Self-Reporting and Market Structure

Matthew D. Rablen[*] Andrew Samuel[†]
University of Sheffield Loyola University Maryland

December 13, 2018

Abstract

Many regulators utilize self-reporting to enforce regulations in a variety of market contexts. This paper studies the effectiveness of self-reporting within the context of an oligopoly. We identify two important consequences of implementing self-reporting (relative to non-reporting) for a welfare-maximizing regulator. First, when the regulator can only control the audit probability and fine, then whether compliance rises or falls upon implementing self-reporting, depends on the level of competition. Second, if the regulator can also control the market size, then the welfare maximizing policy entails self-reporting but with more competition and lower compliance than under no-reporting.

JEL Classification: K42; K32; D62.
Keywords: Law enforcement, self-reporting, Cournot oligopoly, market structure.

[*]Department of Economics, 9 Mappin Street, Sheffield, S1 4DT, UK.
[†]Department of Economics, Loyola University Maryland, 4501 N. Charles Street, Baltimore, MD 21210, USA.

1 Introduction

Many regulators utilize self-reporting to enforce their regulations (Innes, 2000). The U.S. Environmental Protection Agency (EPA) and the U.K. Environment Agency, for example, encourage firms to self-report environmental "crimes" such as spills of oil or of untreated sewerage. Similarly, the U.S. Department of Agriculture (USDA) and the Food and Drug Administration (FDA) have recently adopted self-reporting to regulate firms for compliance with food safety standards.[1]

The essence of self-reporting is that offenders are incentivized to self-report violations in exchange for weaker sanctions, whereas those who do not self-report face stricter sanctions if they are caught. Accordingly, self-reporting is beneficial to a regulator because it need not audit those who confessed to the crime, thereby saving on auditing costs. Indeed, formal analysis of self-reporting has shown that self-reporting can implement a given level of compliance at a lower cost (Malik, 1993; Kaplow and Shavell, 1994). These qualities make self-reporting an attractive policy in an era of smaller budgets for regulators.

Previous research that studies the efficacy of self reporting (Malik, 1993; Kaplow and Shavell, 1994; Innes, 1999, 2001), assumes a large number of (atomistic) agents or price-taking firms. That is, this literature implicitly assumes that regulators are monitoring firms that operate in perfectly competitive industries. We assert that this is not realistic because most regulation occurs in imperfectly competitive markets. For example, the EPA and FDA both regulate an oligopolistic energy and pharmaceutical industry respectively, and the USDA regulates an agricultural industry that is less than perfectly competitive.[2] Despite this, little is known about how self-reporting interacts with market structure; especially whether the effectiveness or impact of self-reporting varies with market structure.

The goal of this paper of this paper is to study the effectiveness of self-reporting under non-perfectly competitive markets; that is, monopolistically competitive and oligopolistic markets. The questions that we wish to address are: how does the optimal self-reporting policy vary by industry structure? Second, under what market conditions will self-reporting yield a higher level of compliance? Last, if a planner is unconstrained and can choose both

[1] See `https://www.epa.gov/compliance/epas-audit-policy`, `https://cedrec.com/news/index.htm?news_id=20823`, and `https://www.fda.gov/medicaldevices/deviceregulationandguidance/postmarketrequirements/reportingadverseevents/default.htm`. See also Toffel and Short (2011) for a discussion of various self-reporting policies.

[2] For example, according to U.S. Census data for 2007, the (C_4) concentration ratio for offshore drilling (which is regulated by the EPA) is 50. Similarly, the FDA regulates animal anti-microbials ($C_4 > 50$) and medical devices ($C_4 = 35$) (U.S. Census, 2007).

the level of enforcement as well as the size of the market, will implementing self-reporting give rise to more or less competitive markets?

To study these questions we develop a "Cournot-style" model in which oligopolistic firms generate a negative externality (e.g., environmental pollution) during production. Firms can reduce this harm by investing in costly abatement. However, since abatement is costly, in the absence of any regulation firms do not abate. To incentivize abatement firms are audited by a regulator who can choose either a self-reporting regime, or a no-reporting regime to fine firms for causing harm. Enforcement, through auditing firms with a given probability, is costly and these costs may either be fixed or variable in nature. Under a fixed cost structure, enforcement cost does not vary with firm size, whereas under a variable cost structure it does.

Analyzing this framework yields three important results. First, by utilizing self-reporting, a regulator can introduce welfare enhancing regulations in markets where regulation would be too costly under non-reporting; that is, without self-reporting the laissez faire policy is "optimal" in a second-best sense. Specifically, if enforcement costs are fixed, we show that there exists a threshold level of competition such that when competition is above this threshold, a regulator will prefer the laissez faire outcome over no-reporting because enforcement costs under no-reporting are too high. But, in this same situation, and indeed for *any level of competition*, the regulator would prefer regulation through self-reporting to the laissez faire outcome. A related result is obtained when enforcement costs are variable.

Second, if a regulator is constrained in that it cannot choose the level of competition, self reporting need not yield a higher level of compliance (relative to no reporting) even though it is always welfare enhancing. Specifically, if enforcement costs are fixed with respect to firm size, then enforcement and compliance is higher (lower) when the market is sufficiently competitive (concentrated). If, however, enforcement costs vary with firm size, then this result is reversed: enforcement and compliance is higher (lower) when the market is sufficiently competitive (concentrated). Thus, whether or not self-reporting yields a higher level of compliance depends on the level of competition. The nature of this effect is importantly mediated by the structure of enforcement costs: fixed or variable.

Third, if the regulator is unconstrained in that it can choose both the level of enforcement and the number of firms, then the regulator always chooses to favor more competition and lower level of compliance, relative to a no-reporting regime. Thus, self-reporting allows for a larger, more competitive, market with larger consumer surplus, but at the expense of lower compliance and greater harm.

It is insightful to relate these findings to the broader literature on self-reporting. The main benefit to self-reporting is that the regulator can save on enforcement costs (Kaplow and Shavell, 1994; Malik, 1993). Innes (1999, 2001) also identifies two further advantages to self-reporting. First, if firms can engage in clean-up activities, then under self-reporting firms always engage in clean-up, whereas, under no-reporting firms only clean-up when they are caught. Since clean-up is welfare improving, self-reporting improves welfare for this additional reason. Second, if firms can invest in costly detection avoidance, then under self-reporting there is less avoidance. Since avoidance is wasteful, self-reporting enhances welfare.

While these studies agree that self-reporting is welfare improving none recognize, with the exception of Innes (1999),[3] the possibility that implementing self-reporting can cause the level of compliance to fall.[4] Further, to date there has been no analysis of the exact conditions under which this will occur. Indeed, as Toffel and Short (2011) recently note in their review of the self-reporting literature:

> [a]lthough this scholarship identifies some important dynamics that underlie self-reporting...[its] connection to improving compliance or reducing harm is unclear.

By introducing market structure into this framework, we show that, in the context of market regulations, this outcome is determined by the level of competition and other market characteristics.

Besides the literature on self-reporting, this paper is contributes to the small, but recently growing literature on the relationship between market structure and various public and private enforcement mechanisms. Dechenaux and Samuel (2018) find that whether a regulator prefers announced or surprise inspections (from a compliance maximization standpoint) depends on whether or not the market is sufficiently concentrated. In the context of private enforcement mechanisms, Reinganum and Daugherty (2006) study the effectiveness of liability rules in various market contexts, and find whether strict liability is preferred to negligence also depends on the market's competitiveness. Our paper contributes to this literature by characterizing the welfare maximizing policy; this has not so far been addressed, perhaps due to its complexity.

The rest of this paper is organized as follows. Section 2 sets up the basic model as well as the market equilibrium. Section 3 studies the welfare maximization problem under

[3] In the conclusion, we elaborate on an important difference between our finding in this regard and Innes'.

[4] Kaplow and Shavell (1994) note that the optimal audit probability may be higher or lower under self-reporting than under no-reporting. But, these authors do not follow this result to its logical conclusion; namely, that this implies that compliance may fall under self-reporting. We are able to show that whether the audit probability is higher or lower under self-reporting is a function of market structure.

self-reporting and no-reporting for a constrained regulator who cannot choose the level of competition. Section 4 conducts the same analysis for an unconstrained regulator, and section 5 concludes. All proofs are provided in the appendix.

2 The model

Consider a market with $N \geq 1$ oligopolistic firms that each produce q_i units of a product. The total market quantity is $Q = \sum_{i=1}^{N} q_i$. The cost of producing each unit is c, and there are no fixed costs of producing q_i. Firms sell products to consumers with quasilinear utility function $U(\mathbf{q}, q_0; \mathbf{a}) = u(\mathbf{q}, \mathbf{a}) + q_0$ where good 0 is the numéraire, with $p_0 = 1$. We assume that U has the Bowley form:

$$U(\mathbf{q}, q_0; \mathbf{a}) = \sum_{i=1}^{N} \beta q_i - \frac{\gamma}{2}\left[\sum_i q_i^2 + \sum_i \sum_{i \neq j} q_i q_j\right] + q_0.$$

Maximizing this utility function with respect to a standard budget constraint yields the linear inverse demand,

$$P = \beta - \gamma Q.$$

Besides the direct costs, producing q_i units imposes a total negative cost (externality) on society $q_i h$. This externality can be abated at the rate $a_i \in (0, 1]$, so that the harm $q_i h$ only occurs with probability $(1 - a_i)$. Abatement, however, costs $k(a_i)$ per unit where we assume that $k(a_i) = k a_i^2/2$. Since abatement is costly, and the harm does not affect a firm's profits, a firm will not choose to abate unless there is some regulation. That is the laissez faire level of abatement is $a_{LF} = 0$.

