Can Self Reporting Reduce Corruption in Law Enforcement?*

Alfredo Burlando†, Alberto Motta‡

Abstract

This paper analyses the impact of self reporting on law enforcement when officers are corruptible. The threat of corruption highlights two additional advantages to the use of self reporting. First, by allowing individuals to self report their unlawful act, the government is able to increase welfare by eliminating officers’ rents and the variations in their wages. Second, due to the reduction in corruption costs, the introduction of self reporting allows governments to fully eliminate corruption.

JEL Classification: K42, L14, O10, D82

Keywords: self reporting, permits, law enforcement, corruption, incentives, leniency program, collusion.

1 Introduction

The literature on law enforcement has long noted the widespread presence of self reporting of criminal acts. Offenders choose to admit their misdeeds directly when they know the government will be lenient to them: they are better off paying a reduced fine for certain

---

*We are especially grateful to Dilip Mookherjee and Antonio Nicolò for their help, suggestions and supervision. We also wish to thank Kevin Lang, Michael Manove and seminar participants at Boston University, University of Padua and Asset 2007 Conference for their useful comments. Any remaining errors are ours. Alberto gratefully acknowledge financial support from the University of Padua (research grant CPDA071899).
†Boston University, Department of Economics, 270 Bay State Road, Boston, USA E-mail: burlando@bu.edu
‡University of New South Wales, School of Economics, Australian School of Business, Sydney 2052, Room Number: 3124, tel +61 293859771 fax +61 293136337 email: motta@unsw.edu.au
rather than facing a harsher sentence if detected. The government also prefers when law-breakers report their own crimes because it reduces enforcement costs. The theory behind self-reporting in law enforcement has already been explored by Malik (1993) and Kaplow and Shavell (1994) (henceforth KS). Using the Becker model (1964), KS demonstrated that self-reporting allows the government to save money by reducing the number of officers needed to monitor the population. These cost savings hinge on one assumption, namely the enforcement agency can easily separate those who reported their crime from the rest of the population. This allows officers to concentrate their attention on the non-reporting population. Since the monitored population has shrunk, fewer officers are employed for a given level of deterrence. When this assumption holds, KS finds that self-reporting benefits law enforcement. Otherwise, self-reporting may not be beneficial.

Despite these restrictions, self-reporting is often found in areas of law enforcement that violate this assumption. Hunting licenses in African national parks, nature reserves or game reserves provide one such compelling example. In many protected areas licensed hunting is introduced as a way for park authorities to cull wild animal populations in a controlled manner. However, in several African countries like Malawi, Tanzania and Zambia, hunting permits are used as a form of self-reporting for a crime - poaching - that is difficult to control. For instance, in 1998 Tanzania introduced animal hunting licenses in game reserves and national parks during a period of falling wild animal populations, intense poaching and worries about enforcement efficacy. Since then, poaching has decreased dramatically and wild animal populations have recovered (Lamotte, 2008). In contrast, poaching in Kenya (where there are no licenses) has continued at sustained levels, and animal populations in the national parks continue to decline.

Despite the wide adoption of hunting permits in African parks, this form of self-reporting

---

1 See also Polinsky and Shavell (2000) for a more general discussion on the law enforcement theory in economic literature.

2 In Tanzania it is possible to legally hunt the vast majority of wild animals, if one is willing to pay the right price. A license to shoot a lion costs $40,000 and one for a leopard $12,000. The 1998 reform legalized hunting for local populations as well, with licenses selling at a much lower cost.
cannot be justified by the standards set in KS. Hunters who purchase licenses can legally shoot wildlife within the park, but rangers cannot easily distinguish them from poachers. Because of this, the licenses do not affect how the park rangers patrol, and do not reduce the number of rangers needed. Why then do we observe such instances of self reporting in situations where enforcement cannot be reduced?

Our answer to that question is that self reporting has another important advantage than the ones highlighted by KS: self reporting can be used as a tool against corruption in law enforcement\(^3\). In this paper, we consider the following problem in a Beckerian model of probabilistic enforcement: apprehended wrong-doers can avoid official sanctions by bribing the police, thus challenging legal enforcement and limiting deterrence. To prevent this, the government may be compelled to offer incentives designed to keep enforcers honest. We focus our attention on those situations where the provision of such incentives is costly to the government: corruption can be eliminated but it is costly to do so. In certain cases, these incentives are onerous enough that a social maximizer would prefer a low-powered incentive system that allows corruption, even if this would lead to weaker enforcement and higher crime levels.

Self reporting provides an alternative to the offender, who can avoid paying either bribe or the full legal sanction by admitting culpability and paying a reduced punishment directly to the government. When the individual chooses this alternative, he avoids the law officer entirely, who - by being excluded - can claim neither bribe nor bonus. Because in a self reporting equilibrium officers do not earn rents, the amount of incentive they are offered is irrelevant. Thus, we show that the social planner can offer strong incentives at no cost, replace a corrupt regime with a clean one, and thus increase welfare through the improvement

---

\(^3\)Whether the difference in the policy outcomes between Tanzania and Kenya can be attributed to an increase in patrolling efficiency and corruption reduction is an issue that is worth exploring in further detail. That some legalization of wildlife hunting and trade can be helpful and beneficial for conservation is a concept that is not lost for several environmental groups. In their 2008 joint report, TRAFFIC and the WWF state that "Policies that criminalize the wild meat trade have not been effective to bringing it under control...greater consideration of alternative management scenarios, including legalizing hunting and trade of certain wild species for meat is therefore required" (Roe, 2008).
in deterrence. In other words, the planner can use self reporting as a way to clean up corrupt enforcement agencies.

It is important to stress that our result requires a tradeoff between enforcement costs and corruption. In a standard beckerian model this tradeoff might not exist: the government can offer very high incentives to officers who report criminals, and make up the expense by lowering the wages of those officers who do not. In that context, corruption is never a welfare-maximizing equilibrium. Since corrupted law enforcement agencies are quite common across the world, it is natural to think that eliminating corruption is not so easy. Here, we present two simple modifications to the basic model that produce the desired tradeoff. First, officers could be risk averse: they must be compensated for taking on a salary that is ex-ante uncertain. In this context (referred as Model A in section 2.1), self reporting eliminates uncertainty, and the risk premium with it. Second, officers' salary could be subject to a "limited liability" constraint - meaning, for instance, that they cannot accept negative wages. If this constraint is binding, we show in section 2.5 (Model B) that the government must transfer rents to keep officers honest, even if they are risk neutral. In that context, self reporting deprives officers of these rents.

