The economics of repeated extortion

Jay Pil Choi*

Marcel Thum**

This paper provides a simple model of repeated extortion. In particular, we ask whether corrupt government officials’ opportunism to demand more once entrepreneurs have made sunk investments entails further distortion in resource allocations. If the choice of technology is left to the entrepreneurs, the dynamic path of demand schedules will induce entrepreneurs to pursue a “fly-by-night” strategy by adopting a technology with an inefficiently low sunk cost component. The unique equilibrium is characterized by a mixed strategy of the government official in future demand. Our model thus explains why arbitrariness is such a central feature of extortion. We also investigate implications of the stability of corrupt regimes for dynamic extortion and discuss alternative applications for our framework.

* Michigan State University & CESifo; Department of Economics, Michigan State University, 101 Marshall Hall, East Lansing, MI 48824, U.S.A., e-mail: choijay@msu.edu

** Dresden University of Technology, ifo Dresden & CESifo; Department of Economics, Dresden University of Technology, 01062 Dresden, Germany, e-mail: marcel.thum@mailbox.tu-dresden.de

This research was begun while the first author was visiting the Center for Economic Studies, University of Munich. We are grateful to Drew Fudenberg, Eric Maskin, John McLaren, Klaus Schmidt, Ken Sennewald, Rafael Di Tella for helpful discussions and Marcus Dittrich, Uwe Kratsch and Gunther Markwardt for careful research assistance. We also thank Editor Raymond Deneckere and two anonymous referees for constructive comments that significantly improved the paper.
1 Introduction

Corruption is deemed by many economists to be detrimental to investment incentives leading to lower economic growth. Government officials, for instance, charge entrepreneurs for permits and licenses that they require to operate a business. These licenses, sold to maximize the officials’ private gains, constitute an additional burden on new businesses (Shleifer and Vishny, 1993). Despite some arguments for beneficial effects of corruption (Leff, 1964; Huntington, 1968), recent empirical findings support the view that corruption is harmful to investment activities. Mauro (1995) shows that corruption lowers private investment, thereby reducing economic growth. In a study of the effect of corruption on foreign direct investment (FDI), Wei (2000) finds that a rise in the corruption level in a host country reduces the inflow of FDI. In particular, an increase in the corruption level from that of Singapore to that of Mexico is shown to be equivalent to an increase in the tax rate of more than 20 percentage points.\(^1\)

This paper is concerned with the dynamics of corruption. More specifically, we ask whether the government officials’ *ex post* opportunism to demand more once entrepreneurs have made sunk investments entails further distortion in resource allocations. To answer this question, we develop a simple model of *repeated* extortion. As in Shleifer and Vishny (1993), we consider the sale of

---

\(^1\) Corruption is not an activity that is confined to any particular corner of the world. Even though there are large differences in the level of corruption between countries, some type of corruption can be found in almost every country. Even for the US which is usually ranked highly in terms of efficient, non-corrupt bureaucracies, Fesler and Kettl (1991) report that “in an FBI ‘sting’ operation, 105 out of 106 offers of bribes to suspected municipal officials in the State of New York were accepted; the 106\(^{th}\) was rejected as too small”. At the other end of the corruption scale are countries like Uganda and Zaire (now Democratic Republic of the Congo). Under Idi Amin’s regime in Uganda, for instance, government became little more than a system of organized crime used to extract rents from the public (World Bank, 1997). See Gould (1980) for a detailed analysis of corruption in Zaire.
government property by government officials as the prototype of extortion activities.\textsuperscript{2} We show that the inability of government officials to commit to future demands does not distort entry decisions any further if the choice of technology is not a decision variable for the entrepreneurs. The government official can properly discount the initial demand in order to induce the appropriate amount of entry. If, however, the choice of technology is left to the entrepreneurs, the dynamic path of demand schedules will induce entrepreneurs to adopt an inefficient "fly-by-night" strategy. They will choose a technology with inefficiently low sunk cost component, which allows them to react more flexibly to future demands from corrupt officials. We characterize the equilibrium behavior of the government officials and the entrepreneurs’ technology choices. In particular, we show that there is no pure strategy equilibrium. Once entry decisions are made by entrepreneurs, the government officials’ optimal strategy is to demand varying amounts of money.

This paper thus provides a new interpretation of the \textit{arbitrariness} that entrepreneurs often face in a corrupt environment; uncertainty is simply an equilibrium property of repeated extortion. There is plenty of evidence that corrupt environments are very susceptible to arbitrariness and unpredictability. For instance, Scott (1972, p. 83) reports on corruption in Indonesia: "(...) in Indonesia the corruption ‘market’ was so disorganized that ‘prices’ were highly unstable and ‘delivery’ by sellers was highly uncertain." According to Robert Klitgaard (1990, pp. 94-95) the price for public services in Equatorial Guinea varied a great deal: "[t]here are no electricity meters so one might be asked to pay a million and a half cefas a month even though one had the generator going for twenty days." Beyond the anecdotal evidence, a recent study by Jakob Svensson (2003) also provides hard evidence using a survey among 250 Ugandan firms which was initiated by the

\textsuperscript{2}Thus, we shall use corruption and extortion activities interchangeably. In doing so, we adopt the "classic" definition of corruption as "use of public office for private gains." See Weinschelbaum (1998) for a different interpretation of corruption as an agency problem where the agent’s incentive to take an action against the principal is motivated by a third player, the "hidden principal."
World Bank and the Uganda Private Sector Foundation. The data set is unique as it contains quantitative information on bribe payments. The respondents were asked how much they had to pay to get access to the public grid and to acquire a telephone line. From a comprehensive analysis of the data, Svensson concludes that "there is considerable variation in reported graft across firms facing similar institutions/policies." For instance, there are large differences even among the firms reporting low bribe payments: the median bribe was $180 while the mean amounted to $280 (with a standard deviation of $280).\(^3\) Despite such extensive evidence and informal commentary emphasizing the arbitrariness and unpredictability of bribery demands as part of the corruption problem, this paper is, to our best knowledge, the first one to demonstrate such a phenomenon as a consequence of the theory.

The repeated demands in extortion are well-documented (see, for instance, John T. Noonan’s (1984) comprehensive study on bribes). How extortion can almost become an art is illustrated by the case of cardinal Tommaso di Capua who became head of the Pope’s Chancery under Innocent III in 1215. His letters written between 1215 and 1239 were even published a few years later in *Summa Dictaminis* - a handbook for correspondence in connection with ‘presents’. Some gifts were sent back: "We have just become friends, don’t rush it." Other letters give an eloquent but clear signal that additional payments are expected: "Be prepared to come back later (...) with fatter recompense (*retributio*). Therefore let your ready hand not grow lukewarm in the future nor put obstacle to later payments so that with repeated benefits you make your friends more devoted” (Noonan, 1984, p. 201). A more contemporary example is the investment history of Gulf Oil Corporation in South Korea. In 1966, when Gulf had invested $200 million in South Korea, the

\(^3\) Variations in bribe payments are also consistent with price discrimination by corrupt officials as pointed out in Svensson’s (2003) model. However, uncertainty still prevails for entrepreneurs as there are also large variations for fairly similar firms.
incumbent party asked for $1 million contribution to finance its election campaign. As John T. Noonan (1984, p. 638) notes, "[t]he request was accompanied by pressure which left little to the imagination." Unfortunately for Gulf, elections are held repeatedly. So four years later, S.K. Kim, a leader of the incumbent party, asked again for a ‘campaign contribution’. This time, the demand was $10 million.

This paper builds on the works by Shleifer and Vishny (1993) and Bliss and Di Tella (1997). Shleifer and Vishny’s main concern is to investigate how the harmful effects of corruption depend on "the industrial organization of corruption.” They argue that when corruption activities are decentralized, the harmful effects of corruption are accentuated. As different agencies set their bribery demands independently in order to maximize their own revenue, they do not take the negative externalities on other agencies’ revenues into account. As in Cournot (1927), corruption among independently bribe-setting officials results in a lower level of entry and thus a lower aggregate level of revenue than the joint optimum. Bliss and Di Tella (1997) investigate the relationship between market competition and corruption. They recognize that the extent of competition is not an exogenous parameter since corruption itself can affect the number of firms in a free-entry equilibrium through the endogenously determined level of graft. In a model where the level of corruption and the extent of entry are co-determined by what they call "deep competition” parameters, they show that there is no simple relationship between competition and corruption, thus questioning the validity of a commonly held belief that competitive pressures in the market can mitigate corruption. In this paper, we are concerned with dynamic aspects of corruption. As in Bliss and Di Tella (1997), we abstract from the issue of coordination among corrupt officials and only consider a monopoly corruption scheme; there is only one official who can issue the license which is required to run a

---

4The seminal paper on the economics of corruption is Rose-Ackerman (1975). For a recent survey of the literature, the reader is referred to Ades and Di Tella (1997).
business and this official demands illegal side payments for his services.\(^5\) However, we extend the analysis to a dynamic situation where the official who has collected the bribe previously comes back to demand more. We analyze whether there are additional harmful effects on resource allocation due to the possibility of repeated extortion.