To incentivize abatement a welfare maximizing regulator may choose to implement either a self-reporting regime or a no-reporting auditing regime, where $z \in \{NR, SR\}$ denotes the no-reporting and the self-reporting regime respectively. In the NR regime, the firm is audited with probability ρ_{NR} and, when harm has occurred (with probability $1 - a_i$), is fined $F_{NR} \in [0, F]$ per unit, where F is the maximal feasible fine. Thus, in the NR regime, a firm's profit is

$$\pi_{i,NR} = \left[\beta - \gamma Q - c - [1 - a_i]\, \rho_{NR} F_{NR} - \frac{k a_i^2}{2}\right] q_i. \tag{1}$$

In a self-reporting regime (SR) if harm occurs the firm self-reports the occurrence of harm with probability $\tau \in [0, 1]$, in which case it is fined $F_{SR} \in [0, F]$. In keeping with Kaplow and Shavell (1994) it is audited with probability ρ_{SR} when it does not make a report (or reports no harm) and is fined at the same rate F_{NR} that applies to unreported harm in the NR regime. Thus, a firm's profit in the SR regime is

$$\pi_{i,SR} = \left[\beta - \gamma Q - c - [1-\tau]\,[1-a_i]\,\rho_{SR} F_{NR} - \tau\,[1-a_i]\,F_{SR} - \frac{k a_i^2}{2}\right] q_i. \tag{2}$$

The timing of this game is as follows:

Stage 1. The regulator chooses $\{\rho_{NR}, F_{NR}\}$ in the no-reporting regime, and $\{\rho_{SR}, F_{SR}\}$ in the reporting regime;

Stage 2. The firm chooses a and q;

Stage 3. Harm is realized or not;

Stage 4. The firm self-reports in a SR regime if the harm occurred (with probability $1-a_i$);

Stage 5. The regulator audits with probability ρ_{NR} in the no-reporting regime, and in the self-reporting regime with probability ρ_{SR} when it does not receive a report.

Using backward induction (and subgame perfection), we first solve the model in the case of the SR regime. In stage 4 (taking quantities and abatement levels as given) the firm chooses τ to maximize profits. The derivative of (2) with respect to τ is

$$\rho_{SR}\,[1-a_i]\,F_{NR} - [1-a_i]\,F_{SR}.$$

Since $1-a_i \geq 0$, if $\rho_{SR}F_{NR} \geq F_{SR}$, then $\tau^* = 1$, otherwise, $\tau^* = 0$.

Although we have not yet introduced the regulator's welfare maximization problem, we find it convenient to note here that, as long as auditing costs are increasing in the audit probability, then the regulator sets $\rho_{SR}F_{NR} = F_{SR}$. Choosing $\rho_{SR}F_{NR} < F_{SR}$ cannot be optimal because then ρ can be lowered (maintaining the equality) while also improving welfare. Also, $\rho_{SR}F_{NR} > F_{SR}$ cannot be a solution since in that case firms would never self-report and the equilibrium would be identical to the NR regime. Thus, $\rho_{SR}F_{NR} = F_{SR}$ is optimal so that firms always self-report when harm occurs.[5] Thus, (2) reduces to

$$\pi_i = \left[\beta - \gamma Q - c - [1-a_i]\,F_{SR} - \frac{k a_i^2}{2}\right] q_i.$$

[5] Clearly, this result also follows from the revelation principle.

The first order condition with respect to a_i yields the profit maximizing level of abatement in the SR regime

$$a^* = \min\{\rho_{SR} F_{NR}/k, 1\} = \min\{F_{SR}/k, 1\}.$$

For now we assume that the solution to a^* is interior (i.e., $F_{SR}/k < 1$), but in assumption 1 (c) we ensure that this condition is always met.

Substituting the value for a^* into the profit function yields

$$\pi_i = \left[\beta - \gamma Q - c - \rho_{SR} F_{NR} + \frac{[\rho_{SR} F_{NR}]^2}{2k}\right] q_i.$$

Maximizing this expression with respect to the firms' quantities yields the symmetric Cournot-Nash equilibrium. This equilibrium is characterized in the following lemma.

Lemma 1 *Denote the firm's full marginal cost by*

$$m = c + \rho F - \frac{[\rho F_{NR}]^2}{2k}.$$

At a symmetric Nash equilibrium, for a given ρ, the firm's equilibrium quantity, profits, and abatement are,

$$q = \frac{\beta - m}{\gamma[1 + N]},$$

$$\pi = \gamma q^2,$$

and

$$a = \frac{\rho F_{NR}}{k}.$$

Note that the two regimes only affect the equilibrium quantity through m. Since fines in the SR regime are chosen such that $F_{SR} = \rho_{SR} F_{NR}$, the full marginal cost m is identical in both regimes, given ρ. Hence, the expressions for a, q, and π are identical in both regimes. Of course, since the optimal levels of ρ will not be the same in the two regimes, the quantities, profits, and abatement levels will also not be identical.

3 Welfare Analysis: constrained social planner

Given the market equilibrium in lemma 1 for some N, we study the regulator's welfare maximizing choice of fines and audit probability. That is, in this section we assume that the regulator is a "constrained social planner" who takes the market size N as given. Further, the regulator acts as a "Stackelberg leader" that chooses its policy anticipating the firms' reaction to its policy, identified in lemma 1. In other words, given the fines, the audit probability, and the regime, firms choose the symmetric Cournot oligopoly quantities and level of abatement derived in the previous section.

To identify the regulator's objective, we follow most of the literature in economics and assume that the regulator is a utilitarian (e.g. Mookherjee and Png, 1995) who maximizes the difference between the benefits and the costs to society. The expected cost of enforcement for the regulator is given by $C(.)$

$$C(\rho, \delta, z) = g\rho N q^{\delta} a^{\mathbf{1}_{z=SR}}; \qquad \delta \in \{0, 1\} \text{ and } g > 0, \tag{3}$$

where $\mathbf{1}_A$ takes the value one when condition A is true, and zero otherwise. When $\delta = 0$ costs are fixed in the sense that firm size does not affect enforcement costs, if $\delta = 1$ then costs are linear in firm size.[6] Note that, under self-reporting ($z = SR$), costs become a function of a, for the regulator need, in expectation, only audit the proportion a of firms who have not self-reported causing harm.

The benefit to society from this industry is given by

$$\Phi(\rho) = q_0 + \beta Q - \frac{\gamma}{2} Q^2 - Q\left[c + k(a) + [1 - a]\, h\right],$$

where $Q = qN$ is the equilibrium market size in the symmetric equilibrium characterized in lemma 1. In this benefit function we assume that fines are transfers from firms to society, therefore, the net cost to society of a fine is zero.

Given these costs and benefits, in the SR regime, the regulator chooses ρ and F to maximize

$$W_{SR} = \Phi(\rho) - C(\rho, \delta, SR),$$

while in the NR regime the regulator's welfare is

$$W_{NR} = \Phi(\rho) - C(\rho, \delta, NR),$$

In these two welfare functions, note that the cost differential between the two welfare functions is critical to the well-known result that self-reporting is optimal. Under self-reporting the regulator only needs to audit the firm whenever no accident has occurred (with probability a). Under a no-reporting regime, the regulator must always audit.

Before proceeding to analyze the socially optimal choices, we make the following assumptions for any regime $z \in \{NR, SR\}$.

[6] If $\delta \in (0, 1)$, costs are concave in firm size. We do not analyze this interior case in this paper.

Assumption 1 *The parameters in our model possess the following properties.*

a. *Demand is sufficiently strong; that is*

$$\beta - c > k,$$

so that full abatement is feasible (for the firm);

b. $h > k$*;*

c. $hF - kg < kF$*;*

d. *The fine* $F < w$ *where* w *is the wealth of a given firm.*

While we leave the algebra to the appendix, the intuitive justification for these assumptions is straightforward. Assumption 1(a) ensures firms produce a positive quantity even under full abatement. 1(b) ensures that the marginal benefit from abatement (a reduction in h) is greater than the marginal cost of abatement, ka, for all $a \in [0, 1]$. Hence, society wants to provide incentives for abatement (through regulation), instead of the alternative, complete deregulation. 1(c) ensures that full abatement is not optimal for the regulator.

Under these assumptions the regulator's welfare maximizing problem involves choosing ρ and F to maximize

$$W_z \qquad z \in \{NR, SR\}$$

subject to the constraint that $\rho \leq k/F$ (for, $a = 1 \Leftrightarrow \rho = k/F$ and it is never optimal to raise ρ once $a = 1$).

