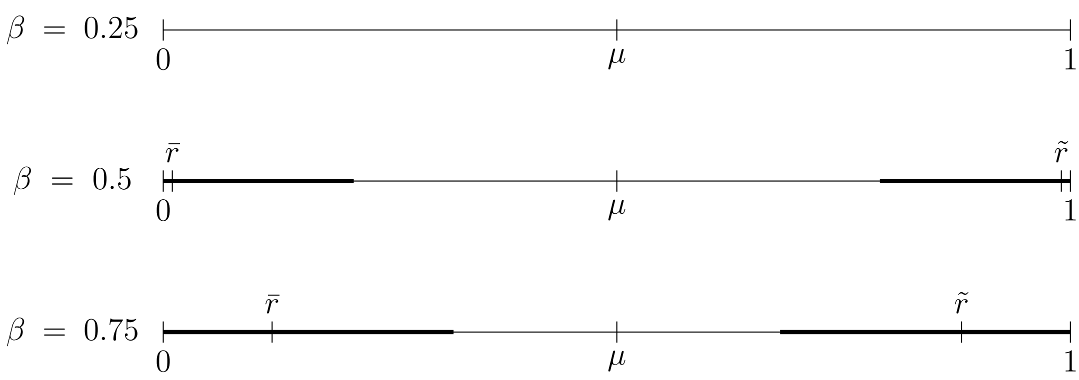


Figure 1: Equilibria for different values of β

[7] $\frac{\partial \bar{r}}{\partial \beta} = (1+\beta)\sqrt{\frac{c}{(\beta(1+\beta))^3}} > 0$

3 Conclusion

This paper investigates first how voters organized in interest groups can influence a politician's policy choice by bundling their votes and second what size and position endogenously formed groups have. Voters organized in interest groups are assumed to have an informational advantage over unorganized voters and can use their better information for punishing the incumbent politician more effectively whenever he chooses a policy that they disapprove of. This punishment takes the form of not voting for the incumbent politician and instead for his challenger.

We have seen that an office-seeking politician implements a group's policy if the group is sufficiently large. Because the politician foregoes votes from the informed unorganized voters when choosing the group's position, the politician prefers groups to be large and moderate. Therefore, in equilibrium the most moderate voter still gaining from founding a group is the equilibrium voter and determines the group's position. The founder's position becomes more extreme the higher the cost of founding and the smaller the share of office-seeking politicians.

Some extensions seem to be worth pursuing. Firstly, we have restricted the analysis to an equilibrium with two groups. It would be interesting to see whether an equilibrium with four groups, i.e. both an extreme and a moderate group on each side of the median, exists and how this equilibrium would look like. This could be done by relaxing the assumption that the members' benefit from joining the group is high. Secondly, the analysis assumes that maximum group size is determined by exogenously given parameters. While the willingness to compromise on ones own position certainly is exogenous to some extent, where examples include environmental or animal rights group occupying positions that range from extreme to moderate, changing the attractiveness of joining a group by providing higher benefits could also influence group size. Endogenizing group size in addition to position, however, allows extreme founders to compensate for extreme positions with a larger group size such that the politician is indifferent between moderate small groups and extreme large groups. An equilibrium in this case will therefore not be as clearcut as in the present analysis where the most moderate voter who is sufficiently extreme to gain from a change in policy is the founder of an interest group.

References

Alesina, A., Cukierman, A. (1990). The politics of ambiguity, *The Quarterly Journal of Economics 105*(4), 829-850.

Anderson, S. P., Glomm, G. (1992). Incumbency effects in political campaigns, *Public Choice 74*(2), 207-219.

Baron, D. P. (1994). Electoral competition with informed and uninformed voters, *The American Political Science Review 68*(1), 33-47.

Becker, G. S. (1983). The theory of competition among pressure groups for political influence, *The Quarterly Journal of Economics 98*(3), 371-400.

Bernhardt, D., Câmara, O., Squintani, F. (2011). Competence and ideology, *The Review of Economic Studies 78*(2), 487-522.

