Article:

Rablen, M. orcid.org/0000-0002-3521-096X and Samuel, A. (2021) Self-reporting and market structure. Economica, 88 (351). pp. 782-808. ISSN 0013-0427

https://doi.org/10.1111/ecca.12365

Economica

Economica (2021) **88**, 782–808
doi:10.1111/ecca.12365

Self-reporting and Market Structure

By MATTHEW D. RABLEN* and ANDREW SAMUEL†

*University of Sheffield †Loyola University Maryland

Final version received 23 December 2020.

Many regulators utilize self-reporting, that is, wrongdoers reporting their own crimes to the authority, 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 no-reporting) for a welfare-maximizing regulator. First, if the regulator can control only 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.

INTRODUCTION

Self-reporting is the reporting of harmful or non-compliant behaviour by the wrongdoer to the enforcement authority. Many regulators utilize self-reporting to enforce their regulations (Innes 2000). The US Environmental Protection Agency (EPA) and the UK Environment Agency, for example, encourage firms to self-report environmental 'crimes' such as spills of oil or of untreated sewerage. Similarly, the US 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 confess to the crime, thereby saving on auditing costs. Indeed, formal analysis of self-reporting (Malik 1993; Kaplow and Shavell 1994) has shown that it can implement a given level of compliance at a lower cost than enforcement without self-reporting ('no-reporting').[2] 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 the FDA regulate, respectively, the oligopolistic energy and pharmaceutical industries, and the USDA regulates an agricultural industry that is less than perfectly competitive.[3] 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 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 as follows. First, 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? Finally, if a planner is unconstrained and can choose both the level of enforcement and 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 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 be either 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.

Analysing this framework yields three important results concerning the value of regulating via self-reporting relative to no-reporting.

First, by utilizing self-reporting, a regulator can introduce welfare-enhancing regulations in markets where regulation would otherwise be inefficient to implement. To elaborate, if the regulations that are needed to correct a market failure (such as an external harm) are too costly, then it may be more efficient to permit the external harm rather than impose even costlier regulation. In these situations the laissez-faire policy of 'no regulation' can be optimal in a *second-best* sense, even though regulating the harm would be welfare-maximizing if regulation were costless (i.e. the *first-best* policy). Framed in our context, in the absence of self-reporting, there exists a threshold level of competition above which regulation becomes so costly that the regulator prefers the laissez-faire outcome over no-reporting. But if the regulator implements self-reporting, then for *any level of competition*, we show that the regulator prefers regulation (through self-reporting) to the laissez-faire policy that would be (second-best) optimal under no-reporting. Thus by utilizing self-reporting, it is always optimal to correct the market failure, whereas under no-reporting it may not always be optimal to do so.

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 the socially optimal audit probability and level of compliance are higher (lower) under self-reporting than under no-reporting when the market is sufficiently competitive (concentrated). If, however, enforcement costs vary with firm size, then this result is reversed: the optimal audit probability and level of compliance are higher (lower) under self-reporting than under no-reporting when the market is sufficiently competitive (concentrated). Thus whether or not implementing self-reporting yields a higher level of compliance, relative to regulation through no-reporting, depends on the level of competition. Importantly, the nature of this effect is 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 favour more competition and a 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. This result, importantly, implies that there should be more partnership and joint enforcement between competition (antitrust) authorities, which determine market concentration levels, and other regulators such as the EPA.

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 clean-up only 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, only Innes (1999)[4] recognizes the possibility that implementing self-reporting can cause the level of compliance to fall.[5] Further, to date there has been no analysis of the exact conditions under which this will occur. Indeed, as Toffel and Short (2011) note in their recent 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 regulation, this outcome is determined by the level of competition and other market characteristics.

Besides the literature on self-reporting, this paper contributes to the small but recently growing literature on the relationship between market structure and various public and private enforcement mechanisms. Dechenaux and Samuel (2019) 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, Daughety and Reinganum (2006) study the effectiveness of liability rules in various market contexts, and find that whether strict liability is preferred to negligence also depends on market competition. 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 I sets up the basic model as well as the market equilibrium. Section II studies the welfare-maximization problem under self-reporting and no-reporting for a constrained regulator that cannot choose the level of competition. Section III conducts the same analysis for an unconstrained regulator, and Section IV concludes. All proofs are provided in the Appendix.

I. 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 $U(\mathbf{q}, q_0) = q_0 + u(\mathbf{q})$, where good 0 is the numeraire, with $p_0 = 1$. We assume that U has the Bowley form

$$U(\mathbf{q}, q_0) = q_0 + \sum_{i=1}^{N} \beta q_i - \frac{\gamma}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} q_i q_j, \quad \gamma > 0.$$

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) $q_i h$ on society. This externality can be abated at the rate $a_i \in [0, 1]$, so that the harm $q_i h$

occurs only with probability $(1-a_i)$. Abatement, however, costs $k(a_i)$ per unit where we assume that $k(a_i)=ka_i^2/2$. 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 zero.

To incentivize abatement, a welfare-maximizing regulator may choose to implement either a self-reporting or a no-reporting regulatory regime, where $z \in \{NR, SR\}$ denotes the no-reporting and self-reporting regimes, respectively. In the NR regime, each firm is audited with probability ρ_{NR}, and when harm has occurred (with probability $1-a_i$) it 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

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

In a self-reporting regime (SR), if harm occurs, then the firm self-reports the occurrence of harm with probability $\tau_i \in [0,1]$, in which case it is fined $F_{SR} \in [0, F]$. In keeping with Kaplow and Shavell (1994), the firm 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.[6] Thus a firm's profit in the SR regime is

$$\pi_{i,SR} = \left(\beta - \gamma Q - c - (1-\tau_i)(1-a_i)\,\rho_{SR} F_{NR} - \tau_i(1-a_i) F_{SR} - \frac{ka_i^2}{2}\right) q_i. \tag{2}$$

The timing of this game is as follows.

1. Stage 1. The regulator chooses $\{\rho_{NR}, F_{NR}\}$ in the no-reporting regime, and $\{\rho_{SR}, F_{SR}\}$ in the self-reporting regime.
2. Stage 2. Firms choose a and q.
3. Stage 3. Harm is realized or not.
4. Stage 4. In the SR regime, if harm occurs (with probability $1-a_i$), then the firm chooses whether or not to self-report it.
5. Stage 5. The regulator audits with probability ρ_{NR} in the no-reporting regime, and with probability ρ_{SR} in the self-reporting regime when it does not receive a report.