The formal structure of the problem is similar to the licensing model by Farrell and Gallini (1987) where consumers incur setup costs to use a new product. This creates a dynamic consistency problem for the monopolist once consumers have incurred these costs. The main result is that second-sourcing through licensing can serve as a price commitment for the future when long-term contracts with consumers are infeasible due to the nonverifiability of quality. In our extortion model, however, there is no such mechanism to commit to a future demand since a corrupt official with the power to extort usually does not possess a credible mechanism to promise ‘competition among corrupt officials’ in the future. In the context of extortion, it is also natural to assume that the long-term contracts are not enforceable in that they involve illegal activities.

The remainder of the paper is organized in the following way. In Section 2, we set up the basic model of dynamic extortion with only one type of technology. We show that the inability to commit to future demand does not entail further efficiency losses because the official can give a discount in the first period for his future opportunistic behavior. In Section 3, we enrich the model

---

\(^5\)It is not always a subordinate government employee who exacts money from businesses in an ex post opportunistic manner. In 17\(^{th}\) century England, it was the Crown itself that expropriated wealth from citizens – against the will of the Parliament. One method used by the Crown was to demand ‘loans’ which were neither voluntarily given nor likely to be repaid (North and Weingast, 1989). In 19\(^{th}\) century US, local ruling party organisations forced utilities or other corporations to pay large bribes by threatening the firm to enact a legislation that would be harmful to the firm’s investment (Miller, 1989). This practice is not completely gone today. As McChesney (1997) reports there is even a separate term for this kind of legislature in some regions of the Midwest. Such bills are called “fetchers” because they allow politicians to fetch campaign contributions.
with the endogenous choice of technologies by potential entrants. In Section 4, we analyze how the stability of a corrupt regime affects economic performance. The dynamic problem analyzed in the paper can also be applied to various other situations where agents have the power to extort such as in organized crime and the expropriation of multinational corporations by host governments. We conclude with a brief discussion of these extensions.

2 The basic model of repeated extortion

We develop a two-period model of repeated extortion. Consider a government official who has the power to issue licenses that allow entrepreneurs to open a shop.\(^6\) The official sets the price of the license to maximize revenues from licensing. In addition to the license fee, entrepreneurs need to incur costs of \(K\) to enter the business, which are assumed to be sunk.

Entrepreneurs are heterogeneous in their ability to generate (net) income in each period, denoted by \(v\). Let us normalise the total population of entrepreneurs to unity. The distribution of abilities is given by the inverse cumulative distribution function \(F(v)\) with continuous density \(F' \leq 0\), that is, \(F'(v)\) denotes the proportion of entrepreneurs who can generate income \textit{more} than \(v\) in each period. The type of entrepreneurs is private information to entrepreneurs. The government official knows only the distribution of types. Once the entrepreneurs have made their sunk investment, the official may require that for the continuation of the business the license be renewed later at a fee.

\(^6\)As pointed out by Stigler (1971), "[t]he state has one basic resource which in pure principle is not shared with even the mightiest of its citizens: the power to coerce." The state’s monopoly on coercion can lead to the abuse of power when public officials have wide discretion and little accountability due to the lack of formal checks and balances (World Bank, 1997).
The static problem

We first analyse a static problem as a benchmark. This preliminary analysis also helps us develop notation. Consider the official’s one-period static problem when the entrepreneur has operating cost of \( c \). All players are assumed to be risk neutral. If the official demands \( m \) for the license, the marginal type who is indifferent between entry and exit is given by \( v = m + c \). Thus, the official solves:

\[
\max_m m \cdot F(c + m)
\]  

(1)

The first order condition for the optimal \( m \) is given by:

\[
F(c + m) + m \cdot F'(c + m) = 0
\]  

(2)

We make the standard assumption that the distribution of types satisfies the monotone hazard rate condition, that is, \(-F'/F\) is increasing:

\[
-F'' \cdot F + (F')^2 > 0
\]  

(3)

This assumption ensures that the official’s objective function is quasi-concave and the second order condition for the maximization problem is satisfied:

\[
2 \cdot F'(c + m) + m \cdot F''(c + m) < 0.7
\]  

(4)

Let \( m^*(c) \) be implicitly defined by (2) and be the solution to the above problem, i.e.,

\[
m^*(c) = \arg\max_m m \cdot F(c + m),
\]  

(5)

\footnote{Using the first order condition, we can rewrite the second order condition as \( 2 \cdot F'(c + m) - F''(c + m) \cdot F(c + m)/F'(c+m) < 0 \). The second order condition holds if the distribution \( F \) satisfies the monotone hazard rate condition. This condition is a standard assumption in the incentive literature and is satisfied by most widely used distributions; see Fudenberg and Tirole (1991).}
and let \( R(c) = m^*(c) \cdot F(c + m^*(c)) \) be the indirect revenue function for the official. Then, the marginal entrepreneur is given by \( v^*(c) = c + m^*(c) \).

To obtain some comparative static results, we totally differentiate the first order condition (2):

\[
\frac{\partial m^*(c)}{\partial c} = -\frac{F' + m \cdot F''}{[\text{s.o.c}]}.
\]

(6)

where \([\text{s.o.c}]\) denotes the second order condition (4) and is negative. By using the first order condition and condition (3), we can verify that the numerator is negative \((F' + m \cdot F'' < 0)\). Thus,

\[
\frac{\partial m^*(c)}{\partial c} < 0.
\]

(7)

As the operating cost of entrepreneurs increases, the optimal monetary demand by the official decreases. However, the overall effect of the increased operating cost on the extent of entry is negative:

\[
\frac{\partial v^*(c)}{\partial c} = \frac{\partial [c + m^*(c)]}{\partial c} = \frac{F'}{[\text{s.o.c}]} > 0
\]

(8)

By using the envelope theorem, we can also verify that

\[
\frac{\partial R(c)}{\partial c} = m^*(c) \cdot F' < 0
\]

(9)

The official’s revenue decreases with the increase of the entrepreneurs’ operating cost.

The dynamic problem with commitment

We now consider a dynamic (two-period) problem where the official can come back to demand more in the second period. The timing is as follows. At the beginning of the first period, the official demands \( m_1 \) as a licensee fee for opening a business. Potential entrepreneurs know their own type \((v)\) and decide whether to enter or not. If they enter, they have to make specific investment of \( K \) which is not recoverable upon exit. Let us assume that there is no further operating cost once
the sunk investment is made. In the second period, the official can demand more money ($m_2$) as a license renewal fee given the number of firms that entered in the first period. The firms who entered in the first period decide whether to stay in the business by paying $m_2$ or exit from the market (see Figure 1). Those firms that did not enter in the first period can potentially enter the market in the second period by paying $m_2$ in addition to the sunk cost of $K$. As we will show, however, there will be no new entry in the second period in equilibrium.

[Figure 1 about here]

The official cannot price discriminate against the existing firms and give discounts to new entrants in the second period. This reflects our assumption about the information structure the government official has in the second period about individual entrepreneurs. We assume that the entrepreneurs are anonymous in that the existing firms can disguise as new entrants if any discounts are offered to new entrants. This implies that there are no incentives for the entrepreneurs to delay their entry to disguise as low types in order to elicit the discount later, as in the ratchet model (Freixas, Guesnerie, and Tirole, 1985; Laffont and Tirole, 1988). Where does the scenario of

---

8This assumption is made without any loss of generality since we can interpret $v$ as the income generated net of any operating cost.

9It may be technically feasible to give discounts to existing firms by demanding them to submit the original license. However, there is no incentive to do that for the official in the second period.

10If the first-period entrants can be ”identified” in the second period, there can be inefficiencies even in the absence of sunk cost. The reason is that those firms that enter in the first period reveal that they are high type entrepreneurs. This updated information in the second period allows the official to price discriminate against the first-period entrants, charging them a higher price while setting a lower price for new entrants. As a result, inefficiencies can arise due to the delayed entry. We confine our analysis to the anonymous case not because the ”identified” entrepreneurs case is not important, but because the qualitative results for the latter case resemble the ratchet model, and thus are relatively well understood. For more details of the ”identified” case, see Choi and Thum (2003).
anonymous firms apply? It is certainly difficult for large corporations relying on large amounts of physical capital – as in the introductory example of Gulf’s FDI in Korea – to disguise as new entrants. For small (and new) enterprises, however, the most important investment is in intangible assets such as human capital and entrepreneurial skills. Here, it is much easier to disguise as new entrants: towards the corrupt official, a firm can simply install a front man and claim that the enterprise is a new entry. (In equilibrium, of course, they never have to.) Anonymity prevails for small firms that rely mostly on human capital investment (and finally most enterprises start out this way).

To some extent, the information structure may be endogenously determined by the official through his decision concerning whether or not to monitor individual entrepreneurs. Import licenses, for instance, can be made anonymous by granting entrepreneurs the right to resell them in the secondary market. It can be shown that the anonymous case generates more revenues for the official.\footnote{See Fudenberg and Tirole (1998) for a related result.} Thus, the anonymous information structure may arise endogenously if the official has some control over the information structure.

The official cannot commit to $m_2$ before entry occurs in the first period. Thus, the official has the temptation to exploit those who incurred sunk costs in the first period. In this setting, we ask whether the official’s \textit{ex post} opportunism distorts the resource allocation any further. To answer this question, however, we first consider the counterfactual case where the official can \textit{commit} to his future demand in the first period before the entry decisions are made. Let $m_1$ and $m_2$ be the monetary demands by the official in period 1 and period 2, respectively. In Appendix A, we establish that the optimum is to have the same number of firms in both periods.\footnote{A similar proof is given in Farrell and Gallini (1987) in the context of technology adoption.} The reason for this outcome is the following: Any deviation from a constant number of firms implies that some
firms stay in the market for one period only. To induce them to enter, the official has to grant a
discount so that these firms can recoup their investments within one period. The discount, however,
does not only apply to those marginal firms but to all firms. Thus the corrupt official is always
better off by keeping the number of firms constant.