Brekke, K. A., Nyborg, K., Rege, M. (2007). The fear of exclusion: Individual effort when group formation is endogenous, *The Scandinavian Journal of Economics 109*(3), 531-550.

Denzau, A. T., Munger, M. C. (1986). Legislators and interest groups: How unorganized interest get represented, *The American Political Science Review 80*(1), 89-106.

Grossman, G. M., Helpman, E. (1996). Electoral competition and special interest politics, *The Review of Economics Studies 63*(2), 265-286.

Hamlin, A., Jennings, C. (2004). Group formation and political conflict: Instrumental and expressive approaches, *Public Choice 118*(3/4), 413-435.

Kartik, N., McAfee, R. P. (2007). Signaling character in electoral competition, *The American Economic Review 97*(3), 852-870.

Maskin, E., Tirole, J. (2004). The politician and the judge: Accountability in government, *The American Economic Review 94*(4), 1034-1054.

Peltzman, S. (1976). Toward a more general theory of regulation, *Journal of Law and Economics 19*(2), 211-240.

Plotnick, R. D. (1986). An interest group model of direct income redistribution, *The Review of Economics and Statistics 68*(4), 594-602.

Potters, J., Sloof, R. (1996). Interest groups: A survey of empirical models that try to assess their influence, *European Journal of Political Economy 12*, 403-442.

Schneider, M. T. (2014). Interest-group size and legislative lobbying, *Journal of Economic Behavior & Organization 106*, 29-41.

Snyder Jr., J. M., Ting, M. M. (2008). Interest groups and the electoral control of politicians, *Journal of Public Economics 92*, 482-500.

Stigler, G. J. (1971). The theory of economic regulation, *The Bell Journal of Economics and Management Science 2*(1), 3-21.

A Proofs

A.1 Proof of Proposition 1

Proof. First, we evaluate the change in reelection probabilities for different values of x. If $x < e_l$, $p^{'}(x) > 0$ for all cases and the politician increases x to e_l.

$$p^{'}(x) = \begin{cases} n\frac{\mu-r}{(\mu-x)^2} + \frac{\alpha}{2} \text{ if } x + \frac{\mu-x}{2} > e_h \\ n\frac{\mu-r}{(\mu-x)^2} \text{ if } e_h > x + \frac{\mu-x}{2} > e_l \\ n\frac{\mu-r}{(\mu-x)^2} + \frac{\alpha}{2} \text{ if } x + \frac{\mu-x}{2} < e_l \end{cases}$$

If $e_l < x < r$, $p^{'}(x) > 0$ for all cases and the politician increases x to r.

$$p^{'}(x) = \begin{cases} n\frac{\mu-r}{(\mu-x)^2} + \frac{\alpha}{2} \text{ if } x + \frac{\mu-x}{2} > e_h \\ n\frac{\mu-r}{(\mu-x)^2} \text{ if } x + \frac{\mu-x}{2} < e_h \end{cases}$$

If $r < x < e_h$:

$$p^{'}(x) = \begin{cases} -\frac{n}{\mu-r} + \frac{\alpha}{2} \text{ if } x + \frac{\mu-x}{2} > e_h \\ -\frac{n}{\mu-r} < 0 \text{ if } x + \frac{\mu-x}{2} < e_h \end{cases}$$

If $x + \frac{\mu-x}{2} > e_h$, the politician reduces x to r if $\frac{\alpha}{2}(\mu - r) < n$. If $x + \frac{\mu-x}{2} < e_h$, the politician always reduces x to r.

If $e_h < x$, the politician reduces x to r if $\frac{\alpha}{2}(\mu - r) < n$.

$$p^{'}(x) = -\frac{n}{\mu - r} + \frac{\alpha}{2}$$

The politician chooses either $x = r$ or $x = \mu$.

The politician still has to compare the reelection probabilities of choosing $x = r$ and $x = \mu$.