Using backwards 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), firms choose τ to maximize profits. The derivative of equation (2) with respect to τ_i 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 $\tau_i^* = 1$; otherwise, $\tau_i^* = 0$.[7]

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, the regulator sets $\rho_{SR} F_{NR} = F_{SR}$. Choosing $\rho_{SR} F_{NR} > F_{SR}$ cannot be optimal because then ρ_{SR} can be lowered (up to the point of equality) while also improving welfare. Also, $\rho_{SR} F_{NR} < F_{SR}$ cannot be optimal as then 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 firms always self-report when harm occurs.[8] Thus equation (2) reduces to

$$\pi_{i,SR} = \left(\beta - \gamma Q - c - (1 - a_i) F_{SR} - \frac{k a_i^2}{2} \right) q_i.$$

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) below we ensure that this condition is always met.

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

$$\pi_{i,SR} = \left(\beta - \gamma Q - c - \rho_{SR} F_{NR} + \frac{(\rho_{SR} F_{NR})^2}{2k} \right) q_i.$$

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

Lemma 1. Denote a firm's full marginal cost in each regime by

$$m_z = \begin{pmatrix} c + \rho_{NR} F_{NR} - \dfrac{(\rho_{NR} F_{NR})^2}{2k} & \text{if } z = NR, \\ c + F_{SR} - \dfrac{(F_{SR})^2}{2k} & \text{if } z = SR. \end{pmatrix}$$

At a symmetric Nash equilibrium, a firm's quantity, profits and abatement are

$$q_z = \frac{\beta - m_z}{\gamma(1+N)}, \quad z \in \{NR, SR\},$$
$$\pi_z = \gamma q_z^2, \quad z \in \{NR, SR\},$$
$$a_z = \begin{pmatrix} \rho_{NR} F_{NR}/k & \text{if } z = NR, \\ F_{SR}/k & \text{if } z = SR. \end{pmatrix}$$

Note that the two regimes affect the equilibrium quantity only through m_z. Since fines in the SR regime are chosen such that $F_{SR} = \rho_{SR} F_{NR}$, the *algebraic* expression of the full marginal cost m is identical in both regimes, for given $\{\rho, F\}$. Accordingly, although Lemma 1 specifies the expression for q_z at the optimal policy under the SR and NR regimes, the expressions for a, q and are identical in both regimes, which is convenient analytically. As, however, the optimal levels of ρ will not be the same in the two regimes, the quantities, profits and abatement levels will not be identical.

II. 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' that takes the market size N as given. Further, the regulator acts as a 'Stackelberg leader' that chooses its policy anticipating 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) that maximizes the difference between the benefits and the costs to society. Then the expected cost of enforcement for the regulator is given by $C(.)$:

$$C(\rho,\delta,z)=gq^{\delta}\rho Na^{\mathbf{1}_{z=SR}},\quad \delta\in\{0,1\}\text{ and }g>0, \tag{3}$$

where $\mathbf{1}_A$ takes the value 1 when condition A is true, and 0 otherwise. Here $C(.)$ is the product of the cost per audit gq^{δ}, where $g>0$ is a scalar, and the expected number of audits $\rho Na^{\mathbf{1}_{z=SR}}$. The parameter δ determines the structure of costs—that is, whether they are fixed or variable in firm size (which is measured by q). When $\delta = 0$, costs are fixed in the sense that firm size does not affect enforcement costs; in this case g is exactly the marginal cost of audit. If $\delta = 1$, then costs are linear in firm size.[9] Importantly, equation (3) also helps to identify the benefit of self-reporting, first recognized in Kaplow and Shavell (1994). Under self-reporting ($z = SR$), costs become a function of a, for in expectation, the regulator need only audit the proportion a of firms who have not self-reported causing harm.

The benefit to society from this industry is given by

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

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; the net cost to society of a fine is therefore 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 maximizes

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

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 needs to audit only those firms that do not cause harm (with probability a). Under the no-reporting regime, the regulator must always audit.

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

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

1. Demand is sufficiently strong; that is, $\beta - c > k$, so full abatement is feasible (for firms).
2. $h > k$.
3. $hF - kg < kF$.
4. The fine F is less than the level of monopoly profit.

While we leave the algebra to the Appendix, the intuitive justification for these assumptions is straightforward. Assumption 1(a) ensures that firms produce a positive quantity even under full abatement. Assumption 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. Hence society wants to provide incentives for abatement (through regulation), instead of the alternative, complete deregulation. Assumption 1(c) ensures that full abatement is not optimal for the regulator. Note that if full abatement is optimal under self-reporting, then there is no longer any gain from self-reporting because the enforcement costs are identical in both regimes. This can be observed easily from equation (3). As we are interested in evaluating self-reporting, we do not explore the case where full abatement is optimal. Assumption 1(d) ensures that when the regulator imposes a fine, it is always feasible for the firm to pay it. This follows because profits are highest under a monopoly.

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

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

subject to the constraint $\rho \leq k/F$ (for $a=1$ if and only if $\rho = k/F$, 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}, which applies to unreported harm in both the self-reporting and no-reporting regimes, is maximal.

Given this result, herein the fine F_{NR} is the maximal fine F. A direct consequence of Lemmas 1 and 2 is that we can write the equilibrium level of abatement in each regime as

$$a_{NR} = \frac{\rho_{NR} F}{k}, \quad a_{SR} = \frac{\rho_{SR} F}{k} = F_{SR} k.$$

These expressions yield various observations that aid our subsequent characterization of the welfare-maximizing policies. First, it follows that when $\rho_{SR} > \rho_{NR}$, abatement will be higher in the SR regime. In this case we say that *enforcement is higher* in the SR regime compared to the NR regime. The reverse will be true when $\rho_{SR} < \rho_{NR}$. Second, when $\rho_z = 0$, abatement is zero. Therefore a policy that implements $\rho_z = 0$ is effectively the laissez-faire policy; if a policy implements $\rho_z > 0$, then some regulation or market intervention is welfare-maximizing. Accordingly, an increase in ρ_z can be described as an increase in enforcement.

Fixed enforcement costs

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

Proposition 1. The welfare-maximizing audit probability in the no-reporting regime, ρ_{NR}, may be higher or lower than the welfare-maximizing audit probability in the self-reporting regime, ρ_{SR}. Whether ρ_{NR} is greater or smaller than ρ_{SR} depends on the level of harm h, the cost of enforcement g, and the level of market competition N. Specifically, there exist thresholds h_1, h_2, h_3 on the level of harm that are functions of g, with $h_1(g) > h_2(g) > h_3(g) > 0$, such that the following hold.