Given that the number of firms staying in the market is constant across periods in the optimum,
the marginal type who is indifferent between entry and exit is defined by $v \cdot (1 + \delta) = K + m_1 + \delta \cdot m_2$, where $\delta < 1$ is the discount factor. For the marginal investor $[v = (K + m_1 + \delta \cdot m_2) / (1 + \delta)]$, the
present value of profits has to be equal to the start-up costs plus the present value of bribes. Thus,
in the commitment solution, the corrupt official solves the following problem:

$$\max_{m_1, m_2} (m_1 + \delta \cdot m_2) \cdot F \left[ \frac{K}{1 + \delta} + \frac{m_1 + \delta \cdot m_2}{1 + \delta} \right]. \quad (10)$$

Let $m = (m_1 + \delta \cdot m_2) / (1 + \delta)$ be the average discounted monetary demand by the official. Then,
the first order condition can be written as:

$$F \left( \frac{K}{1 + \delta} + m \right) + m \cdot F' \left( \frac{K}{1 + \delta} + m \right) = 0 \quad (11)$$

The optimal commitment solution is $m_c = m_c^* (K/(1 + \delta)) = (m_1^c + \delta \cdot m_2^c) / (1 + \delta)$. In other
words, any combination of $(m_1^c, m_2^c)$ that has the same discounted average of $m_c = m^* (K/(1 + \delta))$
can be the optimal solution. For instance, constant demands of $m_1 = m_2 = m^* (K/(1 + \delta))$ are
optimal. The marginal entrant is $v_c = K/(1 + \delta) + m^* (K/(1 + \delta))$.

**Proposition 1.** Any combination of $m_1^c (> 0)$ and $m_2^c$ that has the same discounted average
value of $m_c = m^* (K/(1 + \delta)) = (m_1^c + \delta \cdot m_2^c) / (1 + \delta)$ is optimal for the corrupt official. The
marginal type of entrepreneur who is indifferent between entering and staying out is given by
$v_c = K/(1 + \delta) + m^* (K/(1 + \delta))$. Thus, the solution is equivalent to the repetition of the static
revenue maximization problem when the entrepreneurs’ per period cost is given by $K/(1 + \delta)$. 


The dynamic problem without commitment

Now let us analyze the case where the official cannot commit to the future level of demand before the entry decision is made. In this case too, we can demonstrate that the optimal strategy is to induce a constant number of firms to stay in business for both periods (see Appendix A). Let us denote $v^{NC}$ as the marginal type when no commitment is possible. Then, the official will demand $m_2^{NC} = v^{NC}$ in the second period. Given that the whole surplus is extracted in the second period for the marginal type, the first period demand should be sufficiently low to induce the marginal type to invest in the sunk cost which implies that $m_1^{NC} = v^{NC} - K$. Since the time-consistent demand schedule $(m_1^{NC}, m_2^{NC})$ is uniquely determined by $v^{NC}$, we will find it more convenient to treat $v^{NC}$ as the control variable. Thus, the maximization problem for the official can be written as:

$$\max_{v^{NC}} (m_1 + \delta \cdot m_2) \cdot F(v^{NC}) = \left[ (v^{NC} - K) + \delta \cdot v^{NC} \right] \cdot F(v^{NC})$$

(12)

The first order condition for $v^{NC}$ is given by:

$$F(v^{NC}) + \left[ v^{NC} - \frac{K}{1 + \delta} \right] \cdot F'(v^{NC}) = 0$$

(13)

Thus,

$$v^{NC} - \frac{K}{1 + \delta} = m^* \left( \frac{K}{1 + \delta} \right)$$

(14)

determines the optimal marginal type [see (5) for the definition of the $m^*$-function]. Using the information on the marginal type we obtain a solution for the official’s optimal bribery demands when no commitment is possible:

$$m_1^{NC} = v^{NC} - K = m^* \left( \frac{K}{1 + \delta} \right) - \frac{\delta \cdot K}{1 + \delta}$$

(15)

$$m_2^{NC} = v^{NC} = m^* \left( \frac{K}{1 + \delta} \right) + \frac{K}{1 + \delta}$$

That is, the government official discounts the initial demand by the amount of sunk cost ($m_2^{NC} - m_1^{NC} = K$) to compensate for his ex post opportunistic behavior.
The optimal time-consistent demand schedule above was derived assuming that the first period demand can be negative (i.e., the initial subsidy for entry). This assumption corresponds to the case where the sunk investment is mainly in physical capital so that the official can verify whether the investment has been undertaken by the entrepreneurs who received the subsidy.\footnote{The official in the central office, for instance, can reimburse the entrepreneurs who present the proof of capital purchase or issue vouchers for the capital. Other than that, there would be no further monitoring of the entrepreneurs' activities to keep their identities anonymous. Whether this is possible will partly depend on exogenous considerations.} Alternatively, the official is able to provide the sunk investment himself. Otherwise, the entrepreneurs will just take the money and disappear without any investment. In such a case, we can easily verify that there are no differences between the cases of commitment and no commitment in terms of the number of entrants and the government official's revenue. Since the government official can induce the optimal amount of entry by appropriately discounting his initial demand, there is no additional cost associated with the dynamic consistency requirement.

In most cases, however, we should impose the non-negativity constraint on the initial demand ($m_{1}^{NC} \geq 0$). The sunk investment, for instance, may represent mainly human capital components which cannot be observed. The need to maintain the anonymity of the entrepreneurs may also preclude the aforementioned verification mechanism concerning the usage of the subsidy. Then, the commitment solution can be replicated only when the nonnegative constraint is not binding ($m_{1}^{NC} \geq 0$), in which case the optimal solution is once again given by (15). More specifically, note that $m^{*}(K/(1 + \delta))$ is decreasing in $K$ (Eq. (7)), which makes $m_{1}^{NC} = m^{*}(K/(1 + \delta)) - (\delta \cdot K)/(1 + \delta)$ a monotonically decreasing function of $K$; there exists a unique critical value $K$ such that $m_{1}^{NC} \geq 0$ if and only if $K \leq K$. Therefore, if the cost of sunk investment is substantial ($K > K$), the non-negative constraint is binding and the optimal solution is given by: $m_{1}^{NC} = 0$, $m_{2}^{NC} = v^{NC} = K$. Note that in this case ($K > K$), the no commitment solution
entails an efficiency loss compared to the commitment solution in that there is too little entry
\[ v^{NC} = K > v^C = K/(1 + \delta) + m^*(K/(1 + \delta)) \].

**Proposition 2.** Let \( K \) be the unique value that satisfies \( m^*(K/(1 + \delta)) = (\delta \cdot K)/(1 + \delta) \). If \( K \leq \overline{K} \), the optimal time-consistent extortion schedule is \( m^{NC}_1 = m^*(K/(1 + \delta)) - (\delta \cdot K)/(1 + \delta) \) and \( m^{NC}_2 = m^*(K/(1 + \delta)) + K/(1 + \delta) \). There is no further efficiency loss due to the government official’s inability to commit to future demand in that \( v^{NC} = v^C = K/(1 + \delta) + m^*(K/(1 + \delta)) \).

If \( K > \overline{K} \), the optimal schedule is \( m^{NC}_1 = 0, m^{NC}_2 = v^{NC} = K \). In this case, the optimal solution induces too little entry. In both cases, the official offers a first-period discount for the cost of sunk investment to satisfy the dynamic consistency requirement.

We can conclude that unless the cost of sunk investment is sufficiently large, there is no further inefficiency loss due to repeated extortion. However, we note that the demand schedule is increasing over time due to the initial discount \( (m^{NC}_2 - m^{NC}_1 = K) \). This neutrality result will be a very useful benchmark for our further analysis. We do not want to claim that the neutrality result holds in every instance (as it has already become clear from the case where \( K > \overline{K} \)). However, even though our basic setup is very simple, this result will nevertheless hold in various extensions of the model.

**Some extensions**

**Costs:** So far, we have assumed that the entrepreneurs make an investment in the first period but have no further operating costs in the second period. The outcome of the model, however, would not be affected if there were positive costs \( K_t \) in both periods with \( K_1 > K_2 > 0 \). Then we can interpret \( K_2 \) as the operating costs and \( K_1 - K_2 \) as the fixed investment for market entry. So far, \( v \) has denoted the entrepreneurs’ profits net of operating costs. By simply adding the operating costs, we get an entrepreneur’s gross revenue per period \( (v + K_2) \). The outcome of the model is obviously not affected by this generalization of the model.
Exit: The neutrality result will also go through if some firms exogenously exit in the second period. To simplify the analysis, we can simply forego discounting and interpret $\delta$ as the probability of exit. Except for the interpretation of the discount/exit factor $\delta$, nothing has changed compared to the basic scenario. Risk-neutral entrepreneurs maximize their expected profits. Hence, the marginal investor is given by $v = (K + m_1 + \delta \cdot m_2)/(1 + \delta)$. The amount of entry that is optimal in the commitment case can again be reached by granting a sufficient discount in the first period.