If $\frac{r+\mu}{2} > e_h$, $p(r) = n + \alpha(\frac{\mu+r}{2} - n)$: $p(r) \geq p(\mu)$ if $n + \alpha(\frac{\mu+r}{2} - n) \geq \frac{1}{2}\alpha(1 - n)$. Rewriting gives $n \geq \alpha(\frac{\mu+n-r}{2})$ and finally

$$n \geq \frac{\alpha}{2 - \alpha}(\mu - r)$$

Considering the condition, we derived previously for this case, $\frac{\alpha}{2}(\mu - r) < n$, we can show that

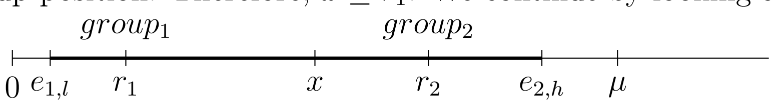
$$\frac{\alpha}{2 - \alpha}(\mu - r) < \frac{\alpha}{2}(\mu - r)$$

reduces to $0 < -\alpha$ and therefore, we need that $n \geq \frac{\alpha}{2-\alpha}(\mu - r)$.

If $\frac{r+\mu}{2} < e_h$, $p(r) = n + \alpha e_l$: $p(r) \geq p(\mu)$ if $n + \alpha e_l \geq \frac{1}{2}\alpha(1-n)$ or $n \geq \frac{\alpha}{2+\alpha}(1-2e_l)$. Note that we can insert $e_l = e_h - n$ in $n + \alpha e_l \geq \frac{\alpha}{2} - \frac{\alpha n}{2}$ and get $n(1-\frac{\alpha}{2}) \geq \frac{\alpha}{2} - \alpha e_h$. From $e_h > \frac{\mu+r}{2}$ follows that $\frac{\alpha}{2} - \alpha e_h < \frac{\alpha}{2} - \frac{\mu+r}{2}$. Therefore, if $n(1-\frac{\alpha}{2}) \geq \frac{\alpha}{2} - \frac{\mu+r}{2}$ so is $n(1-\frac{\alpha}{2}) \geq \frac{\alpha}{2} - \alpha e_h$. The former condition can be rewritten as $n \geq \frac{\alpha}{2-\alpha}(\mu - r)$ □

A.2 Proof of Claim 1

Proof. We first determine the reelection probabilities of the incumbent for different policy choices if there are two groups. Depicted below is the more extreme group, group 1, and the more moderate group, group 2. We assume that there are no unorganized voters between groups such that $e_1^h = e_2^l$. Similar to the case with only one group, the politician never chooses a policy that is more extreme than the most extreme group position. Therefore, $x \geq r_1$. We continue by looking at $r_1 < x < r_2$:



Expected votes from the two groups are $n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-r_2}{\mu-x}$. Expected votes from the informed voters are αe_1^l if $x + \frac{\mu-x}{2} < e_2^h$ and $\alpha(x + \frac{\mu-x}{2} - n_1 - n_2)$ if $x + \frac{\mu-x}{2} > e_2^h$.

If $x + \frac{\mu-x}{2} < e_2^h$, $p(x) = n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-r_2}{\mu-x} + \alpha e_1^l$ such that $p^{'}(x) = -\frac{n_1}{\mu-r_1} + n_2\frac{\mu-r_2}{(\mu-x)^2}$. Because $p^{''}(x) = \frac{2n_2(\mu-r_2)}{(\mu-x)^3} > 0$, x is a local minimum and the politician chooses either r_1 or r_2.

If $x + \frac{\mu-x}{2} > e_2^h$, $p(x) = n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-r_2}{\mu-x} + \alpha(x + \frac{\mu-x}{2} - n_1 - n_2)$ such that $p^{'}(x) = -\frac{n_1}{\mu-r_1} + n_2\frac{\mu-r_2}{(\mu-x)^2} + \frac{\alpha}{2}$. Because $p^{''}(x) = \frac{2n_2(\mu-r_2)}{(\mu-x)^3} > 0$, x is a local minimum and the politician chooses either r_1 or r_2.