1. If the harm is sufficiently high so that $h \geq h_1(g)$, then there exist N_1, N_2, N_3, with $N_3 > N_2 > N_1 \geq 1$, such that we have the following.
   1. If the market is sufficiently concentrated ($N \leq N_1$), then the audit probability is the same in both regimes: $\rho_{NR}^* = \rho_{SR}^* = \min\{k/F, 1\}$.
   2. If the market is moderately concentrated in the sense that $N_1 < N \leq N_2$, then enforcement is higher in the *NR* regime ($\rho_{NR}^* \geq \rho_{SR}^* > 0$).
   3. If the market is moderately competitive, in the sense that $N_2 < N < N_3$, then enforcement is higher in the *SR* regime ($\rho_{SR}^* > \rho_{NR}^* > 0$).
   4. If the market is sufficiently competitive ($N \geq N_3$), then the laissez-faire policy of $\rho_{NR}^* = 0$ is preferred under the *NR* regime, but regulation is still welfare-maximizing under the *SR* regime; that is, $\rho_{SR}^* > 0$.
2. If the harm is moderately high so that $h_2(g) \leq h < h_1(g)$, then there exist N_2, N_3, with $N_3 > N_2 \geq 1$, such that we have the following.
   1. If the market is sufficiently concentrated, in the sense that $N \leq N_2$, then enforcement is higher in the *NR* regime ($\rho_{NR}^* \geq \rho_{SR}^* > 0$).
   2. If the market is moderately concentrated in the sense that $N_2 < N < N_3$, then enforcement is higher in the *SR* regime ($\rho_{SR}^* > \rho_{NR}^* > 0$).
   3. If the market is sufficiently competitive in the sense that $N \geq N_3$, then the laissez-faire policy ($\rho_{NR}^* = 0$) is welfare-maximizing under the *NR* regime, but regulation is still welfare-maximizing under the *SR* regime ($\rho_{SR}^* > 0$).
3. If the harm is moderately low so that $h_3(g) \leq h < h_2(g)$, then for any level of competition (for all N), enforcement is higher in the *SR* regime and there exists an $N_3 \geq 1$ such that if $N \geq N_3$, then the laissez-faire policy is welfare-maximizing under the *NR* regime ($\rho_{NR}^* = 0$).
4. If the harm is sufficiently low so that $h \leq h_3(g)$, then for all levels of market concentration, the laissez-faire policy is welfare-maximizing under the *NR* regime ($\rho_{NR}^* = 0$), but regulation is always welfare-maximizing under the *SR* regime ($\rho_{SR}^* > 0$).

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.

Proposition 1 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, for example, 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 2019). 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 Figure 1(b), if $N < N_2$, then implementing an *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, 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.[10] Thus competition is 'good' for self-reporting in the sense that if markets are sufficiently competitive, then the efficiency gains from self-reporting can be realized fully without raising harm. Indeed, if a regulator were constrained (perhaps politically) by the notion that any new policy implemented must lower the harm—a concern raised in

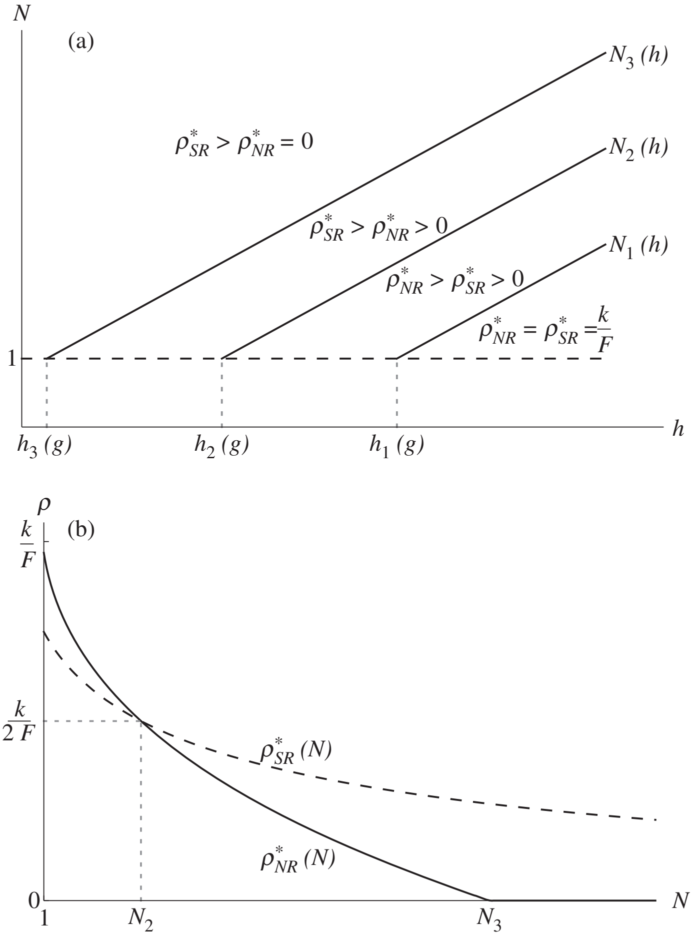

FIGURE 1. (a) Welfare-maximizing audit probability in (h,N)-space. (b) Welfare-maximizing audit probability and market size (competition).

the literature (Toffel and Short 2011)—then it will likely be easier to advocate for self-reporting policies in more competitive markets.[11]

For the welfare-maximizing audit probabilities identified in Proposition 1, the following comparative static result holds at an interior maximum.

Proposition 2. At an interior solution $\rho_z^* \in (0, k/F)$, $z \in \{NR, SR\}$, the welfare-maximizing audit probability possesses the following comparative static properties.

1. Under both the *NR* and *SR* regimes:
   1. ρ_z^* is strictly decreasing in market competition N and the slope of the demand curve γ;
   2. ρ_z^* is increasing in the level of harm h and the strength of the demand $\beta - c$;
   3. ρ_z^* may be increasing with respect to the cost of abatement k—that is, $\partial \rho_z^* / \partial k$ is ambiguous in sign.

2. The fine rate F can affect ρ_z^* differently in the NR and SR regimes:
   1. ρ_{NR}^* is increasing in F if the marginal benefit of an increase in ρ, Φ_ρ, is inelastic with respect to ρ, and is decreasing in F otherwise;
   2. 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), 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 audit probability. Consequently, the optimal audit probability declines with N (Figure 1(b)).