Our first step in identifying the welfare maximizing policy involves the following result concerning the optimal fine in the NR regime.

Lemma 2 *Regardless of the cost structure, the fine* F_{NR} *that applies to unreported harm in both the reporting and no-reporting regimes, is maximal.*

Given this result, herein the fine F_{NR} is the maximal fine F.

3.1 Fixed enforcement costs

Let ρ_z^* represent the welfare maximizing audit probability. When enforcement costs are fixed with respect to firm size ($\delta = 0$), then the socially optimal audit probability possesses the following characteristics with respect to the level of competition and the level of harm.

Proposition 1 *There exists a* $h_1(g) > h_2(g) > h_3(g) > 0$ *such that:*

a. *If* $h > h_1(g)$ *there exists an* $N_3 > N_2 > N_1 \geq 1$ *wherein*

   – *If* $N < N_1$, *then* $\rho^*_{NR} = \rho^*_{SR} = \min[k/F, 1]$;

   – *If* $N_1 < N < N_2$ *then* $\rho^*_{NR} > \rho^*_{SR} > 0$;

   – *If* $N_2 < N < N_3$, *then* $\rho^*_{SR} > \rho^*_{NR} > 0$;

   – *If* $N > N_3$, *then* $\rho^*_{SR} > \rho^*_{NR} = 0$.

b. *If* $h_2(g) < h < h_1(g)$ *there exists an* $N_3 > N_2 \geq 1$ *wherein*

   – *If* $N < N_2$ *then* $\rho^*_{NR} > \rho^*_{SR} > 0$;

   – *If* $N_2 < N < N_3$, *then* $\rho^*_{SR} > \rho^*_{NR} > 0$;

   – *If* $N > N_3$, *then* $\rho^*_{SR} > \rho^*_{NR} = 0$.

c. *If* $h_3(g) < h < h_2(g)$ *then, for all* N, $\rho^*_{SR} > \rho^*_{NR}$ *and there exists an* $N_1 \geq 1$ *such that for all* $N \geq N_2$, $\rho^*_{NR} = 0$.

d. *If* $h < h_3(g)$, *then for all* N, $\rho^*_{SR} > \rho^*_{NR} = 0$.

Proposition 1 is illustrated in Figure 1. Panel (a) depicts the optimal enforcement in (h, N)-space, as described in the proposition, and panel (b) shows ρ^*_{NR} and ρ^*_{SR} as functions of N for the case in which $h_2(g) < h < h_1(g)$. Observe that the optimal probabilities $\{\rho^*_{NR}, \rho^*_{SR}\}$ generally differ because, while the marginal social benefit from ρ is the same in either regime, the costs of enforcement differ at the margin. As first explained by Kaplow and Shavell (1994), on the one hand, the marginal enforcement cost tends to be lower with self-reporting because an increase in the probability of audit applies only to deterred firms. On the other hand, the marginal enforcement cost tends to be higher with self-reporting because an increase in the probability enlarges the pool of firms subject to audit by making harm less likely. The magnitude of the former effect is decreasing in ρ (for, as enforcement is tightened, an increasing proportion of firms generate no harm) while the magnitude of the latter effect is increasing in ρ. It follows that, under the conditions of the proposition, there exists a (unique) point at which marginal costs in the two regimes coincide.

Figure 1 – see p. 27

The proposition offers three key insights into optimal audit probabilities under the SR and NR regimes. First, under the SR regime regardless of the level of harm or the market's concentration, it is always optimal to provide incentives for abatement by auditing firms. In contrast, auditing is not always optimal in the NR regime (for some market structures). Thus, implementing a self-reporting regime permits welfare enhancing regulation in circumstances where the laissez faire outcome is preferred to a no-reporting regime (i.e., no regulation is optimal in a second-best sense because no-reporting is too costly).

Second, whether or not auditing is optimal depends on *both* the level of competition N and the level of harm h, because the total harm to society is proportional to Qh. However, this does not mean (as is the case in, e.g., Polinsky and Shavell (2000) and much of the remaining deterrence literature) that a higher total harm, Qh, implies a greater willingness to audit on the part of the regulator. Given some level of harm $h' > h_1$ the total harm under a monopoly, $Q_M h'$, is less than the total harm under a more competitive market, $Q_c h'$, without enforcement. Nevertheless, in the NR regime the regulator may choose to audit the monopolistic market (where total harm is lower) but not the more competitive market (where total harm is higher), if the latter case falls in the region where $N > N_3$ whereas the former case occurs in the region where $N < N_3$ in figure 1. Indeed, it is only when the harm is sufficiently large and the market sufficiently concentrated that the audit probability is positive under both regimes. The audit probability may even attain its maximum, $\rho = k/F$, if the market is sufficiently concentrated, a case that would essentially amount to continuous monitoring (Dechenaux and Samuel, 2018). Thus, to determine whether or not auditing is optimal, the regulator must account for both the level of competition and the per unit harm h; the total harm Qh is not sufficient.

Third, although implementing the SR regime always allows for harm-reducing regulation (regardless of the level of harm or market concentration), this does not imply that the abatement level under the NR regime is always lower than that produced under the SR regime. As seen in panel (b) of figure 1, if $N < N_2$, then implementing a SR regime can lower abatement (relative to the "status quo" NR regime), whereas the opposite is true if $N > N_2$. Consequently, when the level of competition is sufficiently high, then the level of abatement under self-reporting will be closer to full abatement, whereas when the level of competition is low the level of abatement under no-reporting more closely approximates full abatement.[7] Thus, competition is "good" for self-reporting in the sense that if markets are

[7] The second best level under costly enforcement is higher or lower under the SR or the NR regime because of the distortions introduced by the structure of the enforcement costs.

sufficiently competitive, the efficiency gains from self-reporting can be fully realized without raising crime. Indeed, if a regulator were constrained (perhaps politically) by the notion that any new policy implemented must lower the harm – a concern raised in the literature (Toffel and Short, 2011) – then it will likely be easier to advocate for self-reporting policies in more competitive markets.[8]

For the optimal audit probabilities identified in proposition 1, the following comparative static results hold.

Proposition 2 *The optimal audit probability*

- ρ^* *is*
  - *strictly decreasing in* N *and* γ *and increasing in* h *and* $\beta - c$*;*
  - *ambiguous in sign with respect to* k*.*
- $\partial \rho^*_{NR}/\partial F > 0 \Leftrightarrow \varepsilon_{\Phi_\rho,\rho} < -1$*, and* ρ^*_{SR} *is decreasing in* F*.*

The comparative statics with respect to h and $\beta - c$ are intuitive. As the harm increases the regulator needs to increase the audit intensity. Similarly, when demand is strong (i.e., $\beta - c$ large) then quantity produced increases, and consequently the harm also increases. Thus, audit intensity also rises. The effect of competition on the audit probability, however, is particularly interesting. Increases in competition, as measured by N, increase the marginal cost of raising the enforcement probability. Consequently, the optimal audit probability declines with N (Figure 1b).

An increase in the fine rate has competing effects. On the one hand it incentivizes firms to increase abatement. On the other hand, this increase in abatement induces firms to lower their output. The proof of Proposition 2 establishes that, in the NR regime, the balance of these competing effects depends upon whether the elasticity of the marginal social benefit, Φ_ρ, with respect to the probability of audit is elastic or inelastic. In the inelastic case, an increase in the fine rate increases the optimal audit probability. In the SR regime an increase in F has a third effect: it increases the marginal cost of raising the audit probability ($C_{\rho F} > 0$). This third effect is sufficient to ensure that, in the SR regime, the fine rate and

[8] Kaplow and Shavell (1994) show that in general the audit probability under self-reporting may be higher or lower than the probability under no-reporting. Since the fine is always maximal it implies that the level of crime may be higher or lower under self-reporting than in the no-reporting regime. This proposition "tightens" their result and shows that whether the audit probability in one regime is higher or lower than in the other depends on the level of competition (see figure 1).

the audit rate are substitutes in optimal enforcement. To our knowledge, this relationship between fine rates and optimal enforcement under self-reporting has not previously been explored in the literature.

Proposition 2 may also be used to understand the comparative static properties of N_2, the critical value of N at which $\rho^*_{NR} = \rho^*_{SR}$, and N_3, at which $\rho^*_{NR} = 0$. In particular, the comparative statics of N_3 are identical in sign to those of ρ^*, while, for N_2, only the comparative statics effects for k and F can possibly differ in sign from those of ρ^*. This implies that, as the harm increases, the range of competition ($N \geq N_3$) in which the laissez faire policy is optimal is smaller. In other words, even for relatively high levels of competition, $\rho_{NR} > 0$. Intuitively, because the harm is higher, the regulator chooses to audit under the NR regime even when the level of competition is relatively high.

Further, proposition 2 also claims that N_2 (if it exists) is also increasing in h. Recall that $\rho_{NR} > \rho_{SR}$ for $N < N_2$. Thus, as the harm increases the interval of N for which ρ_{NR} is higher than ρ_{SR} (and hence abatement higher in the NR regime) is larger. Accordingly, when h is large, a switch from the NR to the SR regime will lower the level of abatement for even moderately competitive industries ($N \in (N_1, N_2)$). Whereas when the harm is low ($h < h_2$) a switch to the SR regime increases abatement for all levels of market concentration.

