WORKING PAPERS

N° 1316

October 2024

"Unilateral Practices, Antitrust Enforcement and Commitments"

Michel Polo and Patrick Rey

Toulouse School of Economics

Unilateral Practices, Antitrust Enforcement and Commitments*

Michele Polo[†] Patrick Rey[‡]

September 4th, 2024

Abstract

This paper analyses the impact on antitrust enforcement of commitments, a tool introduced in Europe by the Modernization reform of 2003, and intensively used since then by the European Commission and by National Competition Agencies. We consider a setting where a firm can adopt a costly practice that is either pro- or anti-competitive; the firm knows the nature of the practice and its cost whereas the enforcer has only prior beliefs about them. If the firm adopts the practice, the enforcer then decides whether to open a case. We compare a benchmark regime in which the enforcer can only run a costly investigation that may or may not bring evidence, with policy regimes in which commitments are available. We first analyze a regime reflecting the 2003 regulation, in which the firm can offer a commitment whenever a case is opened. We find that, in most cases, the introduction of commitments does not improve enforcement performance. We then study a potential reform of the regulation giving the enforcer the initiative to propose commitments. We show that this regime dominates the benchmark and current regulations whenever enforcement is desirable.

Keywords: Antitrust enforcement, commitment, remedies, deterrence.

JEL Codes: L40, K21, K42.

*We thank Pierpaolo Battigalli, Massimo Motta, Volker Nocke, Matias Pietola, Marteen Pieter Schinkel and participants at the MaCCI Annual Conference (2016, Mannheim), the EARIE Conference (2016, Lisbon), the CRESSE Conference (2022, Crete), the EEA Conference (2022, Milan), the seminar of the Finnish Competition Authority (2023, Helsinki), the 5th UNIBG IO Symposium (2023, Passo de Tonale) and the MaCCI Summer Institute (2024, Annweiler am Trifels) for useful comments. We gratefully acknowledge financial support from the ANR Investissements d'Avenir program under the grant CHESS ANR-17-EURE-0010 and from the Horizon Europe - European Research Council (ERC) under Grant Agreement No. 101055280-DIGECON

[†]Bocconi University, IGIER and GREEN.

[‡]Toulouse School of Economics, University of Toulouse Capitole, France.

1 Introduction

The modernization reform of 2003 introduced new tools for enforcing European competition law. Notably, companies under investigation can propose remedies to address potential competitive concerns. If the European Commission accepts these remedies, they become binding commitments. Consequently, the case is closed without a fine or a finding of infringement.[1] Since this reform, the European Commission – and national competition authorities, under similar rules – have increasingly relied on commitments,[2] raising a debate on the opportunity of such a widespread adoption. Critics fear that this tool may weaken deterrence and encourage companies to undertake dubious practices, by enabling them, in case of an investigation, to avoid fines and reputational damages – and minimize as well the scope for private damages from follow-on cases. Advocates claim instead that this instrument may speed-up antitrust enforcement and improve legal certainty, compared with the long, costly and uncertain process of running a full-scale investigation exposed to judicial review.

Although commitments are nowadays a very common tool, there is surprisingly little theoretical analysis assessing their impact on antitrust enforcement. In this paper, we consider a setting in which a firm must decide whether to undertake a costly action. The action generates a known profit but can be either pro- or anti-competitive. The firm knows the cost and nature of the action, whereas the enforcer only has prior beliefs about them. If the firm undertakes the action, the enforcer must decide whether to open a case.

In this setting, we identify three channels through which antitrust policy can affect welfare. From an ex-ante perspective, it can discourage the adoption of the practice, either uniformly, regardless of its nature (which we will refer to as *deterrence*), or selectively, by targeting bad actions (*screening*). In addition, it can also improve antitrust enforcement, once the practice has been adopted (*ex-post effectiveness*).

We first consider a benchmark regime in which the only instrument in the enforcer's toolbox is an in-depth investigation. Running an investigation is costly (for both the enforcer and the firm) but, when the action is anti-competitive, the enforcer obtains evidence of infringement with positive probability, in which case it stops the practice and imposes a fine. We show that, in equilibrium, the enforcer runs an investigation with certainty when it has very pessimistic priors, and with positive but decreasing probability when its beliefs become less pessimistic. Thus, enforcement is costly (the enforcer and

[1] Art. 9.1 of Regulation 1/2003 states: "Where the Commission intends to adopt a decision requiring that an infringement be brought to an end and the undertakings concerned offer commitments to meet the concerns expressed to them by the Commission in its preliminary assessment, the Commission may by decision make those commitments binding on the undertakings. Such a decision may be adopted for a specified period and shall conclude that there are no longer grounds for action by the Commission."

[2] According to Mariniello (2014), in the period 2004-2013 out of 47 decisions of the European Commission on Article 102 cases, 27 of them where closed with commitments. Japan has introduced a similar commitment procedure in 2018. See also Gautier and Petit (2018), p.213-6.

the firm face investigation costs, and the firm moreover faces pecuniary sanctions), but it delivers deterrence and screening, although to a lesser extent as priors become less pessimistic.

We then introduce (scalable) commitments that reduce the impact of the practice on welfare as well as on the firm's profit, and consider two alternative policy regimes. The first one corresponds to the procedure established by art. 9.1 of Regulation 1/2003, which gives the firm the initiative in offering and designing any commitment. Consequently, the firm either offers no commitment, or it offers the minimal commitment that the enforcer is willing to accept rather than running an investigation. Three equilibria may arise. In particular, there always exists an equilibrium in which the firm never offers any commitment; the outcome is therefore the same as when commitments are not available. More interestingly, for sufficiently pessimistic priors there also exists a semi-separating equilibrium in which the firm offers a commitment (with some probability) when the action is harmful, whereas the enforcer accepts the commitment whenever offered and otherwise runs an investigation with a probability lower than 1. We show that, whenever this equilibrium is selected, the introduction of commitments undermines enforcement by reducing both deterrence and screening. Finally, for less pessimistic priors, there exists a pooling equilibrium in which, regardless of the type of action, the firm offers the minimal commitment that the enforcer is willing to accept. By construction, this equilibrium entails no screening whatsoever (as both types obtain the same payoff and thus undertake the action with the same probability). Moreover, the enforcer's indifference between accepting the commitment and running an investigation implies that the two regimes are equally cost-effective. Hence, commitments may be socially desirable only if they substantially improve deterrence. Building on these observations, we provide conditions under which enforcement is again undermined compared to the benchmark regime.

The second regime corresponds to a possible reform of the regulation, giving the enforcer the initiative in offering and designing any commitment. Consequently, the enforcer either offers no commitment, or it offers the maximal commitment that a firm is willing to accept to avoid an investigation. We find that there exists an essentially unique equilibrium, the nature of which depends on the enforcer's beliefs. For pessimistic beliefs, a separating equilibrium arises, in which the enforcer offers a high commitment that is accepted if the action is anti-competitive and otherwise rejected, prompting an investigation. Compared with the benchmark regime, this equilibrium performs equally well in terms of deterrence and screening, but is more effective ex-post, as it replaces costly investigations with the maximal acceptable commitment when the action is anti-competitive. For optimistic beliefs, a pooling equilibrium instead arises, in which the enforcer offers a lower commitment that is always accepted. As a result, the enforcer obtains some deterrence, when the other regulations would generate no enforcement whatsoever. For intermediate beliefs, the enforcer randomizes between high and low commitments, in such a way that there is more deterrence and screening as well as improved ex-post effectiveness than in the other regimes. In other words, the reformed regulation weakly dominates the benchmark regime and the current regulation in all three enforcement dimensions, and always performs strictly better in at least one dimension.

Related literature. To the best of our knowledge, only two papers have formally studied the impact of commitments on antitrust enforcement. Choné *et al.* (2014) consider a binary setting that restricts attention to all or nothing commitments, in which the enforcer can decide whether to allow commitments when a case is opened; doing so allows a faster termination of the infringement but weakens deterrence, as the firm avoids paying any fine. They characterize the solution to this trade-off when the enforcer can announce ex-ante its policy, so as to influence firm behavior, and show that this optimal policy may not be credible if the agency can deviate ex post from the announced policy. Gautier and Petit (2018) focus on ex-post enforcement, in a setting in which the enforcer can credibly choose between arbitrary levels of commitments and full-fledged investigations. They show that commitments can be used to discriminate firms according to the social harm of their practices. They also note that this use of commitments can weaken deterrence.

We contribute to this literature by studying the implications of a change of policy regime, reflecting the modernization of the European proceedings, that allow firms to offer commitments, if they wish so, whenever a case is opened. Compared with the binary setting considered by Choné *et al.*, giving the firm the choice of arbitrary levels of commitments tends to limit the enforcer's ability to discriminate different types of practices and further undermines deterrence.[3] As a result, introducing commitments may not be desirable.[4] Moreover, we consider a possible reform giving the enforcer the initiative to propose commitments, and find that it would significantly improve enforcement.

The paper is organized as follows. Section 2 presents the model, and section 3 considers the benchmark case without commitments. Section 4 analyses the case where commitments are introduced, whereas Section 5 studies the proposed reformed regulation. Conclusions follow. All the proofs are in the Appendix.

[3] Another difference is that, in Choné *et al.*, the enforcer can always choose (either ex-ante, or ex-post) to rule out commitments; by contrast, we consider their availability to be determined (ex-ante) by the policy regime.

[4] For informal policy discussions, see, e.g., Wils (2006,2008) and Mariniello (2014). The issues raised by commitments are also related to the literature on settlements, initiated by Shavell (1982) with a focus on litigation costs, and further developed by Bebchuk (1984) and Reinganum and Wilde (1986) by accounting for asymmetric information, and by Polinsky and Rubinfeld (1988) by studying the implications for ex-ante deterrence – for a comprehensive survey, see, e.g., Daughety and Reinganum (2011).

2 The model

We present here the setting, which captures the key relevant ingredients of the antitrust enforcer's problem according to the current regulation on commitments. The enforcer has limited evidence about a potential case, and must decide whether to open it. If it does so, then after a preliminary phase, the enforcer may close the case or proceed with an investigation, which is costly but may bring decisive evidence. Alternatively, the firm may offer a commitment; if accepted by the enforcer, it becomes compulsory and the case is closed with no sanctions nor guilty verdict.

Players, information and outcomes. The *enforcer* implements the antitrust policy and the *firm* can undertake an *action* (business practice) that has uncertain private and social impacts. Specifically, the action yields gross profits normalized to 1 and costs c, which is distributed according to a c.d.f. $F(c)$ over $[0,1]$, with atomless density $f(c)$ and decreasing hazard rate $h(\cdot) = f(\cdot)/F(\cdot)$. With probability $\lambda \in (0,1)$, the action is good ($\theta = G$), in which case welfare increases by $W_G = W > 0$. Otherwise, the action is bad ($\theta = B$) and produces a social loss $W_B = -L < 0$.[5] The welfare impact of the action, W_θ, is independent of the private cost c.[6] The firm observes the action type, $\theta \in \Theta$, and its cost, c, whereas the enforcer only knows their *ex-ante* probability distributions, λ and $F(\cdot)$. In what follows, the type θ will refer interchangeably to the action or the firm undertaking it.

Policy tools and actions. If the firm undertakes the action, the enforcer decides whether to *open a case.* If it does so, the firm chooses which *commitment* $C \in [0,1]$ to offer. The enforcer then decides whether to *accept* the offered commitment or *open an investigation.*[7] We adopt the convention that offering $C = 0$ means "offering no commitment" and accepting $C = 0$ means "closing the case". We moreover interpret the current regulation as requiring the enforcer to accept positive commitments whenever the resulting expected welfare matches that of the best alternative. That is, we do not allow the regulator to randomize over its decision when indifferent between accepting or not a positive commitment.[8]

Payoffs. If a commitment C is accepted, the case is closed with no sanction nor guilty verdict; profit and welfare are cut proportionally and in the same manner: the profit of

[5] An alternative interpretation refers to a population of firms with heterogeneous costs and actions that may have a positive or negative impact on social welfare. $F(c)$ and λ then refer to the distribution of types.

[6] Once it has been sunk in stage 0, the cost c reduces profit and welfare but has no incidence on subsequent decisions; in what follows, we will thus concentrate on the expressions of profit and welfare that are gross of the cost c.

[7] Allowing the enforcer to reject the commitment and close the case would not affect the analysis.

[8] Allowing for such randomization could generate additional semi-separating equilibria, which however are Pareto-dominated from the firm' standpoint; it could also generate additional pooling equilibria, which cannot be Pareto-ranked without introducing additional structure on payoff functions.

the firm is reduced to $1 - C$ and the welfare impact becomes $(1 - C) W_\theta$. Running an investigation costs instead $k > 0$, with sk to the firm and $(1 - s) k$ to the enforcer (where $s \in [0, 1]$), but enables the enforcer to obtain hard evidence with probability ρ when the action is bad (i.e., $\theta = B$). Obtaining hard evidence enables the enforcer to ban the practice and impose a *sanction* S. Otherwise, the enforcer must close the case.[9]

Remark (discounting). We do not explicitly consider the time dimension of the enforcement process. Profit and welfare can be interpreted as discounted values of the corresponding streams. The sanction S, which in practice occurs at a later stage of the process, can be interpreted as the pecuniary fine net of the profits gained in the meantime.

Timing. The timing of the game is as follows:

Stage 0: Nature draws the type $\theta \in \{G, B\}$ and the cost $c \in [0, 1]$, which are privately observed by the firm; the firm then decides whether to undertake the action. If the firm does not undertake the action, the game is over.

Stage 1: If an action is undertaken, the enforcer decides whether to open a case. If it does not, the game is over; otherwise, the firm then offers a commitment $C \in [0, 1]$.

Stage 2: Having observed the offered commitment, the enforcer chooses between accepting it (A) or proceeding to an investigation (I).

Equilibrium. This setting corresponds to a *signalling game* between an informed player (the firm) and an uninformed player (the enforcer), in which the informed player moves first. We will look for the *Perfect Bayesian Equilibria* of this game. If multiple equilibria are Pareto ordered from the point of view of the two types of firm, we will select the Pareto dominant one.

Throughout the paper we adopt the following assumptions:

Assumption 1: $S < (1 - sk - \rho)/\rho$.
This assumption ensures that the firm would always undertake the action (even when bad) if it were costless.

Assumption 2: $k < \rho L$.
This assumption states that running an investigation is socially desirable when the practice is bad; there is therefore room for enforcement when the practice is expected to be sufficiently harmful.

[9] The enforcer may thus commit a type-II error (acquitting a firm despite a bad action), but no type-I error (convicting a firm for a good action).

Assumption 3: $k \geq W$.

This assumption states that running an investigation is so costly that expected welfare is then always negative.[10]

3 Benchmark: No commitments

We start by studying antitrust enforcement in the absence of commitments. In this case, the last two stages of the above timing boil down to:

Stage 1 **(no commitments):** If an action is undertaken, the enforcer then chooses whether to proceed to an investigation or not; for expositional purposes, we interpret the latter as closing the case.

Let $\bar{C}_\theta$ denote the cost of being investigated for type $\theta \in \Theta$. This corresponds to the cost sk of standing up in a case and, for the bad type, also includes the expected lost profit and fine:

$$\bar{C}_G \equiv sk < \bar{C}_B \equiv sk + \rho (1 + S) < 1,$$

where the inequality stems from Assumption 1.

If the firm expects the enforcer to open an investigation with probability $\iota \in [0, 1]$, then undertaking the action gives type $\theta \in \Theta$ an expected profit equal to $1 - \iota\bar{C}_\theta$. Type θ therefore undertakes the action with probability

$$P_\theta(\iota) \equiv F(1 - \iota\bar{C}_\theta), \tag{1}$$

where $P_G(\iota) \geq P_B(\iota)$, with strict inequality when $\iota > 0$, as investigations are more costly for the bad type.

At the beginning of stage 1, upon observing that an action has indeed been undertaken, the enforcer updates its beliefs to:

$$\lambda_1 = \lambda_1^e(\lambda, \iota) \equiv \frac{\lambda P_G(\iota)}{\lambda P_G(\iota) + (1 - \lambda) P_B(\iota)} \geq \lambda, \tag{2}$$

with strict inequality if $\iota > 0$. The posterior $\lambda_1^e(\lambda, \iota)$ increases with λ; furthermore, as investigations are more costly for bad types, it also increases with ι.[11]

If the enforcer closes the case (superscript "c"), expected welfare is then equal to:

$$W^c(\lambda_1) \equiv \lambda_1 W - (1 - \lambda_1) L = (W + L)(\lambda_1 - \lambda^c), \tag{3}$$

[10] This assumption rules out equilibria in which the enforcer would accept low commitments and reject higher ones. See Online appendix A.

[11] See the proof of Proposition 1 for a formal derivation.

with

$$\lambda^c \equiv \frac{L}{W+L} \in (0,1)\,. \tag{4}$$

If instead the enforcer runs an investigation (superscript "i") expected welfare is

$$W^i(\lambda_1) \equiv \lambda_1 W - (1-\lambda_1)(1-\rho)L - k = W^c(\lambda_1) + \rho L(\lambda^{ci} - \lambda_1), \tag{5}$$

with

$$\lambda^{ci} \equiv 1 - \frac{k}{\rho L} \in (0, \lambda^c)\,,$$

where the inequalities $\lambda^{ci} > 0$ and $\lambda^{ci} < \lambda^c$ stem, respectively, from Assumptions 2 and 3.[12] It follows that the enforcer prefers to investigate for $\lambda_1 < \lambda^{ci}$, and to close the case for $\lambda_1 > \lambda^{ci}$.

Three relevant cases can therefore be distinguished for the equilibrium enforcement policy. If $\lambda > \lambda^{ci}$, then from (2), $\lambda_1 \geq \lambda > \lambda^{ci}$, and so the enforcer always closes the case; anticipating this, the firm always undertakes the action, regardless of its type of action.[13] It follows that $\lambda_1 = \lambda$. Likewise, in the boundary case $\lambda = \lambda^{ci}$, investigating with positive probability would yield $\lambda_1 > \lambda = \lambda^{ci}$, a contradiction; hence, the enforcer must still close the case, and $\lambda_1 = \lambda = \lambda^{ci}$.

Consider next the case where $\lambda < \lambda^N$, where $\lambda^N \in \left(0, \lambda^{ci}\right)$ is the unique solution to

$$\lambda_1^e(\lambda^N, 1) = \lambda^{ci}. \tag{6}$$

We then have $\lambda_1 = \lambda_1^e(\lambda, \iota) \leq \lambda_1^e(\lambda, 1) < \lambda^{ci}$, implying that the enforcer investigates with probability 1. Likewise, in the boundary case $\lambda = \lambda^N$, investigating with probability less than 1 would yield $\lambda_1 < \lambda_1^e(\lambda, 1) = \lambda^{ci}$, a contradiction; hence, the enforcer must still investigate with probability 1, and $\lambda_1 = \lambda_1^e(\lambda, 1) = \lambda^{ci}$.

Finally, for $\lambda \in (\lambda^N, \lambda^{ci})$, the enforcer must randomize: if it were to investigate with probability 1, its revised belief $\lambda_1 = \lambda_1^e(\lambda, 1) > \lambda_1^e(\lambda^N, 1) = \lambda^{ci}$ would be too optimistic to warrant intervention; if instead it were to close the case with probability 1, $\lambda_1 = \lambda_1^e(\lambda, 0) = \lambda < \lambda^{ci}$ would be too pessimistic and trigger intervention. The equilibrium probability of investigation must therefore induce the enforcer to be indifferent between investigating or not: $\iota = \tilde{\iota}(\lambda)$, the unique solution to

$$\lambda_1^e(\lambda, \iota) = \lambda^{ci}. \tag{7}$$

Building on this yields:

[12] $\lambda^{ci} < \lambda^c$ amounts to $\rho LW < k(W+L)$, which holds under Assumption 3.

[13] As the cost c is distributed with atomless density over $[0,1]$, the decision of the firm in the particular case where $c = 1$ does not materially affect the analysis.

Proposition 1 (no commitments) *In the benchmark case when commitments are not allowed, there exists a unique equilibrium, $\mathcal{E}_N$, in which the enforcer opens an investigation with probability:*

$$\iota^N(\lambda) \equiv \begin{cases} 1 & \text{if } \lambda \in (0, \lambda^N] \\ \tilde{\iota}(\lambda) \in (0,1) & \text{if } \lambda \in (\lambda^N, \lambda^{ci}) \\ 0 & \text{if } \lambda \in [\lambda^{ci}, 1) \end{cases}, \tag{8}$$

where $\tilde{\iota}(\lambda)$ decreases continuously from 1 for $\lambda = \lambda^N$ to 0 for $\lambda = \lambda^{ci}$. As a result, type θ obtains an expected payoff equal to

$$\Pi_\theta^N(\lambda) \equiv 1 - \iota^N(\lambda)\bar{C}_\theta.$$

Proof. See Appendix A. ■

Hence, the enforcer intervenes less often as it becomes more optimistic; this, in turn, encourages both types of firm to undertake the action. In particular, there is maximal enforcement (and, thus, minimal participation) if $\lambda \leq \lambda^N$, and no enforcement (and full participation) if $\lambda \geq \lambda^{ci}$.

Corollary 1 (participation – no commitments) *The firm undertakes an action of type θ with probability $P_\theta^N(\lambda) \equiv F(\Pi_\theta^N(\lambda))$, which is continuous in λ and equal to $F(1)$ for $\lambda \geq \lambda^{ci}$. For $\lambda < \lambda^{ci}$, participation is strictly increasing in λ and higher for the good action: $P_G^N(\lambda) > P_B^N(\lambda)$. Expected participation, given by*

$$P^N(\lambda) \equiv \lambda P_G^N(\lambda) + (1-\lambda)P_B^N(\lambda), \tag{9}$$

is thus also strictly increasing in λ for $\lambda \in \left(0, \lambda^{ci}\right)$.

Proof. See Appendix B. ■

Enforcement has a *deterrence* effect, as it reduces the participation of both types of firm (e.g., through the litigation costs sk). It has also a *screening* effect, as bad types are further discouraged through the expected sanction and lost profit $\rho(1+S)$. Both effects increase with the probability of investigation, which in turn depends on the enforcer's prior beliefs. As a result, it is maximal for $\lambda \leq \lambda^N$, where the enforcer investigates for sure, decreases in λ in the range $(\lambda^N, \lambda^{ci})$, and vanishes for $\lambda \geq \lambda^{ci}$, where the enforcer stops investigating.

In equilibrium, expected welfare is given by:

$$\mathcal{W}_N(\lambda) \equiv \begin{cases} \lambda P_G^N(\lambda)\left[W - \iota^N(\lambda)k\right] \\ -(1-\lambda)P_B^N(\lambda)\left[(1-\iota^N(\lambda)\rho)L + \iota^N(\lambda)k\right] & \text{if } \lambda \in (0, \lambda^{ci}), \\ \lambda W - (1-\lambda)L & \text{if } \lambda \in \left[\lambda^{ci}, 1\right). \end{cases}$$

The following corollary provides a useful expression of expected welfare for $\lambda \in (0, \lambda^{ci})$.

Corollary 2 (expected welfare – no commitment) *For $\lambda \in (0, \lambda^{ci})$, expected welfare can be expressed as*

$$\mathcal{W}_N(\lambda) \equiv P^N(\lambda) W^i(\lambda_1^N(\lambda)) \tag{10}$$

where

$$\lambda_1^N(\lambda) \equiv \lambda_1^e(\lambda, \iota^N(\lambda)) = \begin{cases} \lambda \frac{P_G^N(\lambda)}{P^N(\lambda)} \in (\lambda, \lambda^{ci}) & \text{if} \quad \lambda \in (0, \lambda^N), \\ \lambda^{ci} & \text{if} \quad \lambda \in \left[\lambda^N, \lambda^{ci}\right). \end{cases} \tag{11}$$

Furthermore, $\mathcal{W}_N(\lambda)$ is continuous in λ, increasing in λ for $\lambda < \lambda^N$ and decreasing in λ for $\lambda > \lambda^N$.

Proof. See Appendix C. ■

As welfare is normalized to zero when the action is not undertaken, expected welfare can be expressed as the probability of observing an action, measured by the participation $P^N(\lambda)$, multiplied by the resulting expected welfare, given the posterior belief $\lambda_1^N(\lambda)$. From Proposition 1, for $\lambda < \lambda^{ci}$ the enforcer either investigates for sure (if $\lambda \leq \lambda^N$), or is indifferent between investigating or not (if $\lambda \in (\lambda^N, \lambda^{ci})$). In both cases, the resulting expected welfare is given by $W^i(\lambda_1^N(\lambda))$.

Corollary 2 points out that, in the range $\lambda \in (\lambda^N, \lambda^{ci})$, expected welfare *decreases* as the enforcer becomes more optimistic. In this range, the enforcer is too optimistic for investigating with certainty, and too pessimistic for not intervening. It follows that, as the enforcer becomes more optimistic, in equilibrium not only the probability of investigation $\iota^N(\lambda)$ decreases, implying that participation increases, but the enforcer's revised belief is maintained to λ^{ci} (to ensure indifference between investigating or not), implying that each undertaken action generates an expected welfare equal to $W^i(\lambda^{ci}) < 0$.

Remark: too much or little enforcement? As just noted, enforcement deters good actions, as well as bad ones, from being undertaken. Depending on which force prevails, the level of enforcement may be excessive or insufficient. To see this, suppose that the enforcer can commit to a given investigation rate ι before the firm chooses whether to undertake the action. Given the firm's response, expected welfare becomes, as a function of ι,

$$\mathcal{W}(\iota) \equiv \lambda P_G(\iota)(W - \iota k) - (1 - \lambda) P_B(\iota) \left[(1 - \iota\rho)L + \iota k\right].$$

So its derivative is given by:

$$\begin{aligned} \mathcal{W}'(\iota) \quad &= \quad \lambda P_G'(\iota)(W - \iota k) - (1 - \lambda) P_B'(\iota) \left[(1 - \iota\rho) L + \iota k\right] \\ &\quad + (1 - \lambda) P_B(\iota)(\rho L - k) - \lambda P_G(\iota) k. \end{aligned}$$

When $\lambda \in (\lambda^N, \lambda^{ci})$, the equilibrium probability $\iota = \tilde{\iota}(\lambda)$ is such that the posterior $\lambda_1^e(\lambda, \iota)$ coincides with λ^{ci}, which in turn implies that the terms in the second line cancel out. Using $P'_\theta(\iota) = -\bar{C}_\theta f(1-\iota\bar{C}_\theta)$, it follows that slightly intensifying the enforcement activity (i.e., increasing ι slightly above $\tilde{\iota}(\lambda)$) would increase expected welfare whenever:

$$\frac{\lambda}{1-\lambda} < \frac{\bar{C}_B}{\bar{C}_G} \frac{f(1-\iota\bar{C}_B)}{f(1-\iota\bar{C}_G)} \frac{L-\iota(\rho L - k)}{W - \iota k}.$$

When instead this condition is not satisfied, there is over-enforcement: expected welfare would increase if the enforcer could commit itself to slightly reduce the frequency of investigations.

4 Commitments

We now revert to the current regulation setting, where commitments are available and are proposed by the firm, moving from stage 2 backwards.

4.1 The enforcer's response

Suppose that the firm undertook the action in stage 0 and offered $C \in [0,1]$ in stage 1, and let $\lambda_2(C)$ denote the enforcer's revised belief at the beginning of stage 2. In the continuation subgame the enforcer must either accept the commitment or launch an investigation. In case of acceptance, expected welfare is given by:

$$W^a(\lambda_2, C) \equiv (1-C)W^c(\lambda_2).$$

If instead the enforcer investigates, expected welfare is given by (5), with the caveat that it must be evaluated with the posterior λ_2. Let

$$\underline{\lambda}(C) \equiv \frac{\rho\lambda^{ci} - C}{\rho\lambda^c - C}\lambda^c. \tag{12}$$

denote the value of the revised belief λ_2 for which the enforcer is indifferent between accepting C or not (i.e., $W^a(\underline{\lambda}(C), C) = W^i(\underline{\lambda}(C))$). Note that $\underline{\lambda}(C)$ decreases from λ^{ci} to 0 as C increases from 0 to $\rho\lambda^{ci}$.

The following Lemma characterizes the enforcer's optimal response to the offered commitment.

Lemma 1 (enforcer's response) *In stage 2, for any offered commitment C and associated revised belief $\lambda_2(C)$:*

- *the enforcer accepts any $C \geq \hat{C} \equiv \rho\lambda^{ci}$ regardless of $\lambda_2(C)$;*

- *the enforcer accepts $C \in (0, \hat{C})$ if $\lambda_2(C) \geq \underline{\lambda}(C)$, and rejects it otherwise;*
- *the enforcer accepts $C = 0$ if $\lambda_2(C) > \lambda^{ci}$, rejects it if $\lambda_2(C) < \lambda^{ci}$, and is indifferent between the two if $\lambda_2(C) = \lambda^{ci}$.*

Proof. See Appendix D. ■

It follows from Lemma 1 that the enforcer accepts the offered commitment C whenever its revised belief $\lambda_2(C)$ is sufficiently optimistic, namely, higher than $\underline{\lambda}(C)$. In particular, the enforcer always accepts C whenever it induces a revised belief exceeding λ^{ci}, implying that it would rather close the case than investigate.[14] If instead the enforcer is more pessimistic (namely, $\lambda_2(C) < \lambda^{ci}$), then it would rather investigate than close the case, and both options moreover generate a negative expected welfare. It is therefore willing to accept the offered commitment only if it is high enough, namely, not lower than $\underline{C}(\lambda_2(C))$, where

$$\underline{C}(\lambda_2) \equiv \underline{\lambda}^{-1}(\lambda_2) = \rho\lambda^c \frac{\lambda^{ci} - \lambda_2}{\lambda^c - \lambda_2}$$

decreases from $\hat{C}$ to 0 as λ_2 varies from 0 to λ^{ci}. The minimal acceptable commitment for a given belief λ_2 is therefore $\underline{C}(\lambda_2)$ if $\lambda_2 \in [0, \lambda^{ci})$ and 0 otherwise, as illustrated by Figure 1.

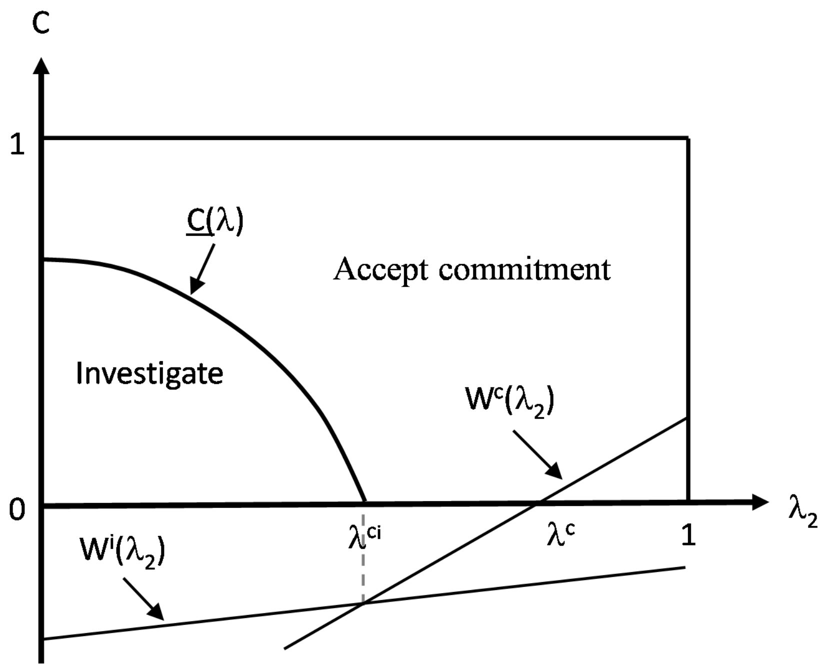

Figure 1.

[14] Recall that the latter option always yields a negative expected welfare, whereas the former one yields a positive welfare if the enforcer's belief exceeds λ^c. Thus, for $\lambda_2 \in (\lambda^{ci}, \lambda^c)$, we have $(1-C)W^c(\lambda_2) > W^c(\lambda_2) > W^i(\lambda_2)$; and for $\lambda_2 \geq \lambda^c$, we have $(1-C)W^c(\lambda_2) \geq 0 > W^i(\lambda_2)$.

Remark: on the efficiency of commitments. In a *complete information environment*, commitments would be an efficient way of dealing with bad actions. Indeed, when the action is known to be bad (i.e., $\lambda = 0$), the minimal commitment that the enforcer is willing to accept is equal to $\left(\hat{C} =\right) \underline{C}(0) = \rho - k/L$, which is lower than the maximal commitment that the firm is willing to offer, $\bar{C}_B = \rho + sk + \rho S$; there is therefore room for a mutually beneficial agreement. As we will see, *asymmetric information* substantially limits the efficiency of commitments.

Remark: minimal equilibrium payoffs. By construction, each type θ can secure $1 - \bar{C}_\theta$ by offering $\bar{C}_\theta$ (regardless of whether it is accepted or rejected). In addition, both types can secure $1 - \hat{C}$ by offering $\hat{C}$ – which the enforcer always accepts, regardless of its beliefs. Hence, in equilibrium, B obtains at least $1 - \hat{C} (> 1 - \bar{C}_B)$ and G obtains at least $1 - \min\{\hat{C}, \bar{C}_G\}$.

4.2 Pareto efficient equilibria

We now turn to the first two stages of the game. If in stage 0 the firm undertakes the action $\theta = G, B$ with probability P_θ, then at the beginning of stage 1, upon observing that an action has been undertaken, the enforcer updates its beliefs to :

$$\lambda_1 = \frac{\lambda P_G}{\lambda P_G + (1 - \lambda) P_B}. \tag{13}$$

Let denote by $\mathcal{C}_\theta$ the set of commitments offered by type $\theta = G, B$, and by $\nu_\theta(C)$ the probability of offering any given $C \in \mathcal{C}_\theta$. Upon observing $C \in \mathcal{C} \equiv \mathcal{C}_G \cup \mathcal{C}_B$, the enforcer revises further its beliefs to:

$$\lambda_2(C) \equiv \frac{\lambda_1 \nu_G(C)}{\lambda_1 \nu_G(C) + (1 - \lambda_1) \nu_B(C)}. \tag{14}$$

We first note that, as in the absence of commitments (see Proposition 1), there is scope for enforcement if and only if $\lambda < \lambda^{ci}$. Specifically, letting Π_θ denote type θ's equilibrium profit when commitments are available, we have:

Proposition 2 (scope for enforcement) *(i) If $\lambda \geq \lambda^{ci}$, then there is a unique Pareto efficient equilibrium outcome, in which the firm offers no commitment (i.e., $C = 0$) and the enforcer closes the case (i.e., accepts $C = 0$); as a result, the firm obtains its maximal payoff:* $\Pi_B = \Pi_G = 1$.

(ii) If instead $\lambda < \lambda^{ci}$, then any equilibrium is such that, with positive probability the enforcer investigates or accepts a positive commitment; as a result, the firm does not obtain its maximal payoff, all the more so if the action is bad: $\Pi_B \leq \Pi_G < 1$.

Proof. See Appendix E. ∎

In what follows we focus on the case where there is room for enforcement, that is, $\lambda < \lambda^{ci}$. The following lemma restricts the set of candidate equilibria.

Lemma 2 (candidate equilibria) *If $\lambda < \lambda^{ci}$, then in equilibrium either $\mathcal{C}_G = \mathcal{C}_B$ or $\mathcal{C}_G \subset \mathcal{C}_B$.*

Proof. See Appendix F ∎

Lemma 2 rules out the possibility that G reveals itself at stage 1 – in particular, it rules out fully separating equilibria. The intuition is that, by revealing itself, G would induce the enforcer to accept its offered commitment. This, in turn, implies that the offered commitment must be low (namely, lower than $\bar{C}_G$, as G can secure a payoff of $1 - \bar{C}_G$ by offering $\bar{C}_G$), which in turn gives B an incentive to mimic G. We are thus left with *pooling* equilibria, in which the two types use the same support, and *semi-separating* equilibria in which G's support is strictly included in B's.[15]

4.2.1 Semi-separating equilibria

We first consider the semi-separating equilibria in which, in stage 1, the two types of firm may offer different commitments (i.e., $\mathcal{C}_G \neq \mathcal{C}_B$); from Lemma 2, it follows that, with positive probability B reveals itself at stage 1 (i.e., $\mathcal{C}_G \subset \mathcal{C}_B$). Furthermore, conditional on doing so, the best strategy for B is to offer $\hat{C}$.[16] Hence, we can restrict attention to semi-separating equilibria in which $\mathcal{C}_B \setminus \mathcal{C}_G = \{\hat{C}\}$. Note that, by construction, B's expected payoff is therefore equal to $1 - \hat{C}$.

If in equilibrium the enforcer were offered another positive commitment C, then the policy rule would require the enforcer to investigate or to accept C with probability one; hence, B could not be indifferent between offering C and $\hat{C}$, a contradiction. It follows that $C_G = \{0\}$; that is, the good type offers no commitment, whereas the bad type randomizes between offering $\hat{C}$ or no commitment. Furthermore, to leave B indifferent between offering $\hat{C}$ and no commitment, the latter option must induce the enforcer to

[15] Pooling and semi-separating equilibria usually refer to equilibria in which $\mathcal{C}_G \cap \mathcal{C}_B$ is a singleton; as we will see, the Pareto-efficient equilibria do have this property.

[16] Recall that B can secure $1 - \hat{C}$ by offering $\hat{C}$, which the enforcer always accepts; conversely, $\hat{C}(= \hat{C})$ is the minimal commitment that the enforcer is willing to accept when B reveals itself.

investigate with probability (where the superscript S refers to semi-separating equilibria)

$$\iota^S \equiv \frac{\hat{C}}{\bar{C}_B}.$$

As both types offer no commitment with positive probability, the equilibrium expected payoffs are $\Pi_\theta = 1 - \iota^S \bar{C}_\theta$. It follows that, in stage 0, the firm is less likely to undertake the action when it is bad; specifically, the participation rates are given by:

$$P_\theta^S \equiv P_\theta\left(\iota^S\right), \tag{15}$$

where $P_\theta(\cdot)$ is defined by (1), and $\bar{C}_B > \bar{C}_G$ implies $P_B^S < P_G^S$.

Finally, to induce the enforcer to randomize between closing the case and opening an investigation, offering no commitment must induce a posterior $\lambda_2(0) = \lambda^{ci}$, where, using (13), (14) and $\nu_G(0) = 1$:

$$\lambda_2(0) = \frac{\lambda P_G^S}{\lambda P_G^S + (1-\lambda) P_B^S \nu_B(0)},$$

where the right-hand side is decreasing in $\nu_B(0)$. It is therefore minimal for $\nu_B(0) = 1$, where it coincides with $\lambda_1^e(\lambda, \iota^S)$, which increases from 0 to 1 with λ. The equilibrium condition $\lambda_2(0) = \lambda^{ci}$ thus imposes an upper bound on the prior λ, which we will denote by λ^S and is such that:

$$\lambda_1^e(\lambda^S, \iota^S) = \lambda^{ci}, \tag{16}$$

which amounts to:

$$\lambda^S \equiv \frac{\lambda^{ci} P_B^S}{\lambda^{ci} P_B^S + (1-\lambda^{ci}) P_G^S} \quad (< \lambda^{ci}). \tag{17}$$

The following proposition shows that, conversely, as long as $\lambda < \lambda^S$ there exists a semi-separating equilibrium $\mathcal{E}_C^S$ (where C stands for current regulation and S for semi-separation), which is, moreover, unique:

Proposition 3 (semi-separating equilibrium) *If $\lambda \in [\lambda^S, \lambda^{ci})$, there is no semi-separating equilibrium. If instead $\lambda < \lambda^S$, there exists a unique semi-separating equilibrium, $\mathcal{E}_C^S$, in which:*

- *G offers no commitment (i.e., $C = 0$), whereas B randomizes between offering $\hat{C}$ and no commitment, in such a way that $\lambda_2(0) = \lambda^{ci}$;*
- *the enforcer accepts $\hat{C}$ and, when offered no commitment, investigates with probability $\iota^S = \hat{C}/\bar{C}_B$.*

In this equilibrium, for any $\lambda < \lambda^S$, the expected payoffs of the firm are given by

$$\Pi_G^S \equiv 1 - \hat{C}\bar{C}_G/\bar{C}_B > \Pi_B^S \equiv 1 - \hat{C} > 0.$$

Proof. See Appendix G. ■

The following corollary characterizes the expected welfare generated by the semi-separating equilibrium:

Corollary 3 (expected welfare – semi-separating equilibrium) *If $\lambda < \lambda^S$, the unique semi-separating equilibrium yields participation rates given by (15) and an expected welfare given by:*

$$\mathcal{W}_C^S(\lambda) \equiv P^S(\lambda) W^i(\lambda_1^S(\lambda)), \tag{18}$$

where

$$P^S(\lambda) \equiv \lambda P_G^S + (1-\lambda) P_B^S \text{ and } \lambda_1^S(\lambda) \equiv \frac{\lambda P_G^S}{P^S(\lambda)} \in (0, \lambda^{ci}) \tag{19}$$

respectively denote the expected participation and the enforcer's interim belief, upon observing that the action has been undertaken.

Proof. See Appendix H. ■

As before, expected welfare can be expressed as the probability of observing an action, measured here by the participation $P^S(\lambda)$, multiplied by the resulting expected welfare, given the interim belief $\lambda_1^S(\lambda)$. From Proposition 3, in the semi-separating equilibrium either B reveals itself and offers $\hat{C}$, or both types offer no commitment, in such a way that the enforcer is always indifferent between investigating or not. Hence, the resulting expected welfare is given by $W^i(.)$.

4.2.2 Pooling equilibria

We now turn to the equilibria in which, in stage 1, both types of firm offer the same commitments (i.e., $\mathcal{C}_G = \mathcal{C}_B$). We first note that, if the enforcer is sufficiently optimistic, namely, for $\lambda \geq \lambda^S$, the equilibrium $\mathcal{E}_N$ characterized by Proposition 1 survives when commitments become available. To see why, recall first that, as λ increases to λ^S, the semi-separating equilibrium $\mathcal{E}_C^S$ from the previous section is such that B is indifferent between offering $\hat{C}$ or no commitment, and chooses the latter with a probability, $\nu_B(0)$, that induces a posterior equal to λ^{ci}. B's indifference, is in turn ensured by the enforcer investigating with probability $\iota^S \equiv \hat{C}/\bar{C}_B$ when no commitment is offered. As λ tends to λ^S, the probability $\nu_B(0)$ tends to 1 and the equilibrium thus converges towards a pooling equilibrium where both types offer no commitment and the interim and posterior beliefs coincide: $\lambda_2(0) = \lambda_1^S(\lambda^S) = \lambda^{ci}$. It follows that $\iota^S = \tilde{\iota}(\lambda^S)$, where $\tilde{\iota}(\cdot)$ is the enforcer's probability of investigation in the $\mathcal{E}_N$ equilibrium with no commitments, which is precisely designed to induce an interim belief λ_1 equal to λ^{ci} (cf. (7)). Conversely, offering no commitment, together with the enforcer's investigating with probability $\tilde{\iota}(\lambda^S)$ and the participation rates given by (15), constitutes an equilibrium when commitments are available, as B is then indifferent between deviating, by offering the minimal acceptable commitment $\hat{C}$, or not (because $\tilde{\iota}(\lambda^S) = \iota^S$), and G strictly prefers to offer no commitment; furthermore, any other deviation can be deterred by interpreting it as signalling a bad type. In other words, for $\lambda = \lambda^S$, $\mathcal{E}_N$ remains an equilibrium (with appropriately expanded strategies) when commitments become available; as $\iota^N(\lambda) = \tilde{\iota}(\lambda)$ is decreasing in λ, $\mathcal{E}_N$ remains an equilibrium for $\lambda > \lambda^S$, where deviations become even less attractive.

However, there also exists another pooling equilibrium $\mathcal{E}_C^P$ (with P for pooling), in which the firm offers the minimal acceptable commitment $\underline{C}(\lambda)$ and thus obtains $1 - \underline{C}(\lambda)$. This equilibrium exists as long as G (which has less to lose from an investigation) is not tempted to deviate and offer no commitment (or an unacceptable one) and be investigated; this is the case as long as $\underline{C}(\lambda) \leq \bar{C}_G$, or $\lambda \geq \underline{\lambda}(\bar{C}_G)$, where $\underline{\lambda}(\cdot)$ is given by (12).

The next proposition shows that the two pooling equilibria just described are the only possible Pareto-efficient ones:

Proposition 4 (Pareto efficient pooling equilibrium)

(i) If $\lambda \in [\lambda^S, \lambda^{ci})$, the equilibrium $\mathcal{E}_N$ arising in the absence of commitment constitutes a pooling equilibrium.[17]

(ii) If $\lambda \in \left[\underline{\lambda}(\bar{C}_G), \lambda^{ci}\right)$ there is also a continuum of pooling equilibria in which the firm offers a commitment $C \in [\underline{C}(\lambda), \min\{\hat{C}, \bar{C}_G\}]$, which is accepted; among them, the Pareto-efficient equilibrium is $\mathcal{E}_C^P$, in which the firm offers the minimum acceptable commitment $\underline{C}(\lambda)$ and obtains $\Pi^P(\lambda) \equiv 1 - \underline{C}(\lambda)$.

Proof. See Appendix I. ■

By construction, the equilibrium $\mathcal{E}_N$ yields an expected welfare equal to $\mathcal{W}_N(\lambda)$, characterized by Corollary 2. The following corollary characterizes instead the expected welfare generated by the second pooling equilibrium $\mathcal{E}_C^P$:

Corollary 4 (expected welfare – pooling equilibrium) *The pooling equilibrium $\mathcal{E}_C^P$ yields an expected welfare given by:*

$$\mathcal{W}_C^P(\lambda) \equiv P^P(\lambda) W^i(\lambda),$$

[17] We slightly abuse notation here: the equilibrium strategies are indeed the same as in the absence of commitments, but they now survive a richer set of potential deviations.

where

$$P^P(\lambda) \equiv F(1 - \underline{C}(\lambda))$$

denotes the equilibrium participation.

Proof. See Appendix J. ■

The expected welfare can thus again be expressed as the probability of observing the action, multiplied by the expected welfare generated by an investigation. Furthermore, as both types of firm obtain the same profit in this pooling equilibrium, there is no updating and the expected welfare from an investigation is thus evaluated at the prior λ.

4.3 Comparison of policy regimes

We now assess the performance of the current regulation (C), compared with the benchmark setting where commitments are not available (N). To this aim, we focus on the Pareto-efficient equilibria among the new ones, namely, the semi-separating equilibrium $\mathcal{E}_C^S$ and the pooling equilibrium $\mathcal{E}_C^P$. We further restrict attention to the case where there is indeed scope for enforcement in both regimes (i.e., $\lambda < \lambda^{ci}$).

The expected welfare generated by policy regime $\tau = N, C$ can be expressed as

$$\mathcal{W}_\tau(\lambda) = P_\tau(\lambda) W_1^\tau(\lambda_1^\tau(\lambda)),$$

where $P_\tau(\lambda)$ denotes total participation, $\lambda_1^\tau(\lambda)$ denotes the enforcer's interim belief, and $W_1^\tau(\cdot)$ denotes the expected welfare, conditional on the enforcer's interim belief. This formulation enables us to distinguish three channels through which the regulation may potential improve enforcement: ex-ante it affects the decision of the two types to undertake the action and it may reduce the participation $P_\tau(\lambda)$ (*deterrence*) and increase the interim belief $\lambda_1^\tau(\lambda)$ by treating bad actions more harshly (*screening*); moreover, it may improve the expected welfare $W_1^\tau(\lambda_1)$, given the interim beliefs (*ex-post effectiveness*).

The following proposition shows that commitments are never desirable if they give rise to the semi-separating equilibrium.

Proposition 5 (semi-separating equilibrium vs. benchmark) *In the relevant range $\lambda \in (0, \lambda^S)$ in which the semi-separating equilibrium $\mathcal{E}_C^S$ exists, $\mathcal{W}_C^S(\lambda) < \mathcal{W}_N(\lambda)$.*

Proof. See Appendix K. ■

In the benchmark regime, the expected welfare is given by $\mathcal{W}_N(\lambda) = P^N(\lambda) W^i\left(\lambda_1^N(\lambda)\right)$. When commitments give rise to the semi-separating equilibrium, it is instead equal to $\mathcal{W}_C^S(\lambda) = P^S(\lambda) W^i\left(\lambda_1^S(\lambda)\right)$. In addition, in both regimes, participation rates are based on the probability of investigation: in the benchmark regime, both types face the same probability, $\iota^N(\lambda)$; under the current regulation, G faces the probability ι^S and B is indifferent between facing that probability or offering the commitment $\hat{C}$. Furthermore, the benchmark regime yields a higher probability: $\iota^N(\lambda) > \iota^S$ in the relevant range $\lambda < \lambda^S$. This is because, when commitments are available, B can secure a profit $1-\hat{C}\left(> 1 - \bar{C}_B\right)$ by offering $\hat{C}$; this option limits in turn, the enforcer's probability of investigation when instead B offers no commitment. It follows that the benchmark regime performs better in terms of both deterrence $(P^N(\lambda) < P^S(\lambda))$ and screening $(\lambda_1^N(\lambda) > \lambda_1^S(\lambda))$. Finally, as the firm has the initiative when designing the commitments, it appropriates all the benefits of enforcement cost savings whenever a commitment is indeed implemented; as a result, given the enforcer's interim belief λ_1, expected welfare is in both cases equal to $W^i(\lambda_1)$. Hence, the current regulation fails to improve ex-post effectiveness.

We now turn to the case where commitments give rise to the pooling equilibrium $\mathcal{E}_C^P$. The next proposition identifies a number of situations in which the benchmark regime yields a higher expected welfare:

Proposition 6 (pooling equilibrium vs. benchmark) *There exists* $\underline{h} > 0$, $\overline{h} > \underline{h}$ *and* $\overline{\lambda} < \lambda^{ci}$ *such that, in the relevant range* $\lambda \in \left[\underline{\lambda}(\bar{C}_G), \lambda^{ci}\right)$ *in which the pooling equilibrium* $\mathcal{E}_C^P$ *exists,* $\mathcal{W}_C^P(\lambda) < \mathcal{W}_N(\lambda)$ *whenever:*

(i) $\lambda \leq \lambda^N$*;*

(ii) $\lambda > \lambda^N$ *and* $h\left(1 - \underline{C}(\lambda^N)\right) < \underline{h}$*;*

(iii) $\lambda > \overline{\lambda}$ *and* $h(1) < \overline{h}$*.*

Proof. See Appendix L. ■

When commitments give rise to the pooling equilibrium, the enforcer accepts the commitment $\underline{C}(\lambda)$ and expected welfare is given by $\mathcal{W}_C^P(\lambda) = P^P(\lambda) W^i(\lambda))$.

By construction, the equilibrium $\mathcal{E}_C^P$ features no screening, as it gives the same payoff to both types and thus induces the same rate of participation: $P_B^P(\lambda) = P_G^P(\lambda) = F(1 - \underline{C}(\lambda))$, leading to $\lambda_1 = \lambda$. By contrast, the benchmark regime always features some screening (although to a lesser extent as the enforcer becomes more optimistic).

Furthermore, the two regimes perform again equally in terms of ex-post effectiveness: as the firm offers the minimum acceptable commitment that leaves the enforcer indifferent between accepting it or investigating, expected welfare, for given interim beliefs, is the same in the two regimes: $W^i(\lambda_1)$.

It follows that commitments can be desirable only if they enhance deterrence, and do so to an extent large enough to compensate the loss of screening. To study this, we must compare the participation in the benchmark regime, $P^N(\lambda) = \lambda F(1 - \iota^N(\lambda)\bar{C}_G) + (1-\lambda)F(1-\iota^N(\lambda)\bar{C}_B)$, with that under the current regulation, $P^P(\lambda) = F(1 - \underline{C}(\lambda))$. Recall that the pooling equilibrium exists for $\lambda \geq \underline{\lambda}(\bar{C}_G)$, which amounts to $\underline{C}(\lambda) \leq \bar{C}_G$.

For $\lambda \leq \lambda^N$, in the benchmark regime the enforcer investigates with probability 1, implying that both types would prefer to offer $\underline{C}(\lambda)$. It follows that commitments cannot enhance deterrence when $\lambda \leq \lambda^N$: we then have $P^P(\lambda) = F(1 - \underline{C}(\lambda)) > F\left(1 - \bar{C}_\theta\right) = P_\theta^N(\lambda)$.

If $\lambda > \lambda^N$, in the benchmark regime the enforcer opens an investigation with probability $\tilde{\imath}(\lambda)$, which is decreasing in λ – the specific pattern of $\tilde{\imath}(\lambda)$ depends also on the distribution of private costs $f(c)$, as (26) shows. As a result, the participation, $P^N(\lambda) = \lambda P_G(\tilde{\imath}(\lambda)) + (1-\lambda) P_B(\tilde{\imath}(\lambda))$, is increasing in λ and equal to 1 for $\lambda \to \lambda^{ci}$ since $\tilde{\imath}\left(\lambda^{ci}\right) = 0$. The participation under the current regulation exhibits a similar pattern: $\underline{C}(\lambda)$ is decreasing in