An increase in the fine rate has competing effects, which implies that whether the fine and audit rate are complements or substitutes depends on the reporting regime. On the one hand, an increase in F incentivizes firms to increase abatement. On the other hand, this increase in abatement induces firms to lower their output. The proof of Proposition 2establishes that in the NR regime, the balance of these competing effects depends on whether the marginal social benefit, Φ_ρ, is elastic or inelastic with respect to the probability of audit. In the inelastic case, an increase in the fine rate increases the optimal audit probability, so that the fine and the optimal audit probability are complements. 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 the audit rate are substitutes in optimal enforcement. To summarize, under the SR regime, F and the optimal audit probability are always substitutes; hence an increase in F allows the planner to lower ρ, thereby reducing enforcement costs. Such cost savings may not be enjoyed under no-reporting since ρ_{NR}^* and F may be complements. To our knowledge, this relationship between fine rates and optimal enforcement under self-reporting has not been explored previously 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 is 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.

Variable enforcement costs

We now consider the case where costs are variable (i.e. δ=1 in equation (3)). Analogous to the previous subsection, we characterize the optimal audit probability as a function of h and N in the following proposition.

Proposition 3. The welfare-maximizing audit probability in the no-reporting regime, ρ_{NR}, may be higher or lower than the welfare-maximizing audit probability in the self-reporting regime, ρ_{SR}. Whether ρ_{NR} is greater or smaller than ρ_{SR} depends on the level of harm h, the cost of enforcement g, and the level of market concentration N. Specifically, there exist thresholds $\tilde{h}_1(g), \tilde{h}_2(g), \tilde{h}_3(g)$ on the level of harm that are functions of g, with $\tilde{h}_3(g) > \tilde{h}_2(g) > \tilde{h}_1(g) > 0$, such that the following hold.

1. If the harm is sufficiently high so that $h > \tilde{h}_3(g)$, then for all levels of market concentration, the level of enforcement is higher in the *NR* regime ($\tilde{\rho}_{NR} = k/F > \tilde{\rho}_{SR}$).
2. If the harm is moderately high so that $\tilde{h}_2(g) < h \leq \tilde{h}_3(g)$, then there exists an $N_2 > 1$ such that we have the following.
    1. If the market is sufficiently concentrated in the sense that $N \leq N_2$, then enforcement is higher in the *SR* regime ($\tilde{\rho}_{SR} \geq \tilde{\rho}_{NR} > 0$).
    2. If the market is sufficiently competitive, in the sense that $N > N_2$, then the level of enforcement is higher in the *NR* regime ($\tilde{\rho}_{NR} > \tilde{\rho}_{SR} > 0$).
3. If the harm is moderately low so that $\tilde{h}_1(g) < h \leq \tilde{h}_2(g)$, then there exist N_1, N_2, with $N_2 > N_1 \geq 1$, such that we have the following.
    1. If the market is sufficiently concentrated, in the sense that $N \leq N_1$, then the laissez-faire policy $\tilde{\rho}_{NR} = 0$ is welfare-maximizing under the *NR* regime, but some regulation is still welfare-maximizing under the *SR* regime ($\tilde{\rho}_{SR} > \tilde{\rho}_{NR} = 0$).
    2. If the market is moderately competitive, in the sense that $N \in (N_1, N_2]$, then the level of enforcement is higher in the *SR* regime ($\tilde{\rho}_{SR} \geq \tilde{\rho}_{NR} > 0$).
    3. If the market is sufficiently competitive, in the sense that $N > N_2$, then the level of enforcement is higher in the *NR* regime ($\tilde{\rho}_{NR} > \tilde{\rho}_{SR} > 0$).
4. If the harm is sufficiently low so that $h \leq \tilde{h}_1(g)$, then for all market structures, the laissez-faire policy $\tilde{\rho}_{NR} = 0$ is welfare-maximizing under the *NR* regime but some regulation is still welfare-maximizing under the *SR* regime ($\tilde{\rho}_{SR} > \tilde{\rho}_{NR} = 0$).

We illustrate the salient features of this proposition graphically in Figure 2. Panel (a) 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 $\tilde{h}_2(g) < h < \tilde{h}_3(g)$.

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(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 higher under the *SR* regime than under the *NR* regime for higher levels of competition. Accordingly, when costs are variable, a regime switch from *NR* to *SR* in a highly competitive industry will lower abatement when costs are variable, whereas when costs are fixed, a switch from *NR* to *SR* will likely raise abatement in a highly competitive industry. Further, as can be seen in Figure 2(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*

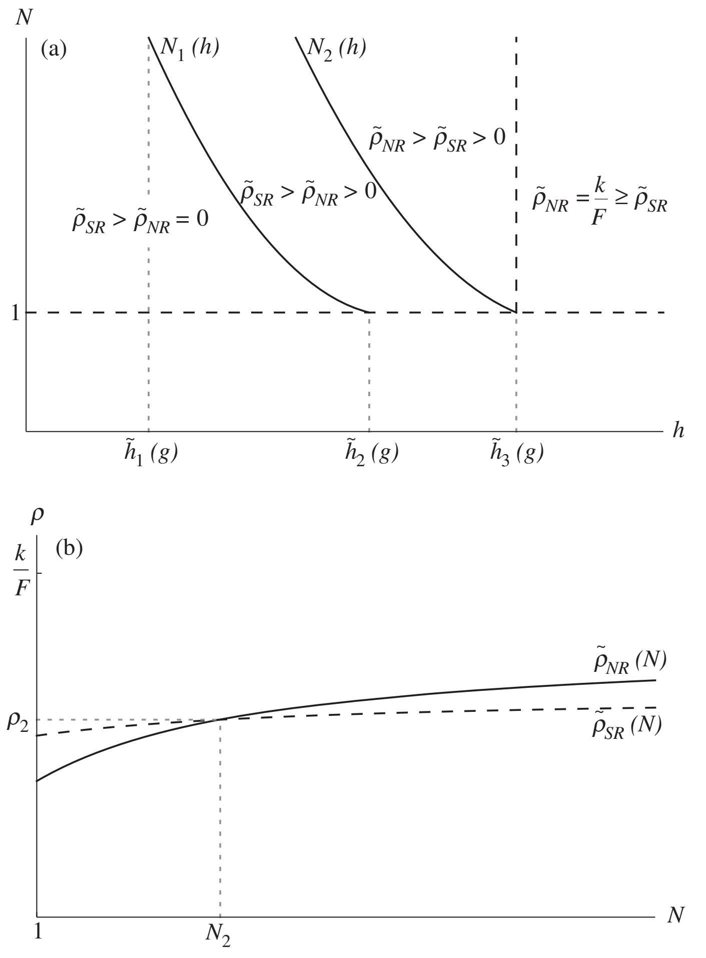

FIGURE 2. (a) Welfare-maximizing audit probability in (h,N)-space under variable costs. (b) Welfare-maximizing audit probability and market size, under variable enforcement costs.

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). This is especially the case for moderate levels of harm between $\tilde{h}_1(g)$ and $\tilde{h}_3$. When the harm is sufficiently high or low (cases (a) and (d) in Proposition 3), however, the audit probability does not depend on market structure.