Finally, both N_1 and N_2 are decreasing in γ. Recall that γ is the slope of the demand curve. Thus, when demand is more inelastic, the range of competition over which self-reporting yields a higher level of abatement grows.

3.2 Variable enforcement costs

We now consider the case where costs are variable (i.e., $\delta = 1$ in equation 3). Analogous to section 3.2, we characterize the optimal audit probability as a function of h and N in the following proposition and graphically depict the salient features of this proposition in figure 2. Panel (a) of the figure depicts optimal enforcement in (h, N)-space; and panel (b) shows $\tilde{\rho}_{NR}$ and $\tilde{\rho}_{SR}$ as functions of N for the case in which $\tilde{h}_2(g) < h < \tilde{h}_3(g)$.

Proposition 3 *There exists $\tilde{h}_3(g) > \tilde{h}_2(g) > \tilde{h}_1(g) > 0$, such that if*

- *$h \leq \tilde{h}_1(g)$, then $\tilde{\rho}_{SR} > \tilde{\rho}_{NR} = 0$ for all N.*
- *$\tilde{h}_1(g) < h < \tilde{h}_2(g)$, then there exists and $N_2 > N_1 \geq 1$ such that if*
    - *$N < N_1$, $\tilde{\rho}_{SR} > \tilde{\rho}_{NR} = 0$;*

- $N \in (N_1, N_2), \tilde{\rho}_{SR} > \tilde{\rho}_{NR} > 0$;
- $N > N_2, \tilde{\rho}_{NR} > \tilde{\rho}_{SR} > 0$.

- $\tilde{h}_2(g) < h < \tilde{h}_3(g)$, *then there exists an* N_2 *such that*
  - *if* $N < N_2, \tilde{\rho}_{SR} > \tilde{\rho}_{NR} > 0$;
  - *if* $N < N_2, \tilde{\rho}_{NR} > \tilde{\rho}_{SR} > 0$;
  - $h > \tilde{h}_3(g)$, *then for all* N $\tilde{\rho}_{NR} = k/F > \tilde{\rho}_{SR}$.

The main lesson from proposition 3 is that the results with respect to N are qualitatively the "inverse" of the case where costs are fixed. Specifically, as seen in Figure 2 panel (b), given some level of harm h, at higher levels of competition (i.e., $N > N_2$) the optimal level of enforcement is lower under the SR regime than under the NR regime. In contrast, when costs were assumed fixed, enforcement was stronger under the SR regime than under the NR for higher levels of competition. Accordingly, when costs are variable, a regime switch from the NR to the SR regime in a highly competitive industry will lower abatement when costs are variable, whereas, when costs are fixed, a switch from an NR to an SR regime will likely raise abatement in a highly competitive industry. Further, as can be seen in panel (b), when $h \in (\tilde{h}_2(g), \tilde{h}_3(g))$, for lower levels of competition enforcement is higher in the SR regime whereas for higher levels of competition enforcement is lower in the SR regime. The main message from our analysis is that the impact of self-reporting on compliance depends critically on both competition and the structure of the marginal cost of enforcement (i.e., whether it is fixed or variable).

Figure 2 – see p. 28

The following proposition further highlights the distinction between the cases $\delta = 0$ and $\delta = 1$.

Proposition 4 *If* $h \in (\tilde{h}_2, \tilde{h}_3)$, *the optimal audit probability* ρ_z *for* $z \in \{NR, SR\}$ *is strictly increasing in* N *and* h.

Proposition 4 proves a clear visual feature in panel (b) of Figure 2: optimal enforcement is increasing in the level of competition, N. The intuition underlying this finding is that, in the variable cost case, an increase in N has two effects on the marginal cost of raising the audit probability, C_ρ. First, a higher N increases C_ρ, as increasing *proportionally* the fraction firms that are audited implies a larger *absolute* number of extra audits, the larger is N. Second, however, higher competition endogenously reduces output per firm q, thereby reducing the per-firm audit cost. In contrast, in the fixed cost case only the first of these effects applies.

4 Welfare Analysis: Unconstrained social planner

We now assume that the social planner can choose N as well as ρ in both the NR and SR regimes. When moving from the NR regime to a SR regime the regulator faces a compromise. Simultaneously increasing N as well as ρ would potentially stimulate competition and reduce harm, but both acts would also raise the marginal cost of enforcement. If this latter effect were too large, therefore, social welfare might instead be maximized by increasing one of N or ρ, and potentially decreasing the other choice variable. Accordingly, the route to maximizing social welfare is not immediately obvious. Here we show that if the social planner can choose N then there will more competition but higher levels of harm in the (socially optimal) SR regime. This result is summarized in proposition 5 below.

Proposition 5 *Let $\hat{\rho}, \hat{N}$ denote the socially optimal level of auditing and market size. If enforcement costs are fixed ($\delta = 0$), this socially optimal policy for an unconstrained social planner possesses the following characteristics.*

- $\hat{\rho}_{NR} > \hat{\rho}_{SR}$;
- $\hat{N}_{NR} < \hat{N}_{SR}$.

Proposition 5 reveals an important finding previously ignored in the literature; that the socially optimal market size is higher when self-reporting policies can be implemented. Specifically, in the fixed cost case, a welfare maximizing regulator will, if switching from a NR regime to a SR regime, choose to lower the audit probability, as a consequence of which market competition increases, as does the level of harm. Since welfare is always raised under self-reporting, it follows that the socially optimal policy consists of implementing self-reporting. Given proposition (5) this, in turn, implies an increase in the market size (N) and, therefore, a reduction in prices and larger market surplus.

Some intuition for this finding comes from Figure 3, which depicts the social optimum in Proposition 5. The two lines $\rho^*_{NR}(N)$ and $\rho^*_{SR}(N)$ depict the optimal choice of audit probability for a given N in the NR and SR regimes. The two lines $N^*_{NR}(\rho)$ and $N^*_{SR}(\rho)$ depict the regulator's optimal choice of N for a given ρ (these functions have been inverted to be drawn in (ρ, N) space). The optimal pair $(\hat{N}_z, \hat{\rho}_z)$, $z \in \{NR, SR\}$ are each found at the intersection of $\rho^*_z(N)$ with $N^*_z(\rho)$. The optimal N^* in the NR regime is seen to be non-monotonic in ρ. At low values of ρ the industry generates large amounts of harm, inducing a regulator to restrict its size. At large values of ρ the marginal enforcement cost of raising N becomes increasingly high, again leading a regulator to wish to restrict N. The highest optimal choices of N therefore arise for intermediate values of ρ at which both harm and marginal enforcement costs are not too high.

Figure 3 – see p. 29

Switching to a SR regime alters the trade-off between harm and marginal enforcement costs. Note that, whereas self-reporting can be associated with either higher or lower marginal enforcement costs with respect to increases in ρ, self-reporting is always associated with lower marginal enforcement costs with respect to increases in N. Intuitively, following an increase of ΔN in N, an extra $\rho \Delta N$ firms must be audited under no-reporting, but only an extra $\rho \Delta N [1 - a]$ firms must be audited under self-reporting. Consistent with this point, in Figure 3 we see that $N^*_{SR}(\rho) > N^*_{NR}(\rho)$ for every value of ρ such that $a > 0$. The higher optimal N under self-reporting acts to reduce ρ, for – as we proved in Proposition 2 – the optimal audit probability is decreasing in N. As well as the optimal N being always higher under self-reporting it is also seen in the figure to vary to a greater degree in the choice of ρ. The interaction between ρ and N in the cost function is given by $C_{\rho N} = N^{-1} C_\rho > 0$. Hence, C_N is more sensitive to variation in ρ the higher is C_ρ. The greater variability in the optimal N^* under self-reporting therefore implies that C_ρ must be higher under self-reporting than under no-reporting. This, in turn, implies that $\rho^*_{SR}(N^*) < \rho^*_{NR}(N^*)$, which places $N^*_{SR}(\rho)$ and $N^*_{NR}(\rho)$ at values below N_2 in Figure 1. Accordingly, with reference to Figure 3, when switching from an NR regime to an SR regime, there are two effects on ρ, both of which are negative. The first is a discrete downwards jump when switching from the line $\rho^*_{NR}(N)$ to the line $\rho^*_{SR}(N)$ at $N = \hat{N}_{NR}$, and the second is a move rightward along the line $\rho^*_{SR}(N)$ from $\hat{N}_{NR}$ to $\hat{N}_{SR}$.