Entry: The entry of new firms might affect the neutrality outcome. Suppose that there are two identical generations of entrepreneurs, i.e. in both periods an identical distribution $F(v)$ of entrepreneurs considers market entry. Both generations of entrepreneurs have to incur the same fixed cost $K$ for entering business. However, the second generation of entrepreneurs only has one period to earn a positive return on their investments. Suppose furthermore that entrepreneurs only have the choice to enter or to stay out of the market. We thus eliminate the possibility that first period entrepreneurs may want to exert the waiting option for a later entry. To see under which conditions a lack of commitment creates further inefficiencies, consider first the counterfactual case where the corrupt official can identify firms according to their entry dates and exercise price discrimination. Then the corrupt official would charge the early entrants $m^*(K/(1 + \delta))$ per period and the second period entrants $m^*(K)$ [see Eq. (5) and Proposition 1]. Note that the marginal entrant of the first generation has a lower type than the marginal entrant in period 2. Now we turn to the relevant case where the corrupt official cannot observe a firm’s entry date. With commitment, the corrupt official can replicate the full information outcome by simply fixing the second period price for the license to $m^*(K)$ and increasing accordingly the price in the first period. Without commitment, this strategy will not work, since the second period price $m^*(K)$ is not time consistent as the corrupt official has an incentive to exploit the early entrants. Hence, the neutrality result will not hold in this case.
We have shown that the neutrality result of Proposition 1 holds in various scenarios but also can be upset by modifying the basic model. We will use the neutrality result as a benchmark for an interesting extension of our model. In the next section, we argue that the entrepreneurs’ incentive to take advantage of the initial discount may lead to inefficient entry behavior if the choice of technology is endogenous. In the remainder of this paper, we will simply assume that $K \leq \bar{K}$ to abstract from inefficiency considerations due to the binding non-negative constraint on initial demand.\footnote{In contrast, Farrell and Gallini’s (1986) analysis of licensing focuses on the case where the non-negativity constraint is \textit{binding}.}

3 Dynamic extortion and the choice of technology

Now we assume that entrepreneurs have available another type of technology with which to enter the market. More specifically, this alternative technology entails less sunk cost and higher per-period operating costs. For simplicity, this technology is assumed to involve no sunk costs; instead there are operating costs of $k$ per period, where $(1 + \delta) \cdot k > K > k$. Thus, the alternative technology is less efficient if production takes place in both periods. However, it protects the entrepreneur from the official’s \textit{ex post} opportunistic behavior because it does not involve any sunk capital. We assume that the official cannot observe which type of technology has been chosen in the first period and thus cannot price discriminate based on the type of technology chosen. This is also in line with our assumption on anonymity. Anonymity prevails when the corrupt official cannot observe whether an applicant for a license has already undertaken an investment in intangible assets. As discussed earlier, this is particularly relevant for investment in specific human capital. In this case, the official will also not be able to identify the type of investment. If he cannot observe \textit{whether} an entrepreneur has acquired specific human capital, he can certainly not verify whether
an entrepreneur has invested in his own human capital \((K)\) or has acquired these skills externally by hiring someone temporarily \((k)\).

With the availability of this short-term investment strategy, the optimal dynamic demand schedule cannot be sustained. To see this, consider the marginal type who was indifferent between entering and staying out in the no commitment case, \(v^{NC} = K/(1 + \delta) + m^* (K/(1 + \delta))\). This marginal type’s surplus was completely extracted with the choice of the \(K\)-technology. The marginal type, however, can do better when the \(k\)-technology becomes available. Facing the demand schedule, \(m_1^{NC} = m^* (K/(1 + \delta)) - (\delta \cdot K)/(1 + \delta)\) and \(m_2^{NC} = m^* (K/(1 + \delta)) + K/(1 + \delta)\), he can enter by choosing the \(k\)-technology in the first period when the discount is offered by the official and exit in the second period when the amount of extortion is increased – a strategy that may be called dynamic cream skimming. Choosing this alternative technology with less sunk costs yields the payoff

\[
v^{NC} - k - m_1^{NC} = K - k > 0. \tag{16}
\]

In fact, with the demand schedule of \((m_1^{NC}, m_2^{NC})\), any entrant below type \(v = v^{NC} + (K - k)/\delta\) will choose the \(k\)-technology and exit from the market in the second period when \(m_2^{NC}\) is demanded. Thus, the optimal schedule identified in the previous section is no longer optimal with the availability of a short-term investment strategy.

What will the equilibrium look like when the entrepreneurs can shield themselves from the official’s ex post opportunism by choosing the short-term investment strategy? To derive the equilibrium strategy profile when the choice of technology is endogenous, we first prove in Appendix

\[15\]Cream skimming in the regulation literature refers to the inefficient firm’s selective entry into the most profitable markets when the regulated incumbent firm practices cross subsidization between markets (Viscusi, Vernon, and Harrington, 1995). In our model, the first period discount in the optimal demand schedule can be interpreted as intertemporal subsidization.
B that the choice of technology is characterized by the following cut-off rule.

Lemma 1. There is a critical value $\hat{v}$ such that all types above it choose the $K$-technology whereas all entrants below it choose the $k$-technology in equilibrium.

When entrepreneurs are uncertain about the future bribes, they can protect themselves to some extent by investing in the flexible technology $k$. If the bribery demand turns out to be high, the entrepreneur can simply exit the market. Only high-ability entrepreneurs will still invest in the $K$-technology. Because of their highly profitable projects, they will stay in the market even with high bribery demands.

Given that the entrepreneurs’ choice of technology is characterized by a cut-off rule, we are now in a position to demonstrate that there is no pure strategy equilibrium in the presence of the $k$-technology.

Proposition 3. There is no pure strategy equilibrium when the choice of technology is endogenous.

Proof. See Appendix C.

The intuition for the non-existence of a pure strategy equilibrium can be explained as follows. In the second period, the official’s optimal choice is either to take advantage of the entrepreneurs who entered with the $K$-technology or to allow additional entrants with the $k$-technology. If the official chooses the first option for sure, the marginal type whose entire surplus is extracted in the second period with the $K$-technology has an incentive to switch to the $k$-technology. Thus, the first option cannot be sustained as equilibrium. If the second option is chosen for sure, the marginal entrant with the $k$-technology can be better off by entering with the $K$-technology in the first period rather than making repeated short-term investments in the $k$-technology. Once again, this upsets the putative pure strategy equilibrium. Thus, there is no pure strategy equilibrium.
To derive the mixed strategy equilibrium, we also use Lemma 1 that the choice of technology is characterized by the cut-off rule. Suppose that all the types above type $\hat{v}$ have entered with the $K$-technology in the first period. In the second period, the relevant state variable for the official is the cut-off value $\hat{v}$ of the entrepreneurs’ types who have chosen the $K$-technology. There is no distinction between those who have entered with the $k$-technology and those who stayed out in the first period because both have to incur the same cost of $k$ to operate in the second period.

**Lemma 2.** For the official to adopt a mixed strategy, the critical type $\hat{v}$ must be larger than $v^*(0) = m^*(0)$.

*Proof.* Suppose not, i.e., $\hat{v} \leq v^*(0)$. Then, the optimal strategy for the official in the second period is to demand $m_2 = v^*(0) = m^*(0)$ with probability 1. Thus, we have a contradiction. *Q.E.D.*

Lemma 2 is illustrated in Figure 2. The revenue function is concave in $v$ with a maximum at $v^*(0)$. As shown, critical types $\hat{v}$ below $v^*(0)$ are not compatible with the mixed strategy equilibrium. There, the number of firms having entered with the $K$-technology is so large that the official’s best strategy would be to demand $m^*(0)$ for sure thus forcing some firms to exit. Hence, the critical type has to be in the downward sloping part of the revenue function. Given that $\hat{v} > v^*(0)$, the optimal demand in the second period is either to charge $\hat{v}$ which prevents any entry with the $k$-technology or to charge $m^*(k)$ which allows entry with the marginal type $v^*(k) = k + m^*(k)$. The corrupt official receives a revenue of $\hat{v} \cdot F(\hat{v})$ in the former case and of $R(k) = m^*(k) \cdot F(k + m^*(k))$ in the latter case. For the official to mix between these two demands, he has to be indifferent between the two strategies:

$$\hat{v} \cdot F(\hat{v}) = m^*(k) \cdot F(k + m^*(k)) \quad (17)$$

Condition (17) pins down the critical type $\hat{v}$ who should be indifferent between entering with the $K$-technology and the $k$-technology; $\hat{v}$ is the larger root to equation (17).
Lemma 3. \( v^* > v^*(k) = k + m^*(k) \) \( v^C = \frac{K}{1+\delta} + m^* \left( \frac{K}{1+\delta} \right) \).

Proof. Recall that \( \hat{v} \) is the larger root to equation (17): \( \hat{v} \cdot F(\hat{v}) = m^*(k) \cdot F(k + m^*(k)) < [k + m^*(k)] \cdot F(k + m^*(k)) \). Since \( \phi(v) = v \cdot F(v) \) is a quasi-concave function of \( v \) and at \( \hat{v} \), \( \phi'(\hat{v}) < 0 \), we have \( \hat{v} > v^*(k) = k + m^*(k) \). Q.E.D.