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

Proposition 4. At an interior solution $\tilde{\rho}_z \in (0, k/F)$, $z \in \{NR, SR\}$, the welfare-maximizing audit probability under variable costs possesses the following comparative static properties.

1. Under both *NR* and *SR*:
   1. $\tilde{\rho}_z$ is strictly increasing in market competition N and the the level of harm h;
   2. $\tilde{\rho}_z$ is independent of the slope of the demand curve γ.

2. The fine rate F, the strength of the demand $\beta - c$, and the cost of abatement k, can each affect $\tilde{\rho}_z$ differently in the NR and SR regimes, but the relative magnitudes of these effects in the two regimes, and their signs, are ambiguous.

Proposition 4 proves a clear visual feature in Figure 2(b): 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 of 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.

A key difference between the fixed and variable cost cases is that the demand parameters $\beta - c$ and γ do not enter the cost function in the fixed cost case, but do in the variable cost case. As a result, whereas a steepening of the demand curve increases the optimal audit probability in the fixed case, it has no impact on the optimal audit probability in the variable case. Also, whereas an increase in the strength of demand increases the optimal audit probability in the fixed case, its impact in the variable case becomes ambiguous in sign. For parameters such as $\beta - c$ that interact with optimal enforcement in a complex way, it is possible, for prescribed parameter values, that the sign of the comparative statics effect differs between the NR and SR regimes. However, as the relative magnitudes of the effect between regimes is also complex, the divergence in signs (when present) can itself go in either direction, depending on parameter values.

Finally, note that when costs are variable, the optimal audit probability and the (maximal) fine may be complements or substitutes in either regime. Whereas when enforcement costs are fixed, the optimal audit probability and fine are necessarily substitutes in the SR regime (Proposition 2).

III. 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 an 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. Therefore if this latter effect were too large, then social welfare might instead be maximized by increasing one of N and ρ, and 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 the next proposition.

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

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

Proposition 5 reveals an important finding: 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 the NR regime to the 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 consumer 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 (N,ρ)-space). The optimal pair $(\hat{N}_z, \hat{\rho}_z)$, $z \in \{NR,SR\}$, are 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.

Switching to the 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 Figure 3 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.

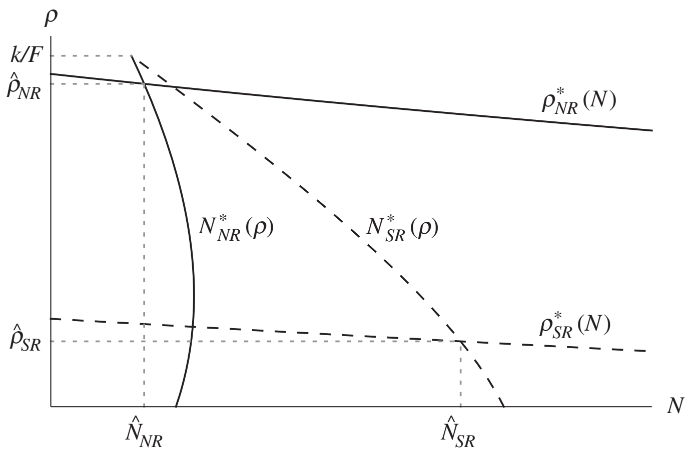

FIGURE 3. Socially optimal audit probability and market size under fixed enforcement costs.

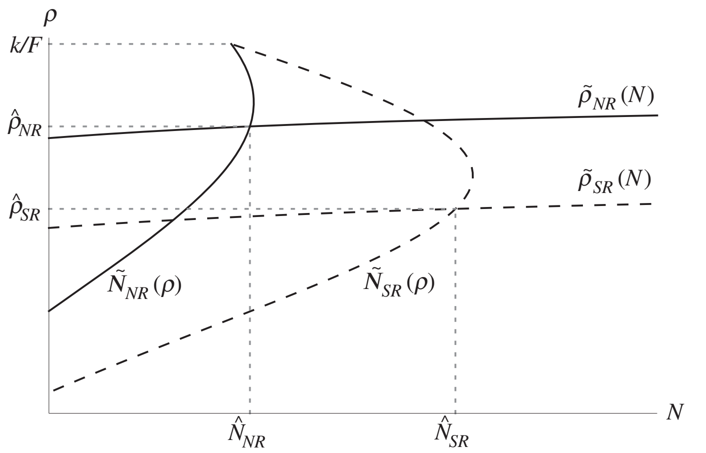

FIGURE 4. Socially optimal audit probability and market size under variable enforcement costs.

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 the *NR* regime to the *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 to the right along the line $\rho^*_{SR}(N)$ from $\hat{N}_{NR}$ to $\hat{N}_{SR}$.

Similar intuitions apply to the variable cost case (δ=1), as depicted in Figure 4. It can be shown that the social optimum again lies in the region where $\tilde{\rho}_{NR}(N) > \tilde{\rho}_{SR}(N)$, albeit this occurs for $\tilde{N}_{NR} > N_2$ rather than $\tilde{N}_{NR} < N_2$. 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 $\tilde{\rho}_{NR}(N)$ to $\tilde{\rho}_{SR}(N)$, this effect is offset by an upward movement in ρ along the line $\tilde{\rho}_{SR}(N)$ 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.[12]

Finally, note that the policy in Proposition 5 can be implemented by introducing a fixed entry cost y>0. Under this policy, firms enter the industry until their profits are zero, given the fixed cost y. As profits tend to zero with *N*, the regulator can choose $\rho = \hat{\rho}_{SR}$ and an entry fee y to achieve $\hat{N}_{SR}$.[13]

IV. 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 improves 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 a 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, though welfare-increasing, need not raise compliance and lower the harm. Accordingly, it may not be appropriate to evaluate the effectiveness of self-reporting by examining whether compliance rises or falls post-implementation.

Our results also imply that regulators introducing self-reporting need to consider the level of competitiveness in order to determine whether harm will rise or fall. This is especially important in the unconstrained social planner's problem where we show that a regulator chooses more competition while also 'permitting more harm' (Proposition 5). This suggests an important policy implication: that there may need to be more coordination between competition (antitrust) authorities and other regulatory bodies that correct for externalities.