Now assume that $r_2 < x$: Expected votes from the two groups are $n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-x}{\mu-r_2}$. Expected votes from the informed voters are αe_1^l if $x + \frac{\mu-x}{2} < e_2^h$ and $\alpha(x + \frac{\mu-x}{2} - n_1 - n_2)$ if $x + \frac{\mu-x}{2} > e_2^h$.

If $x + \frac{\mu-x}{2} < e_2^h$, $p(x) = n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-x}{\mu-r_2} + \alpha e_1^l$ and $p^{'}(x) = -\frac{n_1}{\mu-r_1} - \frac{n_2}{\mu-r_2} < 0$. The optimal choice is to decrease x to r_2.

If $x + \frac{\mu-x}{2} > e_2^h$, $p(x) = n_1\frac{\mu-x}{\mu-r_1} + n_2\frac{\mu-x}{\mu-r_2} + \alpha(x + \frac{\mu-x}{2} - n_1 - n_2)$ and $p^{'}(x) = -\frac{n_1}{\mu-r_1} - \frac{n_2}{\mu-r_2} + \frac{\alpha}{2}$. The optimal choice is to reduce x to r_2 if $\frac{\alpha}{2} < \frac{n_1}{\mu-r_1} + \frac{n_2}{\mu-r_2}$.

Hence, the politician chooses either r_1, r_2 or μ. It remains to show that the politician chooses r_2 if he does not choose μ.

If $x + \frac{\mu-x}{2} < e_2^h$:

$$p(r_1) = n_1 + n_2\frac{\mu - r_2}{\mu - r_1} + \alpha e_1^l$$

$$p(r_2) = n_1\frac{\mu - r_2}{\mu - r_1} + n_2 + \alpha e_1^l$$

$p(r_2) \geq p(r_1)$ if

$$0 \geq (n_1 - n_2)\left(\frac{r_2 - r_1}{\mu - r_1}\right)$$

If $x + \frac{\mu - x}{2} > e_2^h$

$$p(r_1) = n_1 + n_2\frac{\mu - r_2}{\mu - r_1} + \alpha\left(r_1 + \frac{\mu - r_1}{2} - n_1 - n_2\right)$$
$$p(r_2) = n_1\frac{\mu - r_2}{\mu - r_1} + n_2 + \alpha\left(r_2 + \frac{\mu - r_2}{2} - n_1 - n_2\right)$$

$p(r_2) \geq p(r_1)$ if

$$0 \geq (n_1 - n_2)\left(\frac{r_2 - r_1}{\mu - r_1}\right) + \frac{\alpha}{2}(r_1 - r_2)$$

□

A.3 Proof of Claim 3

Proof. The equilibrium group positions are $\bar{r} < \mu < \tilde{r}$. A more extreme voter than the equilibrium founder starts a group with position $r_1 < r_2 = \bar{r}$. We have already shown that the politician never chooses r_1. The reelection probabilities are given by

$$p(\tilde{r}) = \tilde{n} + \alpha\left(\frac{3\mu - \tilde{r}}{2} - \tilde{n}\right)$$
$$p(r_2) = n_1\frac{\mu - r_2}{\mu - r_1} + n_2 + \alpha\left(\frac{\mu + r_2}{2} - n_1 - n_2\right)$$

with $\frac{\mu + r_2}{2} = \frac{3\mu - \tilde{r}}{2}$. Then, $p(\tilde{r}) > p(r_2)$ if

$$\tilde{n} + \alpha\left(\frac{\mu + r_2}{2} - \tilde{n}\right) \geq n_1\frac{\mu - r_2}{\mu - r_1} + n_2 + \alpha\left(\frac{\mu + r_2}{2} - n_1 - n_2\right)$$

This can be rewritten as

$$\tilde{n}(1 - \alpha) \geq n_1\left(\frac{\mu - r_2}{\mu - r_1} - \alpha\right) + n_2(1 - \alpha)$$