Lemma 3 is also illustrated in Figure 2. From Lemma 2, we know that the critical type will be in the downward sloping range of the revenue function. If the official is indifferent between his two feasible strategies \( m^*(k) \) (with the marginal type \( v^*(k) \)) and \( \hat{v} \), \( \phi(v) = v \cdot F(v) \) evaluated at \( v^*(k) \) has to be higher than the one evaluated at \( \hat{v} \), as entrepreneurs also have to finance their investments \( k \). Hence, the critical type \( \hat{v} \) has to be to the right of \( v^*(k) \).

Lemma 4. In the first period, the corrupt official charges \( m^*(k) \) and the marginal entrant is \( v^*(k) = k + m^*(k) \).

Proof. In the first period, there are two possible scenarios: (1) all entrepreneurs enter with the \( K \)-technology, or (2) some entrepreneurs enter with the \( K \)-technology while others enter with the \( k \)-technology.

Under the first scenario, no entrepreneur whose type is below \( \hat{v} \) enters the market in the first period. Let \( \alpha \) be the probability that the official chooses to demand \( \hat{v} \) in the second period. Then, the condition that the \( \hat{v} \) type to be indifferent between entering and not entering in the first period is:

\[
\hat{v} - K - m_1 + \delta \cdot (1 - \alpha) \cdot [\hat{v} - m^*(k)] = 0 + \delta \cdot (1 - \alpha) \cdot [\hat{v} - k - m^*(k)]
\]

(18)

Thus, \( m_1 = \hat{v} - K + \delta \cdot (1 - \alpha) \cdot k \). In this case, the corrupt official’s first period revenue is \( [\hat{v} - K + \delta \cdot (1 - \alpha) \cdot k] \cdot F(\hat{v}) \).

Under the second scenario in which the corrupt official allows entry with the \( k \)-technology, the optimal demand is \( m^*(k) \) which yields the first period revenue of \( m^*(k) \cdot F(k + m^*(k)) \). By Lemma
3, \( m^*(k) \cdot F(k + m^*(k)) = \hat{v} \cdot F(\hat{v}) > [\hat{v} - K + \delta \cdot (1 - \alpha) \cdot k] \cdot F(\hat{v}) \). Thus, we can conclude that the corrupt official charges \( m^*(k) \) in the first period and induces entry with both types of technology. \( Q.E.D. \)

A corollary of Lemmas 3 and 4 is that there are two types of inefficiencies associated with the availability of the short-term investment strategy \( k \). First, there is too little entry compared to the previous case \( (v^*(k) > v^C) \). Second, some of those who enter do so with an inefficient technology \( k \) \( (v \in [v^*(k), \hat{v}]) \).

The mixing probability \( \alpha \) for the official is determined by the critical type \( \hat{v} \) who has to be indifferent between the two technologies:

\[
\hat{v} - K - m_1 + \delta \cdot (1 - \alpha) \cdot [\hat{v} - m^*(k)] = \\
= \hat{v} - k - m_1 + \delta \cdot (1 - \alpha) \cdot [\hat{v} - k - m^*(k)].
\]  \hspace{1cm} (19)

Therefore, the official charges \( \hat{v} \) with probability \( \alpha = [(1 + \delta) \cdot k - K]/[\delta \cdot k] \) and \( m^*(k) \) with probability \( (1 - \alpha) = (K - k)/[\delta \cdot k] \). Our analysis up to now can be summarized in Proposition 4 and Figure 2.

**Proposition 4.** When the choice of technology is endogenous, the corrupt official employs a mixed strategy with respect to his future bribery demand in equilibrium. In the first period, the corrupt official charges \( m^*(k) \) having the entrepreneur with \( v^*(k) \) as the marginal entrant. In the second period, the official charges a high price \( \hat{v} \) with probability \( \alpha = [(1 + \delta) \cdot k - K]/[\delta \cdot k] \) and a low price \( m^*(k) \) with probability \( (1 - \alpha) = [K - k]/[\delta \cdot k] \). In the first case, only those with technology \( K \) stay in the market; in the latter case, entrepreneurs with the flexible technology \( k \) also operate in the market.
Proposition 4 explains why extortion is often associated with arbitrariness and creates uncertainty. It also provides a new interpretation of Wei’s (1997) empirical study on corruption. He examines the effect of corruption-induced uncertainty on foreign direct investment and shows that the second moment (variability) effect is negative, statistically significant and quantitatively large. For instance, an increase in the uncertainty level from that of Singapore to that of Mexico, keeping the average level of corruption constant, is equivalent to raising the tax rate by 32 percentage points. Wei’s empirical result is based on a model where the level of foreign investment is adversely affected by the increase in the variability of the bribe rate due to the foreign investors’ risk aversion.

Our model, however, suggests that there is no causal relationship between uncertainty and inefficient investment behavior. In our model, the entrepreneurs are risk-neutral and the inefficiency stems from the entrepreneurs’ incentive to practice dynamic cream skimming. Uncertainty per se is not a deterrent to investment. Uncertainty is rather a part of equilibrium and is endogenously determined together with the level of investment and the choice of technology.\textsuperscript{16}

In Wei’s model, uncertainty is treated as an exogenous variable. Thus, one obvious policy implication of his model would be to make the bribery schedule as transparent as possible to promote foreign direct investment. In contrast, our model not only explains the genesis of corruption-induced inefficiency but also the link between variability in bribery payments and inefficiency can be derived in our model. In equilibrium, the variance in bribery payments amounts to

\[ \sigma^2 = \alpha \cdot (\hat{v} - Em)^2 + (1 - \alpha) \cdot (m^* - Em)^2 \]

with

\[
\alpha = \frac{(1 + \delta) \cdot k - K}{\delta \cdot k} \quad \text{and} \quad Em = \alpha \cdot \hat{v} + (1 - \alpha) \cdot m^*.
\]

If the expected bribery payment is kept constant \((Em = \alpha \cdot \hat{v} + (1 - \alpha) \cdot m^*)\), we can write the variance as

\[ \sigma^2 = (\hat{v} - Em)^2 \cdot \alpha / (1 - \alpha). \]

Consider a comparative statics exercise in which \(k\) is increased making the \(k\)-technology relatively more inefficient. Note that this change induces a lower \(m^*\), but higher \(v^*(k)\) and \(\hat{v}\), with the result of further efficiency loss. We can find a level of \(K\) that will make \(Em = \alpha \cdot \hat{v} + (1 - \alpha) \cdot m^*\) constant with a small increase in \(k\). However, this change makes the variance higher since \(\alpha\) is increasing in \(k\). Thus, an increase in the variability of bribery payments can be associated with further efficiency loss.

\textsuperscript{16} As in Wei’s paper, the link between variability in bribery payments and inefficiency can be derived in our model.
uncertainty, but also indicates that arbitrariness is a central feature of corruption, which cannot be tackled in isolation.

We point out that the result on arbitrariness is not an artifact of the two-period assumption and is expected to hold as a general feature of corruption in a multi-period extension of the game.\textsuperscript{17} The reason is that we cannot have an equilibrium in which all entrepreneurs enter only with one type of technology in the presence of the official’s \textit{ex post} opportunism. To allow entry only with the $K$-technology in equilibrium, the official needs to give a discount in the first period to make the schedule dynamically consistent. However, the discount in the first period induces the marginal type to choose the $k$-technology, upsetting the putative equilibrium. Likewise, if the official’s demand schedule were to allow entry only with the $k$-technology in equilibrium, the high type entrepreneurs who would be in the market in all periods would be better off by entering with the $K$-technology, once again upsetting the equilibrium. The $K$-technology is more efficient than the $k$-technology in the event of no exit. This implies that the official has to randomize in bribery demands, inducing stochastic exit of firms with the $k$-technology, in order to have an equilibrium with the coexistence of multiple technologies in the market.

Our results are also consistent with empirical evidence that inefficient technologies are used and that sunk investment remains inefficiently low in countries with extensive corruption. First, there are case studies on several countries. For instance, de Soto (1989) points out that in Peru entrepreneurs preferred to remain small to minimise the risks from intervention by government

\textsuperscript{17}If the model were extended to include multiple periods, the randomness result would hold in every period beyond the first one. To see this, consider a model of $N(>2)$ periods with two technologies. The $K$-technology lasts $N$ periods whereas the $k$-technology lasts only one period with no sunk costs, where $k \cdot (1 - \delta^N) / (1 - \delta) > K > k$. Then, it can be easily verified that there is a stationary equilibrium in which the corrupt official randomizes between $\hat{v}$ and $m^*(k)$ with the low price $m^*(k)$ being charged with the probability of $\left[(1 - \delta)/(1 - \delta^N)\right] \cdot [(K - k)/k]$ from the second period on.
institutions. Safavian, Graham and Gonzalez-Vega (2001) analyse the impact of corruption on micro-firms in Russia. Their interview data suggest that entrepreneurs try to protect themselves from corruption by diversifying entrepreneurial activities, i.e. by having several small firms. The diversification of income sources makes the threat of exit to the regulator more credible. Safavian, Graham and Gonzalez-Vega (2001) conclude that this diversification of entrepreneurial skill across several activities is technologically inefficient. In a study on the dynamics of the timber industry, Vincent and Brinkley (1992) report that timber concessionaires in Malaysia showed a remarkable short-run orientation in their investment behaviour due to the prevalent uncertainty created by corruption. One of the most striking examples is the use of floating power stations (instead of the less expensive and more efficient stationary power stations) to facilitate a possible market exit (Rose-Ackerman, 1998).