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). Albeit for very different reasons, he finds that the level of 'care' in preventing accidents is *always lower* under self-reporting than under no-reporting. Notably, 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 new benefits that are achieved from implementing self-reporting. First, under self-reporting, the optimal audit rate and F are always substitutes (when costs are fixed). Thus an increase in F lowers the audit rate, thereby reducing enforcement costs. However, such cost savings will not always be realized in the NR regime, as there the optimal audit rate and the fine need not be substitutes. Second, when both the audit probability and the market size N can be chosen by the regulator (the unconstrained case), the optimal market size—and therefore also consumer surplus—will be higher in a self-reporting regime than in a no-reporting regime. Although previous literature has examined some aspects of optimal enforcement in oligopolies (e.g. Baumann and Friehe 2016), no study considers characteristics of the optimal market size in relation to enforcement. As we see, studying this problem reveals an important finding concerning the benefit of self-reporting.

We end 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 be undertaken by assuming that there is a fixed exogenous cost that is incurred by firms on entry. In this case, the number of firms that enter the industry depends on this cost, similar to the analysis in note 13 (except that here entry costs are exogenous, whereas there they are chosen by the regulator). Once N is determined, 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 high (low), fewer (more) firms enter the industry under the self-reporting regime. Second, while self-reporting generates a welfare 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. Finally, whereas we consider a standard 'static' Cournot setting, self-reporting could be analysed in a continuous time setting with, for example, stochastic revelation of demand. We leave it to future researchers to explore these avenues more closely.

APPENDIX

DISCUSSION OF ASSUMPTION 1

1. Quantity (and hence profits) are positive, that is, $q>0$ when $a=1$. Substituting $a=1$ into the function for quantity yields

$$q=\frac{2(\beta-c-k/2)}{\gamma(1+N)}>0 \quad \text{or} \quad 2(\beta-c)-k>0.$$

2. As discussed in the main text.
3. 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 $hF-kg<kF$.

4. Monopoly profit is given by $\max_q \pi_z=(\beta-\gamma q-m_z)q$. We assume that monopoly profit exceeds the fine rate F, so a firm can always afford to pay the fine F when levied.

PREAMBLE TO PROOFS

The following expressions and their derivatives are utilized in the proofs of Propositions 1–5. $W=\Phi-C$, where

$$\begin{aligned}
&\Phi(\rho,N)=\Phi=Nq(\rho,N)\,w(\rho,N),\\
&q(\rho,N)=q=\frac{\beta-m(\rho)}{\gamma(1+N)},\\
&w(\rho,N)=w=\gamma\left(\frac{N+2}{2}\right)q(\rho,N)-(1-a(\rho))(h-\rho F),\\
&C(\rho,N;\varphi)=C=gNq^{\delta}(\rho,N)\,\rho(1-(1-a(\rho))\varphi),\quad \varphi\in[0,1],\\
&a(\rho)=a=\frac{\rho F}{k}.
\end{aligned}$$

The case of no-reporting corresponds to $\varphi=0$, and the self-reporting case to $\varphi=1$. Next, we establish the expressions for the following derivatives:

(A1) $$q_\rho = -\frac{F(1-a)}{\gamma(1+N)} \leq 0, \quad q_N = -\frac{q}{1+N} < 0,$$

(A2) $$w_\rho = \frac{(h-\rho F)F}{k} - \frac{\gamma N q_\rho}{2} > 0, \quad w_N = \frac{\gamma q_N}{2} < 0,$$

(A3) $$q_{\rho N} = -\frac{q_\rho}{1+N} \geq 0, \quad w_{\rho N} = -\frac{\gamma q_\rho}{2(1+N)} \geq 0.$$

Proof of Lemma 1. Profit is given by $\pi_{i,z} = (\beta - \gamma Q - m_z)q_i$, $z \in \{NR, SR\}$. Differentiating with respect to q_i gives the first-order condition $\beta - \gamma Q - m_z - \gamma q_i = 0$. Imposing $q_i = q_z$ for all i (such that $Q = Nq_z$) and solving for q_z, the results in the lemma follow.

Proof of Lemma 2. Note that ρ is always post-multiplied by F in Φ, but not in C, where ρ appears both independently and post-multiplied by F. Accordingly, social welfare can be written as

(A4) $$W = \Phi(\rho F) - C(\rho, \rho F).$$

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

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

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

(A5) $$\Phi_\rho - C_\rho = 0,$$

where

$$\Phi_\rho = N(q_\rho w + w_\rho q), \quad C_\rho = gN(1 + (2a-1)\varphi).$$

Setting $\rho = k/F$ in the first-order condition (A5), we solve for N to obtain

(A6) $$N_1(\varphi) = \frac{F(h-k)(2(\beta-c)-k)}{2\gamma g k(1+\varphi)} - 1.$$

Next we prove the following claim.

Claim 1. $\partial(\Phi_\rho/N)/\partial N<0$.

Proof. Using the derivatives in (A1)–(A3), we obtain

$$\frac{\partial(\Phi_\rho/N)}{\partial N}=\frac{\Phi_{\rho N}}{N}-\frac{\Phi_\rho}{N^2}=-\frac{\gamma q+2w}{2(1+N)}q<0.$$

Rewriting the first-order condition in (A5) as $N(\Phi_\rho/N-C_\rho/N)=0$, 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 $C_\rho/N=g(1+(2a-1)\varphi)$ 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 $N\leq N_1(\varphi)$. As $N_1(\tau)$ is decreasing in φ, it follows that $\rho=k/F$ for all φ (and therefore in both the NR and SR regimes) when

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

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

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

We now turn to N_2. Setting $\rho=k/(2F)$ ($a=1/2$) and $N=1$ in equation (A5), we obtain

$$\frac{F\big(8(\beta-c)(3k-4h)-4hk+5k^2\big)}{64\gamma k}-g=0.$$

From Claim 1, for $N_2\geq 1$ we must therefore have

$$h\geq\frac{64\gamma g+F(24(\beta-c)-5k)}{8(\beta-c)-k}\frac{k}{4F}\equiv h_2(g). \tag{A7}$$

Turning to N_3, setting $\rho=\varphi=0$ in equation (A5) implies that at an optimum,

$$\Phi_\rho(0,N_3)=gN_3. \tag{A8}$$

Under the condition

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

equation (A8) 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 equation (A5), $\varepsilon_{\Phi_\rho,N} < 1$ when it holds that

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

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

Using the implicit function theorem in equation (A5), we have that for an arbitrary exogenous variable x,

$$\frac{\partial \rho_z^*}{\partial x} = -\frac{W_{\rho x}}{W_{\rho\rho}}, \quad z \in \{NR, SR\}.$$