Because there are no unorganized voters between group 1 and group 2, the voter indifferent between joining group 1 and 2 is located at $r_1 + \frac{r_2 - r_1}{2}$. Group sizes are given by $\tilde{n} = \bar{n} = \sqrt{\frac{b}{k}} + \min\left\{\sqrt{\frac{b}{k}}, \bar{r}\right\}$, $n_1 = \frac{r_2 - r_1}{2} + \min\left\{r_1, \sqrt{\frac{b}{k}}\right\}$, $n_2 = \sqrt{\frac{b}{k}} + \min\left\{\sqrt{\frac{b}{k}}, \frac{r_2 - r_1}{2}\right\}$. Imposing assumption 1, $\min\left\{\sqrt{\frac{b}{k}}, \bar{r}\right\} = \bar{r}$, $\min\left\{\sqrt{\frac{b}{k}}, \frac{r_2 - r_1}{2}\right\} = \frac{r_2 - r_1}{2}$ and $\min\left\{r_1, \sqrt{\frac{b}{k}}\right\} = r_1$ such that $\bar{n} = \sqrt{\frac{b}{k}} + \bar{r}$, $n_1 = \frac{r_2 - r_1}{2} + r_1$, $n_2 = \sqrt{\frac{b}{k}} + \frac{r_2 - r_1}{2}$.

Then we insert $n_2 = \bar{n} - \frac{r_1+r_2}{2}$ and $\tilde{n} = \bar{n}$ in the equation above to get

$$\frac{r_1 + r_2}{2}(1-\alpha) \geq n_1 \left(\frac{\mu - r_2}{\mu - r_1} - \alpha \right).$$

Lastly, we insert that $n_1 = \frac{r_1+r_2}{2}$ and get that

$$r_2 \geq r_1.$$

This is always true. □

A.4 Proof of Claim 4

Proof. There are two possible reelection probabilities if $x = \bar{r}$ and one if $x = \mu$.

$$p(\bar{r}) = \begin{cases} n + \alpha \left(\frac{\mu+\bar{r}}{2} - n \right) + \frac{1}{2}(1-\alpha)(1-2n) \text{ if } \frac{\bar{r}+\mu}{2} \geq e_h \\ n + \frac{1}{2}(1-\alpha)(1-2n) \text{ if } \frac{\bar{r}+\mu}{2} < e_h \end{cases}$$
$$p(\mu) = \frac{1}{2}\alpha(1-2n) + \frac{1}{2}(1-\alpha)(1-2n)$$

Determine which of the two probabilities $p(\bar{r})$ is relevant: $e_h = \bar{r} + \sqrt{\frac{b}{k}} < \frac{\mu+\bar{r}}{2}$ can be rewritten as

$$\sqrt{\frac{b}{k}} < \frac{1}{2}\sqrt{\frac{c}{\beta(2+\beta)}}$$

The politician chooses $x = \bar{r}$ if $p(\bar{r}) \geq p(\mu)$:

If $\frac{\bar{r}+\mu}{2} \geq e_h$, $p(\bar{r}) \geq p(\mu)$ if $n + \alpha\left(\frac{\mu+\bar{r}}{2} - n\right) \geq \frac{1}{2}\alpha(1-2n)$. Rewriting gives $n \geq \frac{\alpha}{2}(1 - \frac{\mu+\bar{r}}{2})$. Insert $n = \bar{r} + \sqrt{\frac{b}{k}}$ to get $\bar{r} \geq \frac{2}{2+\alpha}\left(\frac{\alpha}{4} - \sqrt{\frac{b}{k}}\right)$.

If $\frac{\bar{r}+\mu}{2} < e_h$, $p(\bar{r}) \geq p(\mu)$ if $n \geq \frac{1}{2}\alpha(1-2n)$. Rewriting gives $n \geq \frac{\alpha}{2(1+\alpha)}$. Insert $n = \bar{r} + \sqrt{\frac{b}{k}}$ to get $\bar{r} \geq \frac{\alpha}{2+\alpha} - \sqrt{\frac{b}{k}}$. □