Second, the literature on efficient production frontiers suggests that small enterprises in developing countries where corruption is most prevalent remain inefficiently small. In this framework, the maximum output for a given vector of inputs is estimated and the effective output relative to the production frontier can be used as a measure of efficiency for an entire economy or for a specific sector.¹⁸ Mead and Liedholm (1998) and Liedholm (2001) point out that firms in developing countries are not per se less efficient than firms in industrialised economies. For very small firms (i.e. mostly the self-employed), however, they find significant economies of scale. This result suggests that micro-firms tend to be an inefficient organisational form of production in developing countries. One possible explanation for the over-representation of micro-firms in many developing countries is the threat of being extorted by corrupt government officials which can partly be evaded by remaining small.¹⁹

¹⁸For a critical review of this methodology, see Tybout (2000).

¹⁹Unfortunately, most papers do not explicitly deal with the institutional framework of countries. Hence, it is
4 The stability of a corrupt regime

We now consider how the stability of a corrupt regime affects economic performance. Let us parametrize the stability of the regime by $\beta$ which is the probability that the official will remain in power in the second period. We consider two scenarios. In the first scenario, if the corrupt official loses power in the second period, he is replaced by another corrupt official. In this case, the change of power is irrelevant for the entrepreneurs while it matters a lot to the original corrupt official who loses power. In the second scenario, the change of power takes place through a genuine reform where corruption is eliminated in the second period. In this section, we abstract from the issue of endogenous technology choice; only the $K$-technology is available to entrepreneurs. We also ignore discounting by setting $\delta = 1$.

Succession of corrupt regimes

Consider the case where the official in the first period remains in power only with probability $\beta$ in the second period. Thus, $\beta$ plays the role of a discount factor for the official. If there is a change of power, the office is transferred to another corrupt official. As a result, second period demands will be independent of who is in power. Once again, it can be shown that the optimal strategy in the second period is to extract the whole surplus of the marginal type who entered in the first period without inducing any exit. Thus, the marginal type is given by $\hat{v} = m_1 + K$.

The maximization problem for the official in the first period then is:

$$\max_{m_1} m_1 \cdot F(m_1 + K) + \beta \cdot (m_1 + K) \cdot F(m_1 + K).$$

(20)

difficult to isolate the effect of corruption on the investment volume. An exception is the recent paper by Klein and Luu (2002). They estimate a stochastic efficiency frontier model including measures of the institutional framework and conclude that it is not only factor endowments but differences in institutional environments (such as regulatory frameworks, corruption and political stability) that determine the allocation of factors in a country.
The first order condition is given by:

\[ F(m_1 + K) + \left[ m_1 + \beta \cdot \frac{K}{1 + \beta} \right] \cdot F'(m_1 + K) = 0. \]  
(21)

Totally differentiating Eq. (21) with respect to \( m_1 \) and \( \beta \) yields:

\[ [s.o.c] dm_1 + \frac{K}{(1 + \beta)^2} \cdot F'(m_1 + K) d\beta = 0, \]  
(22)

where \([s.o.c]\) denotes the second order condition for (20) and is negative. Thus, we have \( dm_1 / d\beta < 0 \).

As the corrupted regime becomes more stable, there will be more discounts in the first period demand, which induces more entry in the first period.

**Proposition 5.** When a transfer of power may take place between corrupt officials, the stability of the regime is conducive to economic performance. A policy that fights corruption but cannot ensure that corrupt officials are replaced by honest ones is detrimental to economic activity.

This result is complementary to Shleifer and Vishny (1993). They show in a static setting that *interagency* externalities aggravate the problem of corruption in comparison to the simple monopoly corruption scheme. Weak governments cannot prevent its numerous agencies from setting their own bribes independently, thus maximising the private profit of each agency. When an entrepreneur who wants to start a business needs services from several of these agencies, each agency will neglect the externality it creates on the bribery revenue of other agencies. A higher bribery demand of one agency reduces the willingness to pay for complementary services. Due to this externality, the amount of economic activity is lower in equilibrium than with a monopoly corruption scheme. In our model, a stable regime enables the official to internalize the *intertemporal* externality stemming from the existence of sunk cost. Thus, the official is more willing to invest (discount) in the first period demand when he is more confident that he would reap the benefits in the second period.

Our result also has implications for job transfers often observed in various organizations such as planned enterprises in the former Soviet Union, the U.S. foreign service and military. The practice
can be puzzling since transferring individuals to new jobs sacrifices job-specific human capital (Ickes and Samuelson, 1987). One explanation is that job transfers prevent corruption by ensuring that employees do not occupy a job long enough to reap the benefits of corrupt activities.\textsuperscript{20} Our model, however, suggests that job transfers intended to mitigate corruption may have the exact opposite effect.

**Reform case**

Now consider the case where the corrupt regime may be replaced by an honest government through a genuine reform. In this case, the transfer of power matters not only to the initial corrupt official but also to entrepreneurs. If the corrupt official retains power in the second period, it is optimal to extract the whole surplus of the marginal type who entered in the first period without inducing any exit as in the previous case. However, in the case of genuine reform, entrepreneurs may have an incentive to delay their investment, hoping for the installation of a clean government. With international organisations such as the IMF and the World Bank putting more emphasis on the efficiency of government institutions in aid-receiving countries, this case has particular relevance for the future. In August 1997, for instance, the IMF issued new guidelines making the reform of corrupt institutions a prerequisite for financial aid (IMF, 1997). These conditionalities can be viewed as a sign of national commitment to policy reforms for countries with weak domestic commitment mechanisms (World Bank, 1997).

Given a bribery demand of $m_1$, the marginal type who enters is defined by:

$$
(v - K - m_1) + (1 - \beta) \cdot v = (1 - \beta) \cdot (v - K)
$$

\textsuperscript{20}Other explanations for job transfers include mitigating the ratchet effect, sorting employees into the jobs where they will be the most productive, and allowing potential future managers to gain familiarity with various aspects of an organization’s operations. See Ickes and Samuelson (1987) for details.
The left-hand side of Eq. (23) represents the expected payoff from entering in the first period and the right-hand side represents the option value of delaying the investment until the second period. The marginal type is given by \( \hat{v} = m_1 + \beta \cdot K \). The maximization problem for the official in the first period is:

\[
\max_{m_1} m_1 \cdot F(m_1 + \beta \cdot K) + \beta \cdot (m_1 + \beta \cdot K) \cdot F(m_1 + \beta \cdot K). 
\]

(24)

The first order condition is given by:

\[
F(m_1 + \beta \cdot K) + \left[ m_1 + \frac{\beta^2 \cdot K}{1 + \beta} \right] \cdot F'(m_1 + \beta \cdot K) = 0.
\]

(25)

Thus, the optimal demand in the first period is

\[
m_1 = m^* \left( \frac{\beta \cdot K}{1 + \beta} \right) - \frac{\beta^2 \cdot K}{1 + \beta},
\]

(26)

which we can rewrite as

\[
m_1 = m^* \left( \frac{\beta \cdot K}{1 + \beta} \right) + \frac{\beta \cdot K}{1 + \beta} - \beta \cdot K
\]

(27)

[see (5) for the definition of the \( m^* \)-function]. The instability of a regime increases the first period bribery demand for two reasons. First, as the probability of a new regime increases, the corrupt official has less incentive to internalize the intertemporal externality and thus increases the bribery demand in the first period. Second, the entrepreneurs have less incentives to delay their investments since the risk of \emph{ex post} expropriation is reduced with the increase in the probability of a new regime. Thus, the corrupt official can get away with less discount in the first period.

Hence, we have \( dm_1 / d\beta < 0 \) as in the previous case. To verify this, note that both \( (\beta \cdot K)/(1 + \beta) \) and \( (\beta^2 \cdot K) / (1 + \beta) \) in Eq. (26) are increasing in \( \beta \) and that \( m^*(\cdot) \) is a decreasing function.

The effect of stability on the aggregate investment level, however, is opposite to the previous case. The marginal type is given by:

\[
\hat{v} = m_1 + \beta \cdot K = m^* \left( \frac{\beta \cdot K}{1 + \beta} \right) + \frac{\beta \cdot K}{1 + \beta}.
\]

(28)
Since \((\beta \cdot K)/(1 + \beta)\) is increasing in \(\beta\), the cut-off value \(\hat{v}\) increases with \(\beta\) [see Eq. (8)]. We conclude that the effect of stability of a corrupt regime on the extent of entry is harmful in this case. In other words, the possibility of future reform has a positive effect on current aggregate investment. Thus, it is important to distinguish the two cases of regime changes since the stability of a corrupt regime has different implications for investment activities depending on the nature of regime change.

5 Concluding remarks

In this paper, we have analyzed the consequences of repeated extortion. It turned out that the repeated nature of extortion \textit{per se} does not create further distortions in resource allocations. There are no fewer businesses in operation when corrupt officials can make repeated bribery demands than when there are once-and-for-all bribery payments. The reason is that a corrupt official can properly discount his initial demand in order to induce the appropriate amount of entry. The major inefficiency of repeated extortion emerges only when entrepreneurs have discretionary power over the choice of technology. In that case, entrepreneurs react to the dynamic path of bribery demands by distorting their choice of technology in the form of inefficiently low sunk investments. This type of investment behavior allows them to react more flexibly to future extortion. We have also shown how the stability of a corrupt regime affects extortion and investment activities. A corrupt official who fears to lose his power in the future becomes less willing to discount his future demands, thus driving more entrepreneurs out of business.