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

$$\begin{aligned}
W_{\rho N} &= \Phi_{\rho N} - \frac{C_\rho}{N} = \Phi_{\rho N} - \frac{\Phi_\rho}{N} = -\frac{\Phi_\rho}{N}(1-\varepsilon_{\Phi_\rho,N}) < 0,\\
W_{\rho\gamma} &= \Phi_{\rho\gamma} = -\frac{\Phi_\rho}{\gamma} = -\frac{C_\rho}{\gamma} < 0,\\
W_{\rho h} &= \Phi_{\rho h} = N\left(\frac{qF}{k} - q_\rho(1-a)\right) > 0;\\
W_{\rho,\beta-c} &= \Phi_{\rho,\beta-c} = \frac{FN((h-k)+N(h-\rho F))}{\gamma k(1+N)^2} > 0,
\end{aligned}$$

$$\begin{aligned}
W_{\rho F} &= \Phi_{\rho F} - C_{\rho F} = \frac{\Phi_\rho}{F}(1+\varepsilon_{\Phi_\rho,\rho}) - C_{\rho F}\\
&= \frac{C_\rho}{F}(1+\varepsilon_{\Phi_\rho,\rho}) - C_{\rho F}\\
&= \begin{pmatrix} C_\rho(1+\varepsilon_{\Phi_\rho,\rho})/F & \text{if } \varphi = 0,\\ C_\rho \varepsilon_{\Phi_\rho,\rho}/F < 0 & \text{if } \varphi = 1, \end{pmatrix}
\end{aligned}$$

$$W_{\rho k} = \Phi_{\rho k} - C_{\rho k},$$

where the sign of $W_{\rho N}$ follows from Claim 2. It follows that $\partial \rho_z^*/\partial N < 0$, $\partial \rho_z^*/\partial \gamma < 0$, $\partial \rho_z^*/\partial h > 0$, $\partial \rho_z^*/\partial(\beta - c) > 0$, $\partial \rho_{NR}^*/\partial F \gtrless 0 \Leftrightarrow \varepsilon_{\Phi_\rho,\rho} \gtrless -1$, $\partial \rho_{SR}^*/\partial F < 0$, and $\partial \rho_z^*/\partial k \gtrless 0$.

Proof of Proposition 3. When enforcement costs are variable, we have

$$C(\rho, N;\varphi) = gNq(\rho, N)\,\rho(1-[1-a(\rho)]\varphi).$$

The first-order condition with respect to ρ is

$$\text{(A9)}\qquad \frac{N}{\gamma(1+N)}\big((\beta-m)w_\rho-(1-a)wF\big)=C_\rho.$$

On calculation it can be observed that both the left- and right-hand side terms in equation (A9) are proportional to $N/(\gamma(1+N))$. Cancelling this term, we write (A9) as

$$\text{(A10)}\qquad \underbrace{(\beta-m)w_\rho-(1-a)wF}_{MB}=MC,$$

where

$$MC = MC(\rho, N;\varphi) = g([1+(2a-1)\varphi](\beta-m)-(1-a)[1-(1-a)\varphi]\rho F).$$

Hence, as $MC_{NR}=MC(\rho, N;0)$ and $MC_{SR}=MC(\rho, N;1)$, we may write

$$MC_z=\begin{pmatrix} g(\beta-m-(1-a)\rho F) & \text{if } z=NR, \\ ag(2(\beta-m)-(1-a)\rho F) & \text{if } z=SR. \end{pmatrix}$$

Hence it is clear that marginal benefit is increasing in N, while marginal cost 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 equation (A10) are increasing in N.

Next, we establish the following points.

1. At ρ=0, $MC_{NR}=g(\beta-c)>0$, while $MC_{SR}=0$.
2. At ρ=k/F, $MC_{NR}=g\big(\beta-c-\frac{1}{2}k\big)$, while $MC_{SR}=g(2(\beta-c)-k)$, where we note that $g(2(\beta-c)-k)>g\big(\beta-c-\frac{1}{2}k\big)$.
3. $\partial MC_{NR}/\partial\rho=-2gF<0$ at ρ=0, and $\partial MC_{NR}/\partial\rho=gF>0$ at ρ=k/F.
4. $\partial^2 MC_{NR}/\partial(\rho^2)=3gF^2/k>0$.
5. $\partial MC_{SR}/\partial\rho>0$ for $\rho\in\{0,k/(2F),k/F\}$.
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 the previous point, implies that MC_{SR} is increasing in ρ for $\rho\in(0,k/(2F))$.
7. At ρ=$k/(2F)$, $MC_{SR}-MC_{NR}=\frac{1}{8}gk>0$; that is, $MC_{SR}>MC_{NR}$ at ρ=$k/(2F)$.

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 ρ=$k/(2F)$, there exists a $\hat{\rho}\in(0,k/(2F))$ such that $MC_{NR}>MC_{SR}$ if and only if $\rho<\hat{\rho}$ (and $MC_{NR}=MC_{SR}$ 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 ρ=0, as long as $MB|_{\rho=0,N=1}>0$, we have $\rho_{SR}>0$ for all N:

$$\text{(A11)}\qquad MB|_{\rho=0,N=1}=Fh+\frac{Fh(\beta-c)}{k}-\frac{F(\beta-c)}{2},$$

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

Next, at $\rho{=}0$ and $N{\to}\infty$, we have

$$MB=Fh+\frac{Fh(\beta-c)}{k}.$$

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 given by equation (A11), 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 as MB is increasing in N and given the properties of MC_z concerning ρ, there exist N_1,N_2, with $N_1<N_2$, such that $\rho_{SR}>\rho_{NR}=0$ for all $N<N_1$, $\rho_{SR}>\rho_{NR}>0$ if $N\in[N_1,N_2]$ and $\rho_{SR}<\rho_{NR}$ if $N>N_2$. 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 MC_{NR} at $\rho{=}k/F$, then for all N we have $\rho_{NR}=k/F>\rho_{SR}$. At $\rho{=}k/F$, $MB=\left(F(h-k)/k\right)\left(\beta-k-\frac{1}{2}k\right)$ and $MC=g\left(\beta-k-\frac{1}{2}k\right)$. 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 4. From the first-order condition (A10) we have that for an arbitrary exogenous variable x,

$$\frac{\partial\tilde{\rho}_z}{\partial x}=-\frac{MB_x-MC_x}{MB_\rho-MC_\rho},\quad z\in\{NR,SR\}.$$

As $MB_\rho-MC_\rho<0$ at a maximum, $\partial\tilde{\rho}_z/\partial x$ takes the sign of MB_x-MC_x. We then have

$$\begin{aligned}
&MB_N-MC_N=(\beta-m)w_{\rho N}-(1-a)w_NF>0,\\
&MB_h-MC_h=F\left((1-a)^2+\frac{\beta-m}{k}\right)>0,\\
&MB_\gamma-MC_\gamma=0,\\
&MB_{\beta-c}-MC_{\beta-c}=\frac{F((h-k)+N(h-\rho F))}{k(1+N)}-g(1-(1-2a)\varphi)\gtrless 0,\\
&MB_F-MC_F=\frac{MB}{F}(1+\varepsilon_{MB,\rho})-MC_F\gtrless 0,\\
&MB_k-MC_k\gtrless 0.
\end{aligned}$$

It follows that $\partial\tilde{\rho}_z/\partial N>0$, $\partial\tilde{\rho}_z/\partial h>0$, $\partial\tilde{\rho}_z/\partial\gamma=0$, $\partial\tilde{\rho}_z/\partial(\beta-c)\gtrless 0$, $\partial\tilde{\rho}/\partial F\gtrless 0$ and $\partial\tilde{\rho}/\partial k\gtrless 0$.