Similar intuitions apply to the variable cost case ($\delta = 1$), as depicted in Figure 4. It can be shown that the social optimum again lies in the region where $\rho^*_{NR}(N) > \rho^*_{SR}(N)$, albeit this occurs for $N^*_{NR} > N_2$ rather than $N^*_{NR} < N_2$. An important difference, however, is that the optimal audit probability is *increasing* in N. This implies that, in a switch from the NR regime to the SR regime, although the optimal ρ falls on account of the move downwards from $\rho^*_{NR}(N)$ to $\rho^*_{SR}(N)$, this effect is offset by an upward movement in ρ along the line $\rho^*_{SR}(N)$ arising from an increase in N. Accordingly, whether the optimal ρ increases or decreases from a switch from no-reporting to reporting remains unclear. Intractability precludes a more definite answer.[9]

Figure 4 – see p. 29

5 Conclusion

Although economic analyses of self-reporting show that implementing such a policy always raises welfare, there is still considerable dispute regarding its overall effectiveness (Toffel and Short, 2011). Many empirical studies find little evidence that implementing self-reporting has improved compliance rates (e.g., Esbenshade, 2004; Vidovic and Khanna, 2007). And, other studies find that compliance falls under self-reporting (King and Lenox, 2000). Consequently, some regulators have considered eliminating their self-reporting policies altogether (Toffel and Short, 2011).

In light of this debate, our paper makes an important contribution towards understanding these empirical findings and their implications for evaluating the impact of self-reporting. We show that the impact of self-reporting on compliance is affected by strategic market forces so that whether the optimal level of compliance is higher or lower under no-reporting than under self-reporting depends on the level of competition. Since many regulatory agencies regulate firms in oligopolistic contexts, our findings suggest that self-reporting need not raise compliance and lower the harm, even though such a policy will raise welfare. Accordingly, it may not be appropriate to evaluate the effectiveness of self-reporting by examining whether compliance rises or falls as a result of implementing self-reporting. Further, regulators introducing self-reporting need to consider the level of competitiveness in order to determine whether the harm will rise or fall.

[9] We note an an aside, however, that in Figure 4, and in other numerical examples we have tried, we observe the outcome $\hat{\rho}_{NR} > \hat{\rho}_{SR}$, consistent with Proposition 5.

The result that, under self-reporting, the optimal "permissible" level of harm may be higher than under no-reporting is related to findings in Innes (1999). For completely different reasons he finds that the level of "care" in preventing accidents is *always lower* under self-reporting than under no-reporting. Importantly, we find instead that the level of care (abatement) may be *higher or lower* under self-reporting than under no-reporting *depending on the level of competition.* This suggests that market characteristics should not be ignored when evaluating the benefits of enforcement policies such as self-reporting.

Our paper also identifies a new benefit to self-reporting. When both the audit probability and the market size, N, can be chosen by the regulator (the unconstrained case), then the optimal market size will be higher in a self-reporting regime than in a no-reporting regime. This reveals a new benefit in favor of implementing self-reporting regimes. Namely, that implementing self-reporting leads to a larger market size, and greater consumer surplus. To our knowledge this is a new finding in the literature. Indeed, while other papers have looked at some aspects of optimal enforcement in oligopolies (e.g., Baumann and Friehe, 2015) none of these papers study the characteristics of the optimal market size (N) in relation to enforcement. As we see, studying this problem reveals an important finding concerning the benefit of self-reporting.

We conclude by noting some extensions and ideas for future work. First, we did not consider the possibility of free entry and exit in this market. This could clearly be undertaken by assuming that there is fixed cost e that is incurred by firms upon entry. In this case, our results are broadly similar to the constrained regulator's choices in that if the harm is sufficiently large (small) then the optimal enforcement under the self-reporting regime is higher (lower) than the optimal enforcement in the no-reporting regime. Consequently, when the harm is large (small) fewer (more) firms enter the industry under the self-reporting regime. Second, while self-reporting generates positive surplus in a model with homogeneous firms, it may not do so if firms are sufficiently differentiated. Intuitively, in a vertically differentiated Bertrand duopoly, a firm's decisions to self-report will be a best-response to the other firm's decision to report. Hence, the impact on welfare is unclear. We leave it to future researchers to study these issues more closely.

Appendix

Assumption 1

Assumption 1 (a): Quantity (and hence profits) are positive, i.e., $q > 0$ when $a = 1$. Substituting $a = 1$ into the function for quantity yields $q = \frac{2[\beta - c - k/2]}{\gamma[1+N]} > 0$ or $2[\beta - c] - k > 0$.
Assumption 1 (b): As discussed in the main text.
Assumption 1 (c): Full abatement is not socially optimal for the regulator. To ensure this,

$$\left.\frac{\partial W_z}{\partial \rho_z}\right|_{\rho = k/F} = \left[\frac{N}{1+N}\right] \frac{[2[\beta - c] - k][Fh - k[F+g]]}{2\gamma k}.$$

At $a = 1$ the above expression must be negative, or

$$h < \frac{k}{F}[F+g],$$

which implies that $hF - kg < kF$.

Preamble to Proofs

The following expressions and their derivatives are utilized in the proofs of Propositions 1-4. $W = \Phi - C$ where

$$\begin{aligned}
\Phi &= Nq(\rho, N)\, w(\rho, N); \\
q &= q(\rho, N) = \frac{1}{\gamma[1+N]}\left\{\beta - c - \frac{a}{2}\left[2k - \rho\bar{F}\right]\right\}; \\
w &= w(\rho, N) = \gamma q\left[\frac{N+2}{2}\right] - [1-a]\left[h - \rho\bar{F}\right]; \\
C &= g\rho N\{1 - [1-a]\varphi\}; \\
a &= a(\rho) = \frac{\rho\bar{F}}{k}.
\end{aligned}$$

Recall that the case of no reporting corresponds to $\varphi = 0$, and the reporting case to $\varphi = 1$. Next, we establish the expressions for the following derivatives

$$q_\rho = -\frac{F[1-a]}{\gamma[1+N]} \leq 0; \qquad q_N = -\frac{q}{1+N} < 0; \tag{A.1}$$

$$w_\rho = \frac{[h - \rho F]F}{k} - \frac{\gamma N q_\rho}{2} > 0; \qquad w_N = \frac{\gamma q_N}{2} < 0; \tag{A.2}$$

$$q_{\rho N} = -\frac{q_\rho}{1+N} \geq 0; \qquad w_{\rho N} = -\frac{\gamma q_\rho}{2[1+N]} \geq 0. \tag{A.3}$$

Proof of Lemma 2

Note that ρ is always post-multiplied by F in Φ. Accordingly, social welfare can be written as

$$W = \Phi\left(\rho F\right) - C\left(\rho, \rho F\right). \tag{A.4}$$

Consider lowering ρ and increasing F holding ρF constant. Then $\Phi\left(\rho F\right)$ is unchanged, but $C\left(\rho, \rho F\right)$ falls (thereby increasing W), as

$$\left.\frac{\partial C\left(\rho, \rho F\right)}{\partial \rho}\right|_{\rho F = cons.} = \frac{C\left(\rho, \rho F\right)}{\rho} > 0.$$

This observation implies that W must be maximized with respect to F_{NR} at the maximal choice $F_{NR} = F$. ■

Proof of Proposition 1

We first characterize the optimal ρ^*. Using the expressions in the preamble, the first order condition for ρ is

$$\Phi_\rho - C_\rho = 0, \tag{A.5}$$

where

$$\Phi_\rho = N[q_\rho w + w_\rho q]; \qquad C_\rho = gN\left\{1 + [2a-1]\varphi\right\}.$$

Setting $\rho = k/F$ in the first order condition (A.5) we solve for N to obtain

$$N_1\left(\varphi\right) = \frac{F\left[h-k\right]\left\{2\left[\beta - c\right] - k\right\}}{2\gamma gk\left[1+\varphi\right]} - 1. \tag{A.6}$$

Next we prove the following claim:

Claim 1 $\frac{\partial[\Phi_\rho/N]}{\partial N} < 0$.

Proof. Using the derivatives in (A.1)-(A.3) we obtain

$$\frac{\partial\left[\Phi_\rho/N\right]}{\partial N} = \frac{\Phi_{\rho N}}{N} - \frac{\Phi_\rho}{N^2} = -\frac{\gamma q + 2w}{2\left[1+N\right]}q < 0.$$

■

Rewriting the first order condition in (A.5) as $N\left[\Phi_\rho/N - C_\rho/N\right] = 0$, an increase in N causes Φ_ρ/N to decrease (Claim 1), thereby forcing C_ρ/N to decrease also in order to restore the first order condition. As $C_\rho/N = g\left\{1 + [2a-1]\varphi\right\}$ is independent of N, for it to fall, it must be that ρ (and hence a) falls. It follows that $\rho = k/F$ for all $N \leq N_1\left(\varphi\right)$. As $N_1\left(\tau\right)$ is decreasing in φ, it follows that $\rho = k/F$ for all φ (and therefore in both the NR and SR regimes) when

$$N \leq N_1 = \max_\varphi N_1\left(\varphi\right) = \frac{F\left[h-k\right]\left\{2\left[\beta - c\right] - k\right\}}{2\gamma gk} - 1.$$

Finally to ensure that $N_1 > 1$, which holds if

$$h > k + \frac{4\gamma gk}{F\left\{2\left[\beta - c\right] - k\right\}} \equiv h_1\left(g\right).$$

We now turn to N_2. Setting $\rho = [1/2]\, k/F$ $(a = 1/2)$ and $N = 1$ in (A.5) we obtain

$$\frac{F\left\{8\left[\beta - c\right]\left[3k - 4h\right] - 4hk + 5k^2\right\}}{64\gamma k} - g.$$