We conclude with a discussion on how the dynamic problem analyzed in the paper can be applied to other situations that involve agents who have the power to extort. We also mention how the basic framework can be extended depending on the contexts.
Organized crime

The most obvious example is organized crime.\textsuperscript{21} Gangs charge ‘protection money’ from businesses and this kind of extortion is typically done repeatedly. The dynamic nature of the extortion game has serious consequences for behavior on both sides – the gang and the entrepreneurs. Entrepreneurs become more reluctant to put larger sums in sunk investments. They may choose technologies that are deemed inefficient in the absence of organized crime to be able to react more flexibly to extortive threats. In response to the incomplete information about the types of technologies chosen, the gang may randomly change the sum of ‘protection money’ which it demands from businesses.

One important aspect of extortion that is not considered in our model is how gangs actually try to overcome the information asymmetry in order to improve on their rewards of extortion. This can be achieved, for instance, by forcing businesses to purchase complementary inputs from gang owned firms. In the case of the Fulton Fish Market in Manhattan, fish dealers were forced to hire certain (Mafia owned) companies performing the loading function. This arrangement enabled the gang to observe how much fish was traded by each dealer. The sums extorted from each fish dealer could then be related to the volume of loaded or unloaded fish by charging excessive loading fees.\textsuperscript{22} In 1995, when the City of New York installed a regulatory authority to manage the market in an effort to drive organized crime out of Fulton Fish Market, loading costs dropped by 70 percent (Giuliani 1997).\textsuperscript{23}

\textsuperscript{21}For an economic analysis of extortion by organized crime, see Konrad and Skaperdas (1997, 1998).

\textsuperscript{22}The economic rationale for this arrangement is similar to the price discrimination motives for tying where the practice serves as a metering device (Telser, 1979).

\textsuperscript{23}It is estimated that the control over Fulton Fish Market brought $ 50 million a year into the mob’s treasury (Vulliamy, 1998).
Sale of public offices

Sale of public offices was a prevalent phenomenon in many countries and over long periods in the history of states. Sale of offices, in particular the offices of tax collectors, had great advantages when the honesty of officials was hard to monitor or when the potential for tax revenues in distant regions was largely unknown to the central government. Swart (1970) reports that farming out taxes was a firmly established practice, e.g., in France, Spain, Turkey and China. The sale of offices, however, also gave rise to mechanisms of repeated extortion as described in this paper. When the official’s contract was to be renewed, the ruler (king, emperor...) could use the information from the earlier sale of office, thus reducing the tax collector’s incentive to invest in an efficient infrastructure for tax collection. Even worse, the ruler could expropriate his former tax collector – a strategy that was already known to Roman emperor Titus Flavius Vespasianus (9-79 A.D.).24 He sold the offices of procuratores (tax collectors) to the most greedy men only to sentence them to high fines later on (Sturminger, 1982).25

Expropriation of multinational corporations

When it comes to foreign direct investment (FDI), the interaction between a multinational firm and the government of the host country resembles in its structure the repeated extortion game analyzed in this paper. If the government cannot credibly commit to future tax rates (and certain property rights), this will make the foreign multinational more reluctant to invest in the country. Potential

---

24Vespasian was known for his rigor in tax collection to balance the government budget. He even introduced a tax on urine to be paid by tanners who used urine in the production process. When his son Titus felt disgusted and complained, Vespasian showed him a coin and responded with the now famous dictum: “Non olet” (Money doesn’t stink).

25It is, however, not reported by historians whether this type of behavior was anticipated by the tax collectors and reduced their initial bids for the offices.
investors anticipate that their bargaining position versus the government is weakened once it has incurred the sunk costs of irreversible investments. The investors may mitigate the consequences of the unfavorable *ex post* bargaining position by reducing the capital intensity of their investment projects (Doyle and van Wijnbergen, 1994). As the lack of commitment to future tax rates works against the host country’s own interests, it has an incentive to counteract the consequences of its own opportunistic *ex post* behavior by *ex ante* granting tax holidays. The potentially high tax payments in the future are compensated by a zero tax rate or even by subsidies in the initial period (Bond and Samuelson, 1989).

A Proof that it is optimal to have a constant number of firms in the market

The commitment case

Let us denote the marginal types in period 1 and period 2 as $v_1$ and $v_2$, respectively. We show that it is optimal for the official to have $v_1 = v_2$.

**Case 1. $v_1 > v_2$**

In this case, the official attracts new entrants in the second period. For this to happen, the second period demand should be $m_2 = v_2 - K$. For the marginal type $v_1$ to enter in the first period, the following two conditions must be satisfied:

$$m_1 + \delta \cdot m_2 \leq v_1 \cdot (1 + \delta) - K$$

$$\delta \cdot (v_1 - K - m_2) \leq v_1 \cdot (1 + \delta) - K - (m_1 + \delta \cdot m_2)$$

(A-1)

The first one is that the marginal entrepreneur of type $v_1$ makes nonnegative profit with the first period entry. The second constraint is that he prefers to enter in the first period rather than to
delay the entry until the second period. It can be easily verified that the first constraint is not binding. As a result, \( m_1 = v_1 - (1 - \delta) \cdot K \). We can write the revenue for the official as a function of the marginal types in each period:

\[
R(v_1 > v_2) = [v_1 - (1 - \delta) \cdot K] \cdot F(v_1) + \delta \cdot (v_2 - K) \cdot F(v_2)
\]

(A-2)

which can be rewritten as:

\[
R(v_1 > v_2) = \left[ v_1 - \frac{K}{1 + \delta} \right] \cdot F(v_1) + \delta \cdot \left[ v_2 - \frac{K}{1 + \delta} \right] \cdot F(v_2) + [F(v_1) - F(v_2)] \cdot \frac{\delta^2 \cdot K}{1 + \delta}.
\]

(A-3)

Note that \([F(v_1) - F(v_2)]\) is negative when \(v_1 > v_2\).

Case 2. \( v_1 < v_2 \)

In this case, there is no new entry in the second period and some firms that entered in the first period exit in the second period. For this to happen, the corrupt official has to charge \( m_1 = v_1 - K \) and \( m_2 = v_2 \). Also for Case 2, we calculate the official’s revenue:

\[
R(v_1 < v_2) = (v_1 - K) \cdot F(v_1) + \delta \cdot v_2 \cdot F(v_2) = \\
= \left[ v_1 - \frac{K}{1 + \delta} \right] \cdot F(v_1) + \delta \cdot \left[ v_2 - \frac{K}{1 + \delta} \right] \cdot F(v_2) + [F(v_2) - F(v_1)] \cdot \frac{\delta \cdot K}{1 + \delta}.
\]

(A-4)

Note that \([F(v_2) - F(v_1)]\) is negative when \(v_1 < v_2\).

Case 3. \( v_1 = v_2 \)

In this case, the same firms are in the business in both periods. The marginal type is given by \( v_1 = v_2 = v = (K + m_1 + \delta \cdot m_2)/(1 + \delta) \). The revenue for the official is:

\[
R(v_1 = v_2 = v) = [v \cdot (1 + \delta) - K] \cdot F(v) = \\
= \left[ v_1 - \frac{K}{1 + \delta} \right] \cdot F(v_1) + \delta \cdot \left[ v_2 - \frac{K}{1 + \delta} \right] \cdot F(v_2)
\]

(A-5)

34
The comparison of (A-3), (A-4) and (A-5) makes it clear that the revenue for the official is maximized with monetary demands that induce $v_1 = v_2$.

The non-commitment case

To the contrary, assume that the optimal demand scheme induces $v_1 > v_2$. Then, the second period marginal type is given by $v_2^* = \arg\max (v_2 - K) \cdot F(v_2)$. The overall revenue for the official amounts to

$$R^{NC}(v_1 > v_2) = [v_1 - (1 - \delta) \cdot K] \cdot F(v_1) + \delta \cdot (v_2^* - K) \cdot F(v_2^*) =$$

$$= (v_1 - K) \cdot F(v_1) + \delta \cdot (v_2^* - K) \cdot F(v_2^*) + \delta \cdot K \cdot F(v_1).$$  \hspace{1cm} (A-6)

Thus, reducing $v_1$ closer to $v_2 = v_2^*$ can increase the official’s revenue by increasing both $(v_1 - K) \cdot F(v_1)$ and $\delta \cdot K \cdot F(v_1)$. Thus, $v_1 > v_2$ cannot be part of the optimal demand scheme.

Similarly, assume that the optimal demand scheme induces $v_1 < v_2$. Then, the second period marginal entrepreneur is determined by $v_2^{**} = \arg\max v_2 \cdot F(v_2)$. The overall revenue for the official is given by:

$$R^{NC}(v_1 < v_2) = (v_1 - K) \cdot F(v_1) + \delta \cdot v_2^{**} \cdot F(v_2^{**}) =$$

$$= v_1 \cdot F(v_1) + \delta \cdot v_2^{**} \cdot F(v_2^{**}) - K \cdot F(v_1).$$ \hspace{1cm} (A-7)

Thus, increasing $v_1$ closer to $v_2 = v_2^{**}$ can increase the official’s revenue by increasing $v_1 \cdot F(v_1)$ and decreasing $K \cdot F(v_1)$. Thus, $v_1 < v_2$ cannot be part of the optimal demand scheme, either. Q.E.D.