In practice, self reporting is used by some companies in developing countries as a way to fight corruption. In India, the *tatkal* system was introduced by the phone company to provide people with an alternative to paying bribes for quick phone line installations. The Bharat Sanchar Nigam company guaranteed speedy connections to those willing to sign up and pay into the program. The same *taktal* system was later adopted by the railways as a way for customers to avoid long waiting lists or paying bribes to intermediaries for a quick ticket.

There is a small but growing literature on the economics of corruption in law enforcement.\(^4\) Polinsky and Shavell (2001) consider a setup similar to ours but do not explicitly model the source of the deadweight losses from corruption, and do not explore the role of

self reporting in corruption. Perhaps closest to our paper is Andrianova and Melissas (2008). Using a different setup, they study when governments choose to legalize an activity when officers are corrupt. They assume that once the activity is legalized, the government imposes no reduced fines or taxes, and employs no enforcement. Their paper thus sidesteps the effect of legalization on officers’ pay structure.

Self reporting has been found to have other important properties that are not discussed here. Motta and Polo (2003) show that leniency programs deter long term illegal relationships by creating a prisoner’s dilemma. In this regard, Buccirossi and Spagnolo (2005) show that leniency programs could help sustain occasional sequential illegal transactions. Another advantage, first highlighted by KS, is that self reporting reduces the risk borne by miscreants: in the absence of self reporting, individuals would face uncertain sanctions. Our Model A speaks to that insight by showing how it can be helpful in reducing the risk borne by officers. Innes (1999) shows that self reporting enjoys an additional economic advantages if ex-post benefits of remediation are considered. In this case, the violator can undertake remediation in order to reduce the harm caused. Finally, Innes (2001) argues that self reporting may also be advantageous when violators can engage in "avoidance" activities.

The rest of the paper is organized as follows: section 2 discusses the model; section 3 discusses the implications of some of the assumptions made; section 4 concludes.

2 Model A: Risk Averse Officers

2.1 Structure

There is a measure 1 of risk-neutral citizens who cause a harm to society of \( h \) if they commit an unlawful act or crime. Each citizen derives a private gain \( x \) from committing the act; the gain is distributed with a continuous density function \( g(.) \) and cumulative distribution function \( G(.) \). To minimize the number of crimes the government employs a measure \( p \) officers in the police force\(^5\). They are responsible for monitoring the population and reporting

\(^5\)We refer to the police officer and the citizen respectively as "she" and "he".
violators to a court of law, which in turn imposes fines.

2.1.1 Enforcement and corruption

Monitoring by the police force is costly. The government pays an income $v$ to each officer, which may consist of a base wage $w$ and an incentive $i$. Officers are risk averse individuals with a continuous and twice differentiable utility function $u(.)$ of the expected utility form, such that $u'(.) > 0$ and $u''(.) < 0$. Furthermore, they have reservation utility $u(\bar{v})$ which they must receive in order to remain in the force. Finally, they are potentially corruptible, meaning that they will consider accepting bribes.

Each officer is randomly matched with a citizen and learns whether he is an offender. When an offense is uncovered, the officer can report the violation to the judiciary system, which levies a fine $f$ to the offender. Alternatively, the offender may offer a bribe $b$ to the officer in exchange for her silence. The size of the bribe is reached through a bargaining process between the miscreant and the officer, and we make the following assumption:

**Assumption:** the maximum agreeable bribe cannot be larger than a fraction $\sigma \in (0, 1)$ of the fine.

This assumption simply limits the size of the bribe and has several justifications. First, even if miscreants’ wealth is greater or equal to $f$, they might be unable to transfer more than a fraction $\sigma f$ of this wealth in bribes. This could come from the fact that the fraction $(1-\sigma)f$ is held in assets that are not easily transferrable, or whose ownership is hard to verify at the time of the bribe transfer, such as boats, landholdings, or residential units located in another location or country.\(^6\) Second, the assumption would arise in a model where the transfer of assets itself could be detected by other authorities in the public administration.\(^7\)

\(^6\)While a corrupt officer might be unable to take possession of these assets, a government can reasonably be expected to find, confiscate, and sell these items.

\(^7\)It is common in the literature on corruption to assume that the probability of detecting bribe exchanges is a choice variable of the government. We abstract from it in order to keep the model simple. For an analysis of endogenous bribe detection, see Mookherjee and Png (1995), Polinsky and Shavell (2001). In Adving and Moene (1990), the probability of detection depends on how corrupt the overall system is. Overlapping responsibilities among different bureaucrats may also increase the chances of detection and the size of bribes, since officials may need to solve a coordination game (Rose-Ackerman 1994). See Bhardan (1997) for an insightful review.
In that model, the collusive agreement would be uncovered, the officer would lose the ill-
gotten wealth (through administrative fines or bribes to other officials), and the offender is
made to pay the full fine $f$. Third, the assumption is consistent with a model where part of
the punishment is non-monetary in nature, wherein only a fraction $\sigma$ of the punishment is
collected in monetary form. The remaining part of the punishment would be administered via
revenue-neutral imprisonment\(^8\) or other forms of coercion.\(^9\) Finally, the assumption would
arise in a model where bribe exchanges suffer transaction costs.\(^10\)

As an anti-corruption measure, the government can offer bonuses or incentives $i$ that are
paid when an officer reports an offender. Because acceptance of the bribe implies foregoing
the incentive, the officer will consider only bribes that have the following characteristic:

\[ i \leq b \leq \sigma f. \]

Finally, consider the choices available to an offender: if he is caught, he will prefer the
payment of the bribe when such payment is less than the fine, i.e. $b \leq f$.

The agreement on the bribe is reached through a bargaining process. Regardless of how
the bargain power is distributed between the two parties, to fully eliminate corruption the
government must set incentives high enough according to the following no-collusion condition:

\[ i \geq \sigma f \equiv i^{nc}. \] \hspace{1cm} (1)

\(^8\)Our main result would not be affected if we were to assume that imprisonment is costly for the gov-
ernment. Quite the opposite, assuming costly imprisonment would strengthen our results by highlighting a
further advantage to the use of self-reporting, namely, reducing length and cost of jail sentences.