From Claim 1 to have that $N_2 \geq 1$ it must therefore be that

$$h \geq \frac{64\gamma g + F\left\{24\left[\beta - c\right] - 5k\right\}}{8\left[\beta - c\right] - k}\frac{k}{4F} \equiv h_2\left(g\right). \tag{A.7}$$

Turning to N_3, setting $\rho = \varphi = 0$ in (A.5) implies that, at an optimum,

$$\Phi_\rho\left(0, N_3\right) = gN_3. \tag{A.8}$$

Under the condition

$$h \geq \frac{\left[\beta - c\right]F + 4\gamma g}{\beta - c + k}\frac{k}{2F} \equiv h_3\left(g\right)$$

equation (A.8) has a unique solution satisfying $N_3 > 1$. By Claim 1 it must hold that $\rho^*_{NR} = 0$ for all $N \geq N_3$. ■

Proof of Proposition 2: comparative statics of ρ

Let $\varepsilon_{a,b} \equiv [b/a]\,[\partial a/\partial b]$ be the elasticity of a with respect to b. We first prove a claim

Claim 2 *$\varepsilon_{\Phi_\rho,N} < 1$ for $N \geq 1$.*

Proof. Using (A.5), $\varepsilon_{\Phi_\rho,N} < 1$ when it holds that

$$\frac{1}{1+N} < \left[\frac{h - F\rho}{k - F\rho}\right]\left[\frac{1}{2} + \frac{k\left[1 - a\right]^2}{4\gamma q\left[\rho, 1\right]}\right].$$

Note that for $N \geq 1$ the left-side is does not exceed $1/2$, and the right-side necessarily exceeds $1/2$, hence the inequality holds as claimed. ■

Using the implicit function theorem in (A.5) we have that, for an arbitrary exogenous variable, z,

$$\frac{\partial \rho^*}{\partial z} = -\frac{W_{\rho z}}{W_{\rho\rho}}.$$

As $W_{\rho\rho} < 0$ is the second order condition for a maximum, the sign of $\partial\rho^*/\partial z$ is the sign of $W_{\rho z}$. Noting that $W_{\rho z} = \Phi_{\rho z} - C_{\rho z}$ we have

$$\begin{aligned}
W_{\rho N} &= \Phi_{\rho N} - \frac{C_\rho}{N} = \Phi_{\rho N} - \frac{\Phi_\rho}{N} = -\frac{\Phi_\rho}{N}\left[1 - \varepsilon_{\Phi_\rho, N}\right] < 0; \\
W_{\rho\gamma} &= \Phi_{\rho\gamma} = -\frac{\Phi_\rho}{\gamma} = -\frac{C_\rho}{\gamma} < 0; \\
W_{\rho h} &= \Phi_{\rho h} = N\left[\frac{qa}{\rho} - q_\rho\left[1-a\right]\right] > 0; \\
W_{\rho,\beta-c} &= \Phi_{\rho,\beta-c} = \frac{FN\left\{\left[h-k\right] + N\left[h-\rho F\right]\right\}}{\gamma k\left[1+N\right]^2} > 0; \\
W_{\rho F} &= \Phi_{\rho F} - C_{\rho F} = \frac{\Phi_\rho}{F}\left[1 + \varepsilon_{\Phi_\rho,\rho}\right] - C_{\rho F}; \\
&= \frac{C_\rho}{F}\left[1 + \varepsilon_{\Phi_\rho,\rho}\right] - C_{\rho F}; \\
&= \begin{cases} \frac{C_\rho}{F}\left[1 + \varepsilon_{\Phi_\rho,\rho}\right] & \text{if } \varphi = 0; \\ \frac{C_\rho}{F}\varepsilon_{\Phi_\rho,\rho} < 0 & \text{if } \varphi = 1; \end{cases} \\
W_{\rho k} &= \Phi_{\rho k} - C_{\rho k};
\end{aligned}$$

where the sign of $W_{\rho N}$ follows from Claim 2. It follows that $\partial\rho^*/\partial N < 0$, $\partial\rho^*/\partial\gamma < 0$, $\partial\rho^*/\partial h > 0$, $\partial\rho^*/\partial\left[\beta - c\right] > 0$, $\partial\rho^*/\partial F|_{\psi=0} \gtreqless 0 \Leftrightarrow \varepsilon_{\Phi_\rho,\rho} \gtreqless -1$, $\partial\rho^*/\partial F|_{\psi=1} < 0$, and $\partial\rho^*/\partial k \gtreqless 0$. ■

Proof of Proposition 3

When enforcement costs are variable,

$$C(\rho; z) = \begin{cases} gNq\rho \text{ if } z = NR; \\ gNq\rho a \text{ if } z = SR. \end{cases}$$

The first order condition for ρ is,

$$\frac{N}{\gamma\left[1+N\right]}\left[F\left[1-a\right]^2\left[h-\rho F\right] - \frac{F\left[1-a\right]\left[\beta-m\right]}{1+N} + \frac{F\left[\beta-m\right]\left[h-\rho F\right]}{k}\right] = \frac{\partial C(\rho;z)}{\partial\rho}, \tag{A.9}$$

Upon calculation it can be observed that both left- and right-side terms in condition A.9 possess the term $\frac{N}{\gamma[1+N]}$. Canceling this term write A.9 as,

$$\underbrace{\left[F\left[1-a\right]^2\left[h-\rho F\right] - \frac{F\left[1-a\right]\left[\beta-m\right]}{1+N} + \frac{F\left[\beta-m\right]\left[h-\rho F\right]}{k}\right]}_{MB} = MC_z, \tag{A.10}$$

where

$$MC_z = \begin{cases} g\left[\beta - c - 2\rho F + \frac{3}{2}\frac{\left[\rho F\right]^2}{k}\right] & \text{if } z = NR; \\ \frac{g}{k}\left[2\left[\beta-m\right]\rho F - \left[1-a\right]\rho^2 F\right] & \text{if } z = SR. \end{cases}$$

Hence, it is clear that MB is increasing in N, while the right-side is constant in N. Thus, given the assumption of concavity of welfare with respect to ρ it follows that ρ_{NR} and ρ_{SR} at an interior solution that satisfies A.10 are increasing in N.

Next, we establish the following:

1. at $\rho = 0$, $MC_{NR} = \frac{N}{\gamma[1+N]} g[\beta - c] > 0$, while $MC_{SR} = 0$;
2. at $\rho = k/F$, $MC_{NR} = g\left[\beta - c - \frac{k}{2}\right]$, while $MC_{SR} = g[2[\beta - c] - k]$, where we note that $g[2[\beta - c] - k] > g\left[\beta - c - \frac{k}{2}\right]$;
3. $\frac{\partial MC_{NR}}{\partial \rho} = -2gF < 0$ at $\rho = 0$ and $\frac{\partial MC_{NR}}{\partial \rho} = gF > 0$;
4. $\frac{\partial^2 MC_{NR}}{\partial \rho^2} = \frac{6F}{k} > 0$;
5. $\frac{\partial MC_{SR}}{\partial \rho} > 0$ at $\rho = 0, k/[2F]$, and at k/F;
6. MC_{SR} is convex for all $\rho \in [0, k/[2F])$ and concave for all $\rho \in (k/[2F], k/F]$. This result along with the result in (5.) implies that MC_{SR} is increasing in ρ for $\rho \in (0, k/[2F])$;
7. At $k/[2F]$ $MC_{SR} - MC_{NR} = \frac{gkN}{8\gamma(1+N)} > 0$; that is $MC_{SR} > MC_{NR}$ at $\rho = \frac{k}{2F}$.

Thus, since marginal costs in both regimes are continuous functions in ρ, and MC_{SR} is increasing in ρ for $\rho \in (0, k/[2F])$, and $MC_{SR} > MC_{NR}$ at $\rho = k/[2F]$, there exists a $\hat{\rho} \in (0, k/[2F])$ such that $MC_{NR} > MC_{SR}$ if and only if $\rho < \hat{\rho}$ (and $MC_{NR} = MCSR$ at $\hat{\rho}$). Using these observations, we now establish the claims in proposition 3.

First, we show that $\rho_{SR} > 0$ for all N. Since $MC_{SR} = 0$ at $\rho = 0$, as long as $MB|_{\rho=0,N=1} > 0$, $\rho_{SR} > 0$ for all N.

$$MB|_{\rho=0,N=1} = hF + \frac{Fh[\beta - c]}{k} - \frac{F[\beta - c]}{2},$$

which is strictly positive because the previous expression is increasing in h and positive at the smallest value of h, $h = \beta - c$. Thus, because the MB is increasing in N, $MB > MC$ at $\rho = 0$ therefore $\rho_{SR} > 0$.

Next, at $\rho = 0$ and $N \to \infty$, MB is,

$$hF + \frac{Fh[\beta - c]}{k}.$$

If this expression is less than $g[\beta - c]$, then $\rho_{SR} > \rho_{NR} = 0$ for all N. Simplifying this condition yields

$$h < \frac{g[\beta - c]k}{F[\beta - c] + Fk} \equiv \tilde{h}_1(g).$$

Therefore, if $h < \tilde{h}_1(g)$ then for all N, $\rho_{SR} > \rho_{NR} = 0$.