B Proof of Lemma 1

In general, we allow that the official’s second period demand, $\tilde{m}_2$, can be a random variable to accommodate the possibility of a mixed strategy. Then, given $m_1$, the expected value of entering
with technology $K$ for type $v$ is given by:

$$
\pi(K, v) = v - K - m_1 + \delta \cdot E[\max(v - \tilde{m}_2, 0)].
$$

(B-1)

Similarly, the expected value of entering with the $k$-technology is:

$$
\pi(k, v) = v - k - m_1 + \delta \cdot E[\max(v - k - \tilde{m}_2, 0)].
$$

(B-2)

The difference between these two choices is:

$$
\Delta(v) \equiv \pi(K, v) - \pi(k, v) =
$$

$$
= -(K - k) +
$$

$$
\delta \cdot \{E[\max(v - \tilde{m}_2, 0)] - E[\max(v - k - \tilde{m}_2, 0)]\} =
$$

$$
= -(K - k) + \delta \cdot \left\{E \left[I_{\tilde{m}_2 \leq v} \cdot (v - \tilde{m}_2)\right] -
$$

$$
- E \left[I_{\tilde{m}_2 \leq v - k} \cdot (v - k - \tilde{m}_2)\right]\right\}
$$

(B-3)

where

$$
I_{\tilde{m}_2 \leq x} = \begin{cases} 
1 & \text{for } \tilde{m}_2 \leq x \\
0 & \text{for } \tilde{m}_2 > x
\end{cases}
$$

(B-4)

Now we calculate $\Delta(v) - \Delta(\tau)$ for $v > \tau$:

$$
\Delta(v) - \Delta(\tau) = \delta \cdot (A_1 - A_2).
$$

(B-5)

with $A_1 \equiv v \cdot E \left[I_{\tilde{m}_2 \leq v}\right] - \tau \cdot E \left[I_{\tilde{m}_2 \leq \tau}\right] - E \left[\tilde{m}_2 \cdot I_{\tau < \tilde{m}_2 \leq v}\right]$ and $A_2 \equiv (v - k) \cdot E \left[I_{\tilde{m}_2 \leq v - k}\right] - (\tau - k) \cdot E \left[I_{\tilde{m}_2 \leq \tau - k}\right] - E \left[\tilde{m}_2 \cdot I_{\tau - k < \tilde{m}_2 \leq v - k}\right]$. Let $G(.)$ denote the distribution of $\tilde{m}_2$, that is, $G(x) = \Pr[\tilde{m}_2 \leq x]$. We can now determine the lower and upper bound for $A_1$ and $A_2$, respectively:

$$
A_1 \geq v \cdot G(v) - \tau \cdot G(\tau) - v \cdot [G(v) - G(\tau)] =
$$

$$
= (v - \tau) \cdot G(\tau) \text{ and}
$$

$$
A_2 \leq (v - k) \cdot G(v - k) - (\tau - k) \cdot G(\tau - k) -
$$
Then, we get

\[
\Delta(v) - \Delta(\tau) \geq \delta \cdot (v - \tau) \cdot [G(\tau) - G(v - k)]
\]  

(B-7)

Since \( G \) is increasing in \( v \), we can choose \( \tau \) sufficiently close to \( v \) so that \( v - k < \tau \). This yields

\[
\Delta(v) - \Delta(\tau) \geq 0
\]  

(B-8)

implying that \( \pi(K, v) - \pi(k, v) \) is an increasing function of \( v \). Hence, the choice of technology is characterized by a cut-off rule.

**C Proof of Proposition 3**

We prove Proposition 3 by using a series of lemmas.

**Lemma C1.** In any pure strategy equilibrium, \( \hat{v} > v^*(0) = m^*(0) \).

**Proof.** Suppose not, i.e., \( \hat{v} \leq v^*(0) \). Then, the optimal strategy for the official in the second period is to demand \( m_2 = v^*(0) = m^*(0) \). This implies that the critical type \( \hat{v} \) gets no surplus in the second period with his overall payoff of \( \hat{v} - K - m_1 \). The critical type \( \hat{v} \) can achieve a higher payoff of \( \hat{v} - k - m_1 \) with the deviation of choosing the \( k \)-technology in the first period, which is a contradiction. \( Q.E.D. \)

Lemma C1 implies that for \( m_2 \) to be the optimal pure strategy choice for the official in the second period, \( m_2 \) should be either \( \hat{v} \) or \( m^*(k) \). More precisely, the relevant state variable for the official is the cut-off value \( \hat{v} \) of the entrepreneurs’ types who have chosen the \( K \)-technology. Let \( \hat{v}^o \) be defined as the larger root to the equation \( \hat{v}^o \cdot F(\hat{v}^o) = m^*(k) \cdot F(k + m^*(k)) \). Then, if \( \hat{v} > \hat{v}^o \),
$m_2 = m^*(k)$ is the optimal demand and if $\hat{v} < \hat{v}^o$, $m_2 = \hat{v}$ is the optimal demand in the second period.

We consider two cases depending on whether $m_2$ is $\hat{v}$ or $m^*(k)$. The following lemmas, however, show that neither of them can be part of a pure strategy equilibrium. We consider the two cases separately.

**Case 1.** $m_2 = m^*(k)$.

*Lemma C2.* For $m^*(k)$ to be the pure strategy optimal demand in the second period, there should be no entrant with the $k$-technology in the first period.

*Proof.* For $m^*(k)$ to be the pure strategy optimal demand in the second period, we must have $m^*(k) + k = v^*(k) < \hat{v}$. Suppose to the contrary that there are entrants with the $k$-technology. Then, there should be an entrant with the $k$-technology whose type belongs to $(v^*(k), \hat{v})$. This type, however, will prefer to enter with the $K$-technology rather than making repeated short-term investments in the $k$-technology when the second period demand is $m^*(k)$. Once again, we have a contradiction. *Q.E.D.*

*Lemma C3.* For $m^*(k)$ to be the pure strategy optimal demand in the second period, $m_1 - m^*(k) > (1 + \delta) \cdot k - K > 0$ and $\hat{v} = m_1 + (K - \delta \cdot k)$.

*Proof.* Lemma C2 implies that with the demand of $m_2 = m^*(k)$ to be optimal in the second period, the critical type $\hat{v}$ should satisfy the following condition:

$$(\hat{v} - K - m_1) + \delta \cdot [\hat{v} - m^*(k)] = \delta \cdot [\hat{v} - k - m^*(k)].$$  (C-1)

The LHS of (C-1) is the type $\hat{v}$ entrepreneur’s payoff of entering in the first period with the $K$-technology and the expression in the RHS of (C-1) represents the type $\hat{v}$’s payoff of delaying
his entry until the next period in which he enters with the k-technology. Eq. (C-1) defines the relationship between the critical type $\hat{v}$ and the first-period demand $m_1$ when $m_2 = m^*(k)$, which is given by $\hat{v} = m_1 + (K - \delta \cdot k)$. By substituting $\hat{v} = m_1 + (K - \delta \cdot k)$ into the condition that $m^*(k) + k = v^*(k) < \hat{v}$ yields $m_1 - m^*(k) > (1 + \delta) \cdot k - K > 0$. Q.E.D.

Lemma C4. There is no pure strategy equilibrium in which $m_2 = m^*(k)$.

Proof. Given $m_2 = m^*(k)$ in the second period, the corrupted official’s optimization problem is equivalent to the maximization of

$$R(\hat{v}) = [\hat{v} - (K - \delta \cdot k)] \cdot F(\hat{v})$$

subject to $\hat{v} \geq \hat{v}^o$ (i.e., $\hat{v} = \hat{v}^o$). The official charges $m_1 = \hat{v}^o - (K - \delta \cdot k)$ in the first period. The overall payoff of the corrupted official in any pure strategy equilibrium with $m_2 = m^*(k)$ thus is given by:

$$[\hat{v}^o - (K - \delta \cdot k)] \cdot F(\hat{v}^o) + \delta \cdot m^*(k) \cdot F[k + m^*(k)].$$

(C-3)

In contrast, if the corrupted official charges $m_1 = m^*(k)$, he can guarantee at least the payoff of $(1 + \delta) \cdot m^*(k) \cdot F[k + m^*(k)]$. Since $m^*(k) \cdot F[k + m^*(k)] = \hat{v}^o \cdot F(\hat{v}^o) > [\hat{v}^o - (K - \delta \cdot k)] \cdot F(\hat{v}^o)$, $m_1 = \hat{v}^o - (K - \delta \cdot k)$ and $m_2 = m^*(k)$ cannot be part of pure strategy equilibrium. Q.E.D.

Case 2. $m_2 = \hat{v}$.

Lemma C5. There is no pure strategy equilibrium in which $m_2 = \hat{v}$, either.
Proof. In this case, by proceeding exactly the same way as in the proof of Lemma C1, we can show that the critical type $\hat{\nu}$ can strictly increase his payoff by deviating with the $k$-technology, deriving a contradiction. Q.E.D.

The five lemmas taken together prove that there is no pure strategy equilibrium in the second period demand.

References


Figures

Choi / Thum

RJE

Fig. 1 of 2

---

Figure 1: The Timing of Repeated Extortion Game

Official Demands | Firms Enter | Official Demands | Firms Enter

\[ m_1 \quad v_1 \quad m_2 \quad v_2 \]

First Period | Second Period
Figure 2: Equilibrium with Endogenous Choice of Technologies