\(^9\)For instance, denote as $l$ the monetary equivalent of lifelong incarceration, and denote as $m$ the offender’s
budget. For example, if $f$ is maximal, i.e., $f = l + m$, then $\sigma = \frac{m}{l+m}$. The model under this third explanation
evolves somewhat differently from what is explained in this paper, although all the results are maintained.
For simplicity, we work under assumption of complete detection.

\(^{10}\)A more traditional approach would have the parameter $\sigma$ reflect the level of transactional imperfections
between the colluding parties (see Tirole [1992] for a review of this transaction cost strand of the literature).
For instance, the offender incurs a cost $(1 + \tau)b$ in transferring a bribe $b$ to the inspector. In this case, once
the offender is caught, he will consider offering a bribe only if $b \leq \frac{1}{(1+\tau)} f$, where $\sigma = \frac{1}{(1+\tau)}$. The transaction
cost $\tau b$ is equivalent to a deadweight loss in terms of social welfare, requiring a slight modification of the
model presented here. In the framework we present, it can be shown that a transaction cost assumption
does not introduce a tradeoff between enforcement cost and criminality, and that corruption would never be
allowed by a welfare maximizing government.
2.1.2 Government expenditures and social welfare

Government expenditures $B$ consist of wage payments net of revenues from fines. $B$ is raised through distortionary taxation from the citizenship. Taxes cost taxpayers $(1 + \lambda)B$, where $\lambda$ is a parameter that measures the size of the dead weight loss. When $\lambda = 0$, resources are costlessly shifted from taxpayers to the government, resulting in no welfare losses; when $\lambda > 0$, there is a welfare loss equivalent to $\lambda B$. The government maximizes a weighted average of the welfare of all members of society, which include both civilians and enforcement officers.

2.1.3 Equilibrium considerations

In this simple model with homogeneous officers, there are only extreme outcomes. When $i < i^{nc}$, all officers accept bribes, and all criminals offer them. Since delinquents are never reported, the government neither pays incentives, nor levies fines. When $i \geq i^{nc}$, no bribes are ever exchanged, and therefore all violations are reported, all fines are levied, and all incentives are paid. We label these two states by the index $j = c, nc$. Citizens choose to commit the unlawful act only when their private benefit exceeds the expected sanction, which can be either the expected fine or the expected bribe. Enforcement determines a threshold level of gain $\hat{x}_j$ such that only individuals whose private benefit exceeds $\hat{x}_j$ commit the act. In this case, the harm to society due to criminality is $[1 - G(\hat{x}_j)] h$. The social welfare in state $j$ is

$$W_j = \int_{\hat{x}_j}^{\infty} (x - h)g(x)dx - \lambda B_j \quad \text{for } j = c, nc. \quad (2)$$

The canonical approach in KS or Polinsky and Shavell (2001) doesn’t consider a government budget which include the revenues from fines collected and the effect of distortionary taxation. From this perspective, our framework is similar in nature to setups more commonly found in the literature of corruption, such as in Besley and McClaren (1993) and Laffont and Tirole (1993). This departure is justified by the nature of the problem we are analyzing: corruption affects efficiency and depends on incentives, and our paper analyzes the trade-offs...
between costly incentives and police force efficiency.

In the next section, we consider the welfare functions for \( j = c, nc \) under the assumption that self reporting is not possible. The government can affect the state \( j \) and the level of welfare \( W_j \) by changing the incentive paid to officers and their number \( p \).

### 2.2 No Self Reporting

#### 2.2.1 Clean Regime

Consider the choice faced by a citizen when the regime is clean, i.e. officers are never corrupted (because they are paid high incentives, \( i \geq i^{nc} \)). Since his probability of being audited is \( p \), by committing the act he expects to pay a fine of \( pf \). This determines the threshold gain from the act, \( \hat{x}_{nc} \):

\[
\hat{x}_{nc} = pf.
\]

The threshold determines the number of unlawful acts, \( 1 - G(pf) \), and is also the probability that a given audit uncovers an offender. The payment structure must allow officers to earn their reservation utility, so that the following participation constraint must be met:

\[
[1 - G(pf)] u(w_{nc} + i) + G(pf)u(w_{nc}) \geq u(\bar{v})
\]

Next, consider the government’s decision. The government hires \( p \) officers, pays them \( v^{E}_{nc} \) on average, and collects fines from caught miscreants, \( f \). Accounting for wage expenses (which include incentive payouts) and fine income, the social welfare function is a function

---

\(^{11}\)The fine \( f \) is also a policy instrument that is optimally set by the government. As in the rest of the enforcement literature, the fine is always maximal: \( f \) is set to coincide with the wealth of the citizen. For a proof, Becker (1968), Kaplow and Shavell (1994), Polinsky and Shavell (2001). There are some exceptions to this: Malik (1990) shows that when criminals engage in detection avoidance, fines are not maximal. Livernois and McKenna (1999) claim that under plausible condition higher compliance rates are achieved with lower fines. However, Ines (2001) shows that when self reporting is introduced, there is no need to engage in avoidance and therefore the Becker principle of maximal fines applies again.

\(^{12}\)A modified version of costly imposition of fines is found in Polinsky and Shavell (2000).
of $p$, $w_{nc}$, and $i$:

$$W_{nc}(p, w_{nc}, i) = \int_{pf}^{\infty} (x - h) g(x) dx - \lambda p \{w_{nc} + (i - f)[1 - G(pf)]\}. \quad (4)$$

subject to the no-collusion bonus pay constraint (1) and the participation constraint (3).

Note that the only effect of wages and incentive pay is to increase the government budget. Enforcement is neither helped nor worsened by high payouts to officers. Thus, welfare is maximized when the average income earned by officers is minimized. It is straightforward to show that, due to the concavity of the utility function, this happens when both constraints (1) and (3) bind. At the optimum, the incentive is $i^* = i^{nc} = \sigma f$ and $w^*(p; \bar{v})$ is the base wage that satisfies (3) with equality. The objective function to be maximized is then reduced to a function of $p$ only:

$$W_{nc}(p) = \int_{pf}^{\infty} (x - h) g(x) dx - \lambda p \{w^*(p; \bar{v}) + f(\sigma - 1)[1 - G(pf)]\}. \quad (5)$$

The three terms in the welfare function are the following: the first term is the welfare loss due to illegal acts; the second term is the social cost of hiring $p$ officers which are provided a base wage equal to $w^*(p; \bar{v})$; the last term is the net revenues to the government, that is, fine income minus incentive pay.