Proof of Proposition 5. Using the characterization of the regulator's objective function provided in the subsection 'Preamble to proofs' above, the first-order conditions for $\{\rho,N\}$ can be written as

$$\Phi_\rho - C_\rho = 0, \quad \Phi_N - C_N = 0,$$

where, when δ=0,

$$C_\rho = gN(1+(2a-1)\varphi), \tag{A12}$$

$$C_N = g\rho(1-(1-a)\varphi) > 0. \tag{A13}$$

Using this framework, the result in Proposition 5 is obtained by proving each of Claims 3–6 below.

Claim 3. At any solution of the first-order conditions for ρ and N, $\partial N/\partial\rho<0$.

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' gives

$$\rho = \frac{C'}{gN(1-(1-a)\varphi)}.$$

Therefore N and ρ are substitutes, and at the optimum, $\partial N/\partial\rho<0$.

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

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

$$\frac{1-(1-a)\varphi}{1+(2a-1)\varphi} \le 1,$$

it must hold that $\rho\Phi_\rho - N\Phi_N \ge 0$. Using the derivatives in (A1)–(A3), 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

$$w > \frac{\gamma q(1+N)}{2}, \tag{A14}$$

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 inequality (A14) is not satisfied at the highest value of β (which is h), then it is not satisfied for all β. Similarly, if inequality (A14) is not satisfied at the lowest value of N (which is 1), then it is not satisfied for all N. Thus we have that if

$$\frac{2h - a(2k - \rho F)}{8} - (1-a)(h - \rho F) \le 0, \tag{A15}$$

then ρ cannot be part of a social optimum. Moreover, if inequality (A15) holds at some ρ', then it holds for all $\rho \le \rho'$ (so a social optimum must satisfy $\rho > \rho'$). Set $\rho = k/(2F)$. Then inequality (A15) becomes

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

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

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

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. $\partial\hat{\rho}/\partial\varphi < 0$.

Proof. We have from Claim 4 that $2a-1>0$. So from the first-order condition for ρ in equation (A5),

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

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

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

Hence $\hat{N}_{SR} > \hat{N}_{NR}$, 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

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

Hence $\hat{\rho}_{NR} > \hat{\rho}_{SR}$, from Claims 3, 5 and 6.

NOTES

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 (all accessed 11 January 2021). See also Toffel and Short (2011) for a discussion of various self-reporting policies.
2. The term ‘no-reporting’ refers to a standard ‘Beckerian’ enforcement framework in which all firms are audited and sanctioned with a given probability.
3. For example, according to US Census data for 2007 (US Census 2007), the (C_4) concentration ratio for offshore drilling (which is regulated by the EPA) is 50. Similarly, the FDA regulates animal antimicrobials ($C_4 > 50$) and medical devices ($C_4 = 35$).
4. In Section IV, we elaborate on an important difference between our finding in this regard and that of Innes.
5. 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 observation to its logical conclusion, namely, that it 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.
6. *A priori*, there is no reason why the fine for unreported harm in the *SR* regime must be the same as the fine for unreported harm in the *NR* regime. Nevertheless, as we show in Lemma 2, the optimal fine in both these

cases is maximal. Thus although we have not yet derived this result, we make this assumption to avoid excessive notation.

7. Strictly, firms are indifferent between self-reporting or not when $\rho_{SR}F_{NR} = F_{SR}$. In keeping with the mechanism design literature, however, we assume that when they are indifferent, firms report truthfully.
8. This result also follows from the revelation principle.
9. If $\delta \in (0,1)$, then costs are concave in firm size. We do not analyse this interior case here.
10. The second -best level under costly enforcement is higher or lower under the SR or the NR regime because of the differential structure of enforcement costs between the two regimes.
11. 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. As the fine is always maximal, it implies that the level of harm may be higher or lower under self-reporting than in the no-reporting regime. Proposition 1 '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).
12. We note as an aside, however, that in Figure 4—and in other numerical examples that we have tried—we observe the outcome $\hat{\rho}_{NR} > \hat{\rho}_{SR}$, consistent with Proposition 5.
13. A straightforward calculation shows that $\hat{N}$ in Proposition 5 is implemented by an entry cost $y(\hat{N}) = \gamma^{-1}(\beta - m_z)^2(1+\hat{N})^{-2}$, $z \in \{NR, SR\}$.

REFERENCES

BAUMANN, F. and FRIEHE, T. (2016). Competitive pressure and corporate crime. *B. E. Journal of Economic Analysis and Policy*, **16**(2), 647–87.

DAUGHETY, A. F. and REINGANUM, J. F. (2006). Markets, torts, and social inefficiency. *RAND Journal of Economics*, **37**(2), 300–23.

DECHENAUX, E. and SAMUEL, A. (2019). Announced or surprise inspections and oligopoly competition. *B. E. Journal of Economic Analysis and Policy*, **19**(1), 1–20.

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), 379–93.

INNES, R. (2000). Self-reporting in optimal law enforcement when violators have heterogeneous probabilities of apprehension. *Journal of Legal Studies*, **29**(1), 287–300.

INNES, R. (2001). Violator avoidance activities and self reporting in law enforcement. *Journal of Law, Economics, and Organization*, **17**(1), 239–56.

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

KING, A. A. and LENOX, M. J. (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), 241–57.

MOOKHERJEE, D. and PNG, I. P. L. (1995). Corruptible law-enforcers: how should they be compensated? *Economic Journal*, **105**(428), 145–59.

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

TOFFEL, M. W. and SHORT, J. (2011). Coming clean and cleaning up: does voluntary self-reporting indicate effective self-policing? *Journal of Law and Economics,* **54**(3), 609–49.

US CENSUS (2007). United States Census: Business and industry [Data file]; available online at https://www.census.gov/econ/concentration.html (accessed 14 January 2021).

VIDOVIC, M. and KHANNA, N. (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), 180–95.