Next at $\rho = 0$ and $N = 1$, MB is

$$hF + \frac{Fh[\beta - c]}{k} - \frac{F[\beta - c]}{2},$$

which is less than $g\,[\beta - c]$ if and only if

$$h < \frac{[\beta - c]\,k}{[\beta - c + k]\,F}\left[g + \frac{F}{2}\right] = \tilde{h}_2(g) > \tilde{h}_1(g).$$

If $h \in [\tilde{h}_1(g), \tilde{h}_2(g)]$, then since MB is increasing in N and given the properties of MC_z concerning $\hat{\rho}$, there exists an $N_1 < N_2$ such that $\rho_{SR} > \rho_{NR} = 0$ for all $N < N_1$ and $\rho_{SR} > \rho_{NR} > 0$ if $N \in [N_1, N_2]$ and $\rho_{SR} < \rho_{NR}$ if $N > N_2$. Finally, if $h > \tilde{h}_2(g)$ then $\rho_{NR} > 0$, and there exists an N_2 such that $\rho_{NR} < \rho_{SR}$ if and only if $N < N_2$.

Finally, if MB at $\rho = k/F$ and $N = 1$ is greater than the MC_{NR} at $\rho = k/F$, then for all N $\rho_{NR} = k/F > \rho_{SR}$. At $\rho = k/F$, $MB = \{F\,[h - k]\,/k\}\,[\beta - c - k/2]$ and $MC = g\,[\beta - c - k/2]$. Therefore $MB > MC$ if

$$h \geq \frac{gk}{F} + k \equiv h_3(g) > h_2(g),$$

then $\rho_{NR} = k/F > \rho_{SR}$.

Proof of Proposition 5

Using the characterization of the regulator's objective function provided in Section 6.5, the first order conditions for $\{\rho, N\}$ can be written as

$$\Phi_\rho - C_\rho = 0; \qquad \Phi_N - C_N = 0;$$

where, when $\delta = 0$,

$$C_\rho = gN\,\{1 + [2a - 1]\,\varphi\}\,; \tag{A.11}$$
$$C_N = g\rho\,\{1 - [1 - a]\,\varphi\} > 0. \tag{A.12}$$

Using this framework, the result in Proposition 5 is obtained by proving each of the following claims. ■

Claim 3 *At any solution to the first order conditions for ρ and N, $\frac{\partial N}{\partial \rho} < 0$.*

Proof. The proof of this claim follows directly from the first order conditions. At any solution

$$\frac{\phi_\rho}{\phi_N} = \frac{C_\rho}{C_N}.$$

That is, at the optimal solution the marginal rate of substitution between ρ and N with respect to Φ must equal their rate of substitution with respect to costs. A straightforward calculation shows that keeping the total costs fixed at C'

$$\rho = \frac{C'}{gN\,\{1 - [1 - a]\,\varphi\}}.$$

Therefore, N and ρ are substitutes and at the optimum, $\frac{\partial N}{\partial \rho} < 0$. ■

Claim 4 *At the social optimum (in the SR regime),* $\hat{\rho}_{SR} > [1/2]\, k/F$.

Proof. As C is homogeneous of degree 1 in N, we have $C_N = C/N$, so $C = NC_N$. Hence $W = \Phi - C = \Phi - NC_N$. At $N = N^*$ we have $\Phi_N = C_N$, hence $W = \Phi - NC_N = \Phi - N\Phi_N$. By similar reasoning, at $\rho = \rho^*$, $W = \Phi - \rho\Phi_\rho \left[\frac{1-[1-a]\varphi}{1+[2a-1]\varphi}\right]$. It follows that,at a social optimum, $N\Phi_N = \rho\Phi_\rho \left[\frac{1-[1-a]\varphi}{1+[2a-1]\varphi}\right]$. Noting that $\frac{1-[1-a]\varphi}{1+[2a-1]\varphi} \leq 1$, it must hold that $\rho\Phi_\rho - N\Phi_N \geq 0$. Using the derivatives in (A.1)-(A.3). We obtain

$$\rho\Phi_\rho - N\Phi_N = \frac{1}{2}N^2 q \left[\frac{2w}{1+N} - \gamma q\right],$$

so it must hold, at a social optimum, that

$$w > \frac{\gamma q\,[1+N]}{2}, \tag{A.13}$$

which is equivalent to

$$w - \frac{\beta - m}{2} > 0.$$

Define $\zeta \equiv w - \frac{\beta-m}{2}$ then

$$\begin{aligned}
\frac{\partial\zeta}{\partial N} &= w_N < 0; \\
\frac{\partial\zeta}{\partial\beta} &= w_\beta - \frac{1}{2} = \frac{1}{2}\left[\frac{2+N}{1+N} - 1\right] > 0; \\
\frac{\partial\zeta}{\partial\rho} &= w_\rho + m_\rho > 0.
\end{aligned}$$

Thus, if (A.13) is not satisfied at the highest value of β (which is h) it is not satisfied for all β. Similarly, if (A.13) is not satisfied at the lowest value of N (which is 1), it is not satisfied for all N. Thus, we have that if

$$\frac{2h - a\,[2k - \rho F]}{8} - [1-a]\,[h - \rho F] \leq 0 \tag{A.14}$$

then ρ cannot be part of a social optimum. Moreover if (A.14) holds at some ρ' it holds for all $\rho \leq \rho'$ (so a social optimum must satisfy $\rho > \rho'$). Set $\rho = [1/2]\,k/F$ then (A.14) becomes

$$-\frac{8h - 5k}{32} < 0.$$

Hence $\rho = [1/2]\,k/F$ cannot be part of a social optimum. Rather, it must hold at a social optimum that $\hat{\rho} > [1/2]\,k/F$. ■

Claim 5 $\frac{\partial\hat{N}}{\partial\varphi} > 0$.

Proof. The first order condition for N can be written as $\Phi_N - g\rho\{1-[1-a]\tau\} = 0$. Then

$$\frac{\partial \hat{N}}{\partial \tau} = -\frac{g\rho[1-a]}{W_{NN}} = -\frac{g\rho[1-a]}{W_{NN}} > 0.$$

■

Claim 6 $\frac{\partial \hat{\rho}}{\partial \varphi} < 0$.

Proof. We have from Claim 4 that $2a-1>0$. So, from the first order condition for ρ in (A.5),

$$\frac{\partial \hat{\rho}}{\partial \varphi} = \frac{gN[2a-1]}{W_{\rho\rho}} < 0.$$

■

We now prove $\hat{N}_{SR} > \hat{N}_{NR}$. Using the chain rule, the total effect on $\hat{N}$ of an increase in φ is given by

$$\frac{d\hat{N}}{d\varphi} = \frac{\partial \hat{N}}{\partial \varphi} + \frac{\partial \rho}{\partial \varphi}\frac{\partial \hat{N}}{\partial \rho} > 0.$$

Hence $\hat{N}_{SR} > \hat{N}_{NR}$, where this inequality follows from Claims 3, 5, and 6. Again using the chain rule, the total effect on $\hat{\rho}$ of an increase in φ is given by

$$\frac{d\hat{\rho}}{d\varphi} = \frac{\partial \hat{\rho}}{\partial \varphi} + \frac{\partial N}{\partial \varphi}\frac{\partial \hat{\rho}}{\partial N} < 0.$$

Hence $\hat{\rho}_{NR} > \hat{\rho}_{SR}$, where this inequality follows from Claims 3, 5, and 6.■

References

Bayer, R., and F. Cowell. (2009). Tax compliance and firms' strategic interdependence. Journal of Public Economics, 93(11-12), 1131–1143.

Baumann, F., and T. Friehe. (2015). Competitive pressure and corporate crime. B.E. Journal of Economic Analysis and Policy, 16(2).

Daughety, A., and J. Reinganum. 2006. Markets, torts, and social efficiency. RAND Journal of Economics, 32(2), 300–323.

Dechenaux, E., and A. Samuel. 2018. Announced or Surprise Inspections and Oligopoly Competition. B.E. Journal of Economic Analysis and Policy. Forthcoming.

Esbenshade, J. 2004. Codes of conduct: Challenges and opportunities for workers' rights. Social Justice, 31(3), 40–59.

Innes, R., 1999. Remediation and self-reporting in optimal law enforcement. Journal of Public Economics 72(3), 379–393.

Innes, R., 2001. Violator avoidance activities and self reporting in law enforcement. Journal of Law, Economics, and Organization, 17(1), 239–256.

Kaplow, L., and S. Shavell. 1994. Optimal law enforcement with self-reporting of behavior. Journal of Political Economy, 102(3), 583–606.

King, A. A. and M. J. Lenox. 2000. Industry self-regulation without sanctions: The chemical industry's responsible care program. Academy of Management Journal, 43(4), 698–716.

Malik, A., 1993. Self-reporting and the design of policies for regulating stochastic pollution. Journal of Environmental Economics and Management, 24(3), 241–257.

Mookherjee D., and I.P.L. Png. 1995. Corruptible law-enforcers: how should they be compensated? The Economic Journal, 105 (428), 145-159.

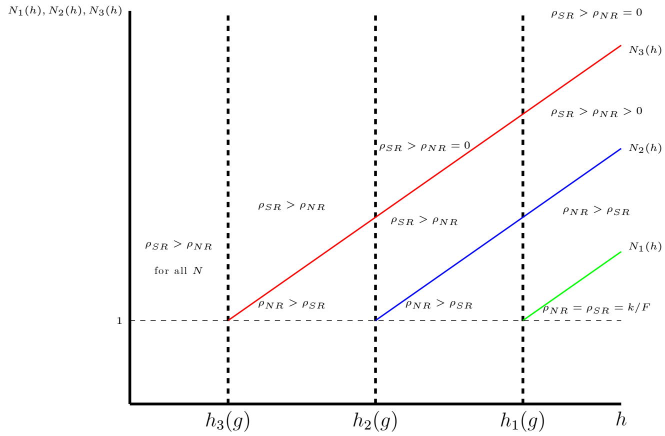
Polinksy, M. and S. Shavell. 2000. The economic theory of public enforcement of law. Journal of Economic Literature, 38(1), 45–76.

Toffel., M. W. and J. Short. 2011. Coming clean and cleaning up: Does voluntary self-reporting indicate effective self-policing? Journal of Law and Economics, 54(3), 609–649.

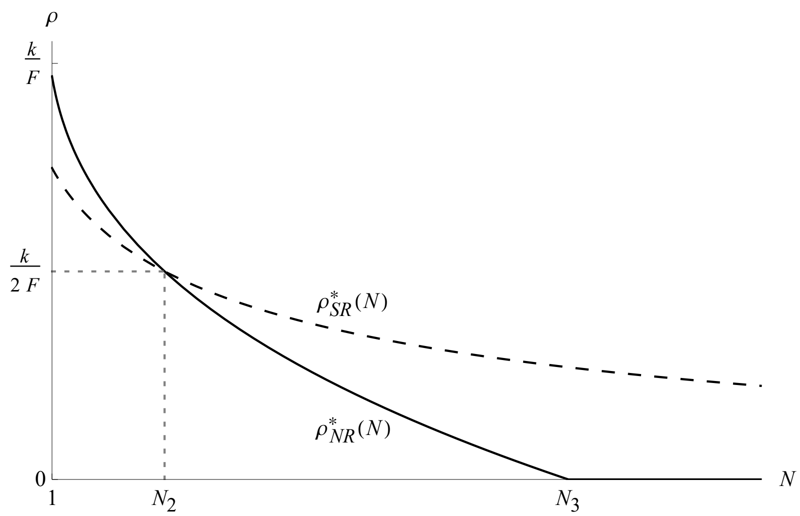
U.S. Census. 2007. United States Census: Business and industry [Data file]. Retrieved from `https://www.census.gov/econ/concentration.html`

Vidovic, M. and N. Khanna. 2007. Can voluntary pollution prevention programs fulfill their promises? Further evidence from the EPA's 33/50 program. Journal of Environmental Economics and Management, 53(2), 180–195.

Figures

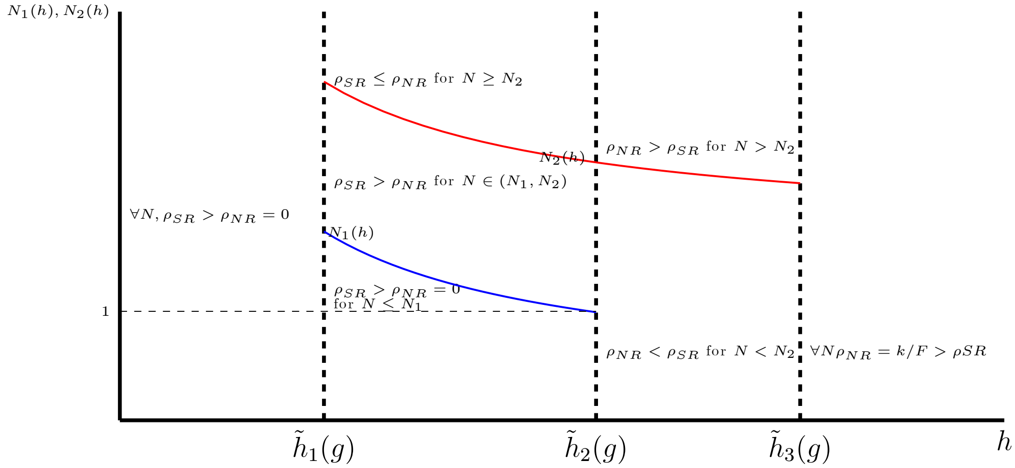


(a)

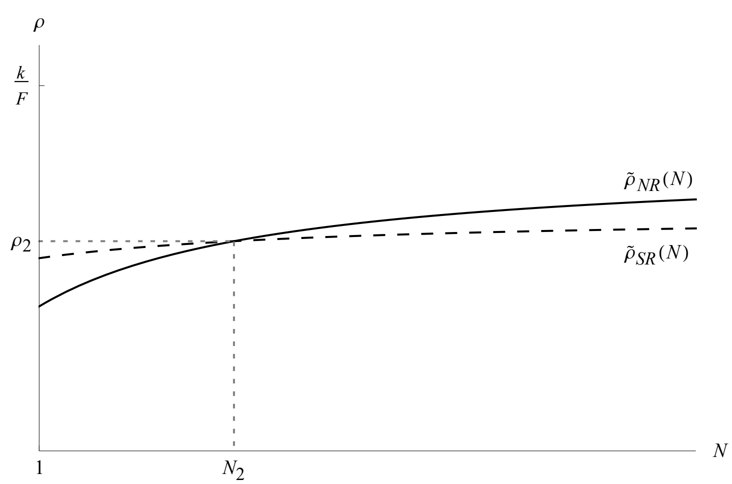


(b)

Figure 1: (a) Optimal enforcement in (h, N) space. (b) Optimal ρ



(a)



(b)

Figure 2: (a) Optimal enforcement in (h, N) space under variable costs. (b) Optimal enforcement probability under fixed enforcement costs.

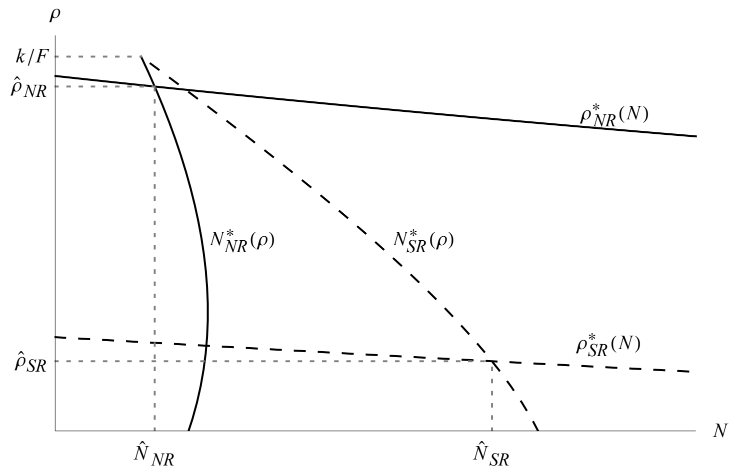


Figure 3: Socially optimal enforcement probability under fixed enforcement costs.

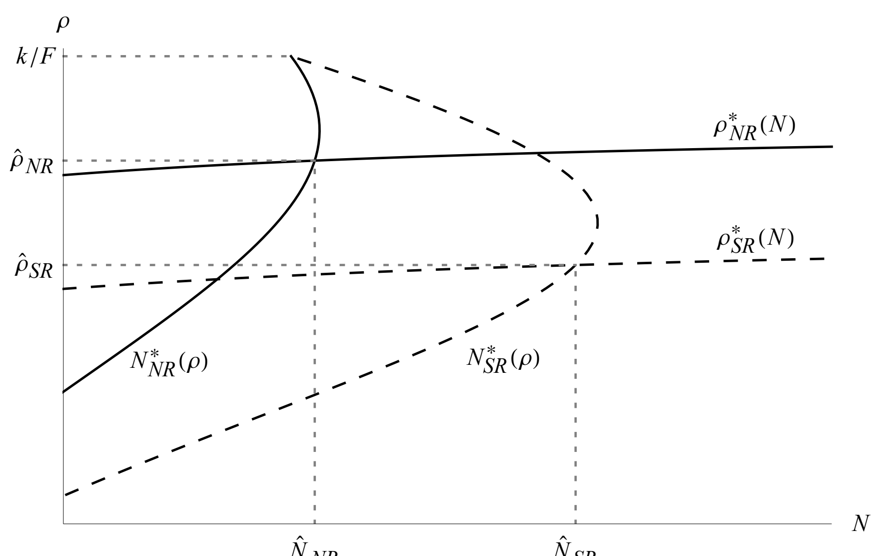


Figure 4: Socially optimal enforcement probability under variable enforcement costs.