### 2.2.2 Corrupt Regime

Suppose that incentives are too low, i.e. $i < \sigma f$: in this case there is scope for bribe exchange and the regime is corrupt. The agreement on the bribe is reached through Nash bargaining, with weights of $\mu \in (0, 1)$. The equilibrium bribe is given by the following expression:

$$b = \min \{(1 - \mu)i + \mu f, \sigma f\}. \quad (6)$$
The potential law breaker’s choice is either to do nothing, or commit the act and pay with probability $p$ a bribe $b$ as defined by (6). The threshold condition is

$$\hat{x}_c = pb.$$  \hspace{1cm} (7)

Officers have two sources of income: their base wage and the bribes. In this scenario, the government pays only base wages to its officers and receives no fine income, so the government budget is simply $pw_c$. As before, the wage received by officers need to satisfy a participation constraint, which now takes the following form:

$$[1 - G(pb)] u(w_c + b) + G(pb)u(w_c) \geq u(\bar{v})$$ \hspace{1cm} (8)

The welfare function takes the form

$$W_c(i, p) = \int_{pb}^{\infty} (x - h)g(x)dx - \lambda pw_c$$ \hspace{1cm} (9)

Subject to constraints (6), (8) and for values $i < \sigma f$.

**Claim 1** *In the corrupt regime, the incentive is set to a value such that the bribe is maximal ($b = \sigma f$) and officers’ participation constraint binds with equality at a wage $w_c^*(p; \bar{v})$*

**Proof.** For a given value of $p$, $b$ and $w_c$, consider a corrupt regime where $i < \sigma f$. A small increase in $i$ induces the equilibrium bribe $b$ to increase to $\tilde{b}$ according to (6), and the participation constraint (8) slackens. The government can then choose to hire fewer officers $\tilde{p}$ such that $\tilde{p}\tilde{b} = pb$, and keep the pay $w_c$. With fewer officers, the wage bill is reduced without affecting the level of criminality. Thus, for any wage $w_c$ that satisfies (8), $i$ is set to a value such that $b = \sigma f$.\(^{13}\) Next, note that the participation constraint must be biding with equality. Otherwise, the government can reduce $w_c$ without changing either $b$ or $p$, and the wage bill is reduced in a way that does not impact the amount of criminality. $\blacksquare$

\(^{13}\)The set of values $i^*$ such that $b(i^*) = \sigma f$ is $\left[ \max \left[ 0, \frac{\sigma - \mu}{1 - \mu} \right], \sigma f \right]$.  

11
Note that, in this case, incentives are offered but never paid out. However, they do play a role: they raise the equilibrium bribe and therefore increase the compliance threshold (7). In the optimum, the government increases $i$ to the point where bribes are maximal ($b = \sigma f$).

The corrupted loss function is then expressed in terms of the number of officers $p$ only:

$$W_c(p) = \int_{\sigma pf}^{\infty} (x - h)g(x)dx - \lambda pw_c^*(p; \bar{v}). \quad (10)$$

### 2.2.3 When is the corrupt regime optimal?

The government can move from one regime to the other by modifying the incentive pay $i$. The final outcome is determined by maximizing both $W_{nc}$ and $W_c$ in equations (5) and (10), and then choosing whichever is larger. When choosing which regime to adopt, the government essentially trades off lower agency costs (lower wages) in a corrupt regime, with the revenue that is collected when the regime is clean. A regime based on bribes is desirable in that having more criminals effectively reduces officers’s wage, since $w_c^*(p; \bar{v}) \leq w^*(p; \bar{v})$.\(^{14}\)

On the other hand, a clean regime brings about extra revenues: When $\sigma$ is small, officers collect in fines substantially more than they cost in incentive pay. But when $\sigma$ is large, fines collected can barely cover the expense of keeping officers honest.\(^{15}\) Depending on the parameters of the model, either regime could prevail.

Aside from the profitability of officers, the government also finds that bribes are less desirable because they limit deterrence: for any given $p$, there are more criminals in a corrupt regime than in a clean one. Still, when $\sigma$ is high, bribe amounts are high, and that provides a relatively large deterrence. Hence, a society characterized by high $\sigma$ could be better off adopting an enforcement system based on bribes.

\(^{14}\)Notice that in the optimum we have $i = b = \sigma f$. Therefore, $u_c^E = w_c + \sigma f[1 - G(\sigma f)]$ and $u_{nc}^E = w_{nc} + \sigma f[1 - G(pf)]$. A fast inspection reveals that the probability of apprehending a criminal is higher in the corrupted regime, $1 - G(\sigma f) \geq 1 - G(pf)$. In equilibrium, the base wage that makes the officer indifferent between accepting the job or not must be lower in such a regime, i.e., $w_c^*(p; \bar{v}) \leq w^*(p; \bar{v})$ for a given value of $p$.

\(^{15}\)Governments often exert great effort to keep highly profitable enforcement agencies (such as tax administrations and customs) as honest as possible, even amid the widespread corruption of other forces (i.e., police forces and anti-crime units).
Whichever regime is chosen, officers benefit from their corruptibility, since they are able to supplement their wage with rents in the shape of bonuses or bribes.

2.3 Self Reporting

We now introduce self reporting in both regimes $j = nc, c$. Following KS, we assume that an individual who committed an unlawful act may admit it before the inspection has taken place. As a reward for the admission, the judiciary imposes a reduced or discounted fine $r$, which is less than the fine $f$. By self reporting and paying $r$, the individual avoids paying either the fine $f$ or a bribe $b$. However, we diverge from the original Kaplow and Shavell model by assuming that those who self reported do not necessarily avoid enforcement audits, because officers cannot easily distinguish them from the general, non-self reporting population. By doing this, we abstract from the ‘enforcement reduction’ aspect of self reporting highlighted in their paper.

In general, the reduced fine follows a different judicial path than the full fine. A citizen who is reported to the judiciary by a law enforcer has the right to defend himself in a court of law, so demonstrating his culpability is time consuming and expensive. The same cannot be said when a person reports his unlawful act: by confessing, he gives up the right to proclaim his innocence. As a consequence, the bureaucratic procedure needed to process the reduced fine is more efficient and less expensive. This point, already highlighted by KS, adds to the appeal of self reporting. However, we show that the benefits highlighted in this paper do not depend on the superior bureaucratic efficiency of self reporting.

In a self-reporting regime, an individual who commits an unlawful act can either accept the chance of being caught and face a punishment $f_j$ (equal to $f$ in the clean regime or the bribe in the corrupt regime), or report himself and pay the reduced fine $r_j$ - whichever is more convenient (less expensive) in expectation to him. An individual with private gain $x$ may commit the act if his private benefit from the act exceeds the cost:
\[ x_j \geq \min[r_j, pf_j] \equiv \hat{x}^{sr}_j, \]

where \( r_j \) is now part of the set of policy instruments available to the government. To get the optimal level of \( r_j \), consider first the case \( r_j > pf_j \). Because self reporting is more expensive than the expected full sanction, criminals do not report their act. In this case, self reporting is possible, but no one employs it.

Now suppose \( r_j \leq pf_j \). All unlawful acts are reported by miscreants who pay only reduced fines to the government. Since \( \hat{x}^{sr}_j = r_j \), the total number of crimes committed is \( 1 - G(r_j) \). The welfare achieved is

\[
W_j^{sr}(r_j, p, w, i) = \int_{r_j}^{\infty} (x - h)g(x)dx - \lambda \left\{ p\nu_j^{sr,E} - r_j[1 - G(r_j)] \right\}, \tag{11}
\]

where \( \nu_j^{sr,E} \) is the average payout given to officers (wages and incentives). This welfare can be achieved subject to the usual constraints on officers that dictates that their expected utility from their earnings \( E(u(\nu_j^{sr,E})) \) under the self reporting regime cannot fall below the reservation utility \( u(\bar{v}) \).

Proposition 2 explains the optimal self reporting policy under the clean regime.

**Proposition 2** When the regime is clean and self reporting is adopted, all miscreants self report and:

(i) \( r_{nc} = pf \) (the reduced fine is equal to the expected full punishment when the regime is clean);

(ii) \( i \geq \sigma f \) (the no-collusion condition (1) is implemented)

(iii) \( \nu_{nc}^{sr,E} = \bar{v} \) (officers’ base wage is equal to their reservation wage).

**Proof.** see the Appendix. ■

With this proposition, we can replace \( r \) with \( pf \), and \( \nu_{nc}^{sr,E} \) with \( \bar{v} \). The welfare function
to be maximized becomes

\[ W_{nc}^{sr}(p) = \int_{pf}^{\infty} (x - h)g(x)dx - \lambda p \{ \overline{v} - f[1 - G(pf)] \} . \] (12)

The first term indicates the welfare loss due to crime; the second term represents the wage bill \( \overline{v} \) paid to officers; the third term is the revenues from self reporting.

The critical difference between a self reporting regime and one that does not rely on it is that in the latter officers must be compensated for bearing the risk of uncertain wages. By introducing self reporting, officers receive only their base wage, they do not observe wage uncertainty, and therefore they do not need to be compensated for the risk. In section 2.4 we prove that this is a welfare-enhancing difference.

Even though inspectors never uncover unreported miscreants, the presence of officers does provide a credible threat against those criminals who fail to self report: such criminals could get caught and subjected to pay the full fine \( f \). This mechanism resembles the ‘threat’ found in the enforcement of parking regulations. Cities allow drivers to park in certain areas only if they pay a certain reduced fine at the curb by feeding a meter. People feed the meter in the off chance that a parking inspector passes by and fines those who did not.

Proposition 3 explains the optimal self reporting policy under the corrupt regime.

**Proposition 3** When the regime is corrupted and self reporting is adopted, all miscreants self report and:

(i) \( r_c = pb \) (the reduced fine is equal to the expected bribe payment)

(ii) \( b = \sigma f \) (bribes are maximal)

(iii) \( i \in \left[ \max \left[ 0, \frac{\sigma - \mu}{1 - \mu} f \right], \sigma f \right] \).

(iv) \( v_c^{sr,E} = \overline{v} \) (officers’ base wage is equal to their reservation wage)

**Proof.** see the Appendix.  ■

With this proposition, the objective function reduces to a function of \( p \) only:
Again, the first term is the social loss due to criminality, and the term in parenthesis is the wage paid net of revenues from self reporting. Note that, in essence, the revenues from self reporting are in fact the bribes that are now diverted from the hands of corrupted officials into the state’s coffers. This bribe diversion has no direct enforcement effect, but it has a significant revenue effect. It also has no direct effect on the corruptibility of officials, who would take bribes if in the position to do so.

### 2.4 Comparative Statics

We are now able to compare the different regimes with and without self reporting, and demonstrate the main result of the paper: that self reporting in a clean regime is the best enforcement policy. To reach that conclusion, we show in the next two lemmas that self reporting is more efficient in both clean and corrupt regimes. We start by evaluating self reporting in a clean regime.

**Lemma 4** Self reporting improves welfare under a clean regime.

**Proof.** It is sufficient to show that for any $\bar{p}$ that is chosen under a policy without self reporting, a self reporting policy $\bar{r} = \bar{p}f$ yields greater social welfare. Denoting the welfare without self reporting by $W_{nc}(\bar{p})$ and the welfare with self reporting policy $\bar{r}$ by $W_{nc}^{sr}(\bar{p})$, the increase in welfare under self reporting is

$$W_{nc}^{sr}(\bar{p}) - W_{nc}(\bar{p}) = \lambda p \{ (w^*(\bar{p}; \bar{v}) + \sigma f[1 - G(\bar{pf})] - \bar{v}) > 0$$

To see this point formally, let us consider the no self reporting regime. Note that since $E(v) = v_{nc}^E = w^*(\bar{p}; \bar{v}) + \sigma f[1 - G(\bar{pf})]$, $E[u(v)] = u(\bar{v})$ and $u'' < 0$, Jensen’s inequality tells us that $u[E(v)] = u(v_{nc}^E) > u(\bar{v})$. But we know that in the self reporting regime
\( u(v_{nc}^{sr,E}) = u(\overline{v}) \), and so \( u(v_{nc}^{E}) > u(v_{nc}^{sr,E}) \). It follows that \( v_{nc}^{E} > v_{nc}^{sr,E} \); substituting we have that \( w^*(\hat{p}; \overline{v}) + \sigma f [1 - G(\hat{p}f)] > \overline{v} \). \hfill \blacksquare

The improvement is not necessarily caused by a reduction of crime rates. Indeed, the optimal policy under self reporting may or may not reduce the number of criminals with respect to the optimal policy under no self reporting. Nonetheless, for any given level of deterrence, self reporting allows the government to save in enforcement expenditures. This improvement is made from the incentives that are no longer paid out to officers. Since trasgressors report their harmful acts directly, they do not need to be inspected, and so officers do not receive any bonuses. The introduction of self reporting effectively accomplishes two goals: it induces criminals to report truthfully and it fully insures officers against income risk.

Next, consider the corrupted regime with and without self reporting.

**Lemma 5** Self reporting improves welfare under a corrupt regime.

**Proof.** For any \( \hat{p} \) that is chosen under a policy without self reporting, the government can introduce self reporting by choosing \( \hat{r}_c = \sigma \hat{p} f \). The gain in welfare is given by the difference between (10) and (13):

\[
W_c^{sr}(\hat{p}) - W_c(\hat{p}) = \lambda \hat{p} \{ w_c^*(\hat{p}; \overline{v}) + \sigma f [1 - G(\hat{p}f)] - \overline{v} \} > 0
\]

To see why the difference is strictly positive, consider the no self reporting regime. Note that since \( E(v) = v_c^{E} = w_c^*(\hat{p}; \overline{v}) + \sigma f [1 - G(\sigma \hat{p}f)] \), \( E[u(w)] = u(\overline{v}) \) and \( u'' < 0 \), Jensen’s inequality tells us that \( u[E(v)] = u(v_c^{E}) > u(\overline{v}) \). But we know that in the self reporting regime \( u(v_c^{sr,E}) = u(\overline{v}) \), and so \( u(v_c^{E}) > u(v_c^{sr,E}) \). It follows that \( v_c^{E} > v_c^{sr,E} \); substituting we have that \( w_c^*(\hat{p}; \overline{v}) + \sigma f [1 - G(\sigma \hat{p}f)] > \overline{v} \). \hfill \blacksquare

Moving to a regime of self-reporting allows the government to add an income stream: bribes from officers that now find their way into the hands of the government. The principle established for the clean regime then translates also to the corrupt regime: officers cannot
earn rents under self reporting.

The two lemmas show that self reporting improves welfare both under a clean and a corrupt regime. We now establish the third result: once self reporting is introduced, the clean regime always dominates.

**Proposition 6** When self reporting is introduced, for any policy that induces corruption, there exists another policy under a clean regime that is strictly preferable.

**Proof.** Consider a regime of corruption with self reporting, where the number of officers is $\tilde{p}$ and the reduced fine is $\tilde{r}_c = \sigma \tilde{p} f$. We now show that the government would be strictly better off if it eliminates corruption (by choosing $i \geq i^{ac}$), reduces the workforce from $\tilde{p}$ to $\tilde{p} = \sigma \tilde{p}$, and keeps the self reporting fine at $\tilde{r}_c$. The change in welfare is then

$$W_{nc}^{sr}(\tilde{p}) - W_{c}^{sr}(\tilde{p}) = \lambda \tilde{p} (1 - \sigma) \tau > 0$$

The fact that corruption is never optimal should not come as a surprise: the main reason for allowing corruption when there is no self reporting is that the government would forgo the expense of paying bonuses to its officials; but under self-reporting, bonuses are never paid, and in either regime officers earn their outside wage $v$ only. With the main benefit of corruption gone, what is left is the negative aspect of corruption, namely, that it reduces deterrence. But the level of deterrence under a corrupt regime can be achieved with lesser expense (fewer officers) under a clean one.

### 2.5 Model B: Limited Liability

So far we have assumed that the officers are risk averse, but we imposed no limits on how low their wages could fall. These are somewhat restrictive assumptions. The first assumption is violated if officers are risk neutral. Or, while they are risk averse, they audit many citizens, so that total earnings are subject to little variability. The second assumption implies that
officers can make negative wages, essentially "paying" the public administration a fixed amount for the right to collect bribes or incentive pay. While such situation has been documented (refer, for a number of interesting instances, to Besley and McClaren, 1993), it is not generally common: even very corrupt enforcement agencies pay a (sometimes nominal) wage to their employees, although in many countries you do find a practice of not paying civil servants for months on end.

In this section, we consider risk-neutral officers that are subject to a wage floor, and without loss of generality we set the floor at zero. The latter entails that the following limited liability constraints must be met:

\[ w_{nc} + \max(i, b) \geq 0, \]

\[ w_{nc} \geq 0, \tag{14} \]

where the first and the second inequality ensure that the officer receives at least 0 irrespective of her being matched with an offender or a honest citizen. Take the clean regime with no self-reporting. The incentive is \( i^{nc} \equiv \sigma f \) and the following participation constraint must be met:

\[ [1 - G(pf)] (w_{nc} + \sigma f) + G(pf)w_{nc} \geq \bar{v} \tag{15} \]

The participation constraint binds when

\[ w_{nc} = \bar{v} - [1 - G(pf)] i^{nc} = \tilde{w}_{nc}. \tag{16} \]

If \( w_{nc} = \tilde{w}_{nc} > 0 \), the limited liability constraint (14) is slack: the officer receives an expected income equal to her outside option and the government can eliminate corruption at no cost. Thus, corruption is not a problem and the rest of the model becomes trivial. However, if \( \tilde{w}_{nc} \) is negative the limited liability constraint (14) binds. In this case the government sets \( w_{nc} = 0 \), and has to forgo a rent equal to the absolute value of \( \tilde{w}_{nc} \). The objective function
to be minimized in the clean regime is then reduced to,

$$W_{nc}(p) = \int_{p_f}^{\infty} (x - h)g(x)dx + \lambda pf(\sigma - 1)[1 - G(pf)].$$  \hspace{1cm} (17)

In the corrupted regime without self reporting the government optimally increases $i$ to the point where $b = \sigma f$. As before, the wage received by officers need to satisfy a participation constraint, which now takes the following form:

$$[1 - G(p\sigma f)](w_c + \sigma f) + G(p\sigma f)w_c \geq \bar{v}$$  \hspace{1cm} (18)

The participation constraint binds when

$$w_c = \bar{v} - [1 - G(p\sigma f)] \sigma f \equiv \tilde{w}_c.$$

Recall that we are considering the case where $\tilde{w}_{nc} < 0$, i.e., in the clean regime the wage floor binds and corruption is costly to eliminate. This implies that $\tilde{w}_c < 0$ because $0 > \tilde{w}_{nc} > \tilde{w}_c$. In this case the government optimally set $w_c = 0$. The welfare function takes the form

$$W_c(i, p) = \int_{\sigma pf}^{\infty} (x - h)g(x)dx.$$  \hspace{1cm} (19)

Note that the assumption on the wage floor produces the usual tradeoff between enforcement efficiency under the clean regime, and low enforcement costs under the corrupt one, so a welfare-maximizing state could well choose one or the other.

Under the clean regime with self reporting, the welfare function to be maximized becomes

$$W_{nc}^{sr}(p) = \int_{p_f}^{\infty} (x - h)g(x)dx - \lambda p\bar{v} + \lambda pf[1 - G(pf)].$$

$$W_{nc}^{sr}(p) = \int_{p_f}^{\infty} (x - h)g(x)dx - \lambda p\bar{v} + \lambda pf[1 - G(pf)].$$
When the regime is corrupted and self reporting is adopted:

\[
W_{c}^{sr}(p) = \int_{\sigma_p f}^{\infty} (x - h)g(x)dx - \lambda p\bar{v} + \lambda p\sigma f[1 - G(\sigma pf)]
\]

Comparing the different regimes with and without self reporting it is easy to note that the main result of our paper still holds: self reporting in a clean regime is the best enforcement policy. To see this point note that

\[
W_{nc}^{sr}(p) - W_{nc}(p) = \lambda p ([1 - G(pf)]\sigma f - \bar{v}) > 0.
\]

The latter inequality holds because when (14) binds, from (16) follows

\[
[1 - G(pf)]\sigma f > \bar{v}.
\]  \hspace{1cm} (20)

The comparisons for the remaining regimes are straightforward. Denote as \(p^*\) the optimal number of officers in the corrupted regimes, and set \(p = \sigma p^*\) for the clean regime with self reporting. It is easy to see that,

\[
W_{nc}^{sr}(\sigma p^*) - W_{c}^{sr}(p^*) = \lambda p^* \bar{v}(1 - \sigma) > 0
\]

and

\[
W_{nc}^{sr}(\sigma p^*) - W_{c}(p^*) = \lambda \sigma p^* (f[1 - G(\sigma p^* f)] - \bar{v}) > 0
\]

where the last inequality follows from (20).
3 Discussion

3.1 Weakening of enforcement effort

The first limitation of self reporting in law enforcement is that it may create moral hazard problems of its own among the officers. In our model, the probability of detecting a criminal act does not depend on the officers’ effort. In reality, the intensity of effort exerted is likely to change the chance that an unlawful act is uncovered. If all criminal acts were self reported, law enforcers would see no benefit in working hard, and this would reduce the probability of detection for everyone. How this weakening in enforcement impacts the overall equilibrium and the implementation of self reporting depends on how effort is modeled.

While this limitation may be substantial in some settings, it may not be as important in instances where either officer exertion is unimportant or it is easily monitored by the enforcement agency. Effort may be unimportant when the officer must perform many tasks, and only one of them is to check whether an individual has committed a certain crime. For example, customs officials at a port of entry perform a series of tasks on a random selection of incoming containers, such as ensuring that contents match the documentation. In the process, they may determine whether other regulations have been violated without making significant extra effort: whether all import duties have been paid, whether illegal substances or restricted materials are found. In other instances, where effort matters, the government can monitor effort. For example, many tasks can be standardized and reduced to checklists or forms that must be completed by the officers. Many instances of tax evasion are captured in this way, since officers must first of all check that forms sent from different sources match the income report.

3.2 Framing of innocent civilians

If officers receive a bonus every time they discover a criminal, they are more likely to abuse their position and frame the innocent to gain bonuses. In a normal setting without self reporting, the government either prevents corruption, or prevents framing, but not both at
the same time (see Polinsky and Shavell, 2001). Under self reporting, however, all criminals self report so that officers are always matched with innocent civilians. Since all crimes reported by the police are likely to be instances of framing, a government that can commit to investigate all officer-produced cases can stem framing at virtually no cost.

Complications may arise whenever the government cannot make such a commitment. For instance, suppose that the government needs to hire investigators in charge of detecting framing before framing happens (and, for simplicity, assume that such investigators are incorruptible). In that case, the probability of an instance of framing being detected depends on the number of such investigators and on the number of other framing cases. The framing-deterring Bayesian outcome is likely to require a positive number of investigators and a more complex choice of self reporting policies.

While a full analysis of framing under self reporting is beyond the scope of this paper, we suspect that self reporting may in fact incentivize the government to perform monitoring of framing. In equilibrium, all criminals self report, and officers receive incentive payments only from those innocent civilians they were able to frame. Investigators are, therefore, very efficient in detecting framing. Absent self reporting, framing cases will be mixed with legitimate cases, and agents in charge of monitoring officers are much less efficient in uncovering abuses.\footnote{An alternative way to avoid framing would be to offer efficiency wages to officers. The threat of detection of extortion or framing and the subsequent loss of the high income could deter officers from abuse of power. We leave for future research the analysis of such a model of law enforcement based on efficiency wages.}

### 3.3 Adverse selection of officers

A second aspect worth considering is adverse selection among officers. Officers may have different degrees of ability in performing their job: some may have a higher probability of uncovering offenders than others. Clearly, self reporting eliminates these differences, since the chance of encountering an unreported violation is zero for both ‘good’ and ‘bad’ officials. This may be a problem for the enforcement agency if selection is important in other aspects
of its activities. For example, the agency may want to observe individual ability so that it can promote good workers to higher ranks. In that case self reporting is still worth it if the agency has other means to measure ability.

### 3.4 Failure to self report

In practice, offenders often fail to self report even when such an option is available. We can think of four reasons for this. First, it may be that individuals have heterogeneous probabilities of detection, levels of risk aversion, or discount factors (Innes 2000, McCrary 2010). Second, a person may have more to hide than the crime itself: self reporting on one crime may lead investigators to audit more thoroughly other aspects of a person’s life, the cost of which is not ‘priced in’ the self reporting fine. Thus, a driver may prefer to ‘hit and run’ a bystander rather than stop to help if he is carrying a stash of drugs with him, for which he is guaranteed a harsh punishment. Third, individuals may face some uncertainty when self reporting due to the complexity of the law. He cannot be sure that, after self reported for one crime, he won’t be held accountable for another act which he did not think was illegal.

Whenever these circumstances arise in a way that creates additional heterogeneity in citizens’ preferences, some individuals do not self report. The implication is that in a clean regime the government cannot avoid paying incentives to some of its officers. This reduces the effect of self reporting against corruption. While proposition 5 may therefore be violated, we can show that self reporting remains the optimal policy under either clean or corrupt regime: this is because in either regime, the government can reduce the number of officers taking bribes or bonuses. Moreover, the introduction of self reporting still allows some governments (but not all) to fully eliminate corruption.\(^{18}\)

---

\(^{17}\) It is also possible that individuals actively engage in avoidance (Malik 1990, Innes 2001); however, unless individuals have different cost functions of avoidance, there is no reason for agents to avoid self reporting.

\(^{18}\) Notes available from the authors upon request.
4 Concluding Comments

We analyze the role of self reporting as an anti-corruption instrument in the practice of law enforcement when the enforcers are corruptible. Enforcement agencies which suffer from widespread corruption within the ranks are fairly common in many countries of the world, one reason being that cleansing can be painful and expensive. Our paper suggests that when reform aimed to eliminate corruption is implemented in conjunction with self reporting, some of the costs of reform can be eliminated. This is due to the self reporting propriety of reducing agency costs associated with corruption by eliminating rents to officers and variations in their wages.

5 Appendix

5.1 Proof of Proposition 2

In what follows, we prove the three statements in Proposition 1, i.e. when the regime is clean and self reporting is allowed: (i) \( r = pf \), (ii) \( i \geq \sigma f \) and (iii) \( v^{sr,E}_{nc} = \overline{\nu} \).

(i) Suppose \( r < pf \). Then, all criminals self report, and the government could slightly decrease \( p \) without changing the number of crimes. Neither the integral nor the last term in (11) would change. The second term would decrease if \( v^{sr,E}_{nc} \) is nonincreasing in \( p \). Since all offenders self report, there are no incentives paid and \( v^{sr,E}_{c} = \overline{\nu} \) which is independent of \( p \). Thus, welfare would increase. Since \( r < pf \) is not optimal, it must be that \( r \geq pf \). If \( r > pf \) self reporting is not binding, and therefore the welfare function to be maximized is equivalent to (5). That is, self reporting is not adopted. When \( r = pf \) self reporting is used by criminals and the welfare function is given by (11).

(ii) Suppose \( i < \sigma f \). An offender who does not pay the reduced fine \( r = pf \) expects to be apprehended with probability \( p \). In that case, the officer will accept the bribe \( b \) as defined by (6). It follows that the violator’s expected cost when not reporting himself equals \( p[(1 - \mu)i + \mu\sigma f] \leq p(\sigma f) < pf = r \). In this case, the regime is corrupt. Therefore it must be that \( i \geq \sigma f \).
(iii) When offenders self report, officers only inspect innocent citizens, and therefore they earn no incentive pay: the only salary paid is the base wage. To maximize equation (11), the government sets base wage as low as possible, to $\overline{v}$.

5.2 Proof of Proposition 3

In what follows, we show that when the regime is corrupted and self reporting is allowed: (i) $r_c = pb$, (ii) $b \to \sigma f$ and (iii) $v_{c}^{sr,E} = \overline{v}$.

(i) Suppose $r_c < pb$. Then, all criminals self report, and the government could slightly decrease $p$ without changing the number of crimes, which remain at $\hat{x} = r_c$. Neither the integral nor the last term in (11) would change. The second term would decrease if $v_{c}^{sr,E}$ is non-increasing in $p$. Since all offenders self report, there are no bribes and therefore the expected wage does not depend on $p$, i.e., $v_{c}^{sr,E} = \overline{v}$. Thus, reducing $p$ increases welfare, which means that $r < pf$ is not optimal. Hence, it must be that $r \geq pf$. When $r > pb$, no offender self reports, so self reporting is not adopted and welfare function (10) applies. It follows that self report binds when $r = pb$ and social welfare is given by (11). We can then rewrite (11) as:

$$W_c^{sr}(r_c, p, w, i) = \int_{pb}^{\infty} (x - h) g(x) dx - \lambda \{ p v_{c}^{sr,E} - pb [1 - G(pb)] \}$$

(21)

(ii) / (iii) To show that $b = \sigma f$, we need to show that the government wants to select $i$ from the non-empty set $\left[ \max \left[ 0, \frac{\sigma - \mu}{1 - \mu} f \right], \sigma f \right]$. Suppose that $i < \max \left[ 0, \frac{\sigma - \mu}{1 - \mu} f \right]$, so that $b = (1 - \mu)i + \mu f < \sigma f$. Now consider a raise in $i$ to $\bar{i}$, such that equilibrium bribe is $\tilde{b} > b$, and a corresponding decrease of $p$ to $\tilde{p}$ such that $\tilde{p} \tilde{b} = pb$. Then, the integral in equation (21) does not change, whereas the second term decreases. Since reducing $p$ increases welfare, $i$ cannot be optimal. On the other hand, if $i \geq \sigma f$ the regime is a clean regime. Thus, $b = \sigma f$ and $i \in \left[ \max \left[ 0, \frac{\sigma - \mu}{1 - \mu} f \right], \sigma f \right]$.

(iv) So far, all criminal acts are reported to the government directly, so officers do not earn any bribes. Wages $v_{c}^{sr,E}$ enter the welfare function (21) negatively, so in order to maximize welfare, the base wage needs to be as low as the reservation wage $\overline{v}$.

\[\text{Given that } \mu \in (0, 1), \sigma \in (0, 1), \text{ and } f > 0, \text{ it is easy to see that this set is non-empty.}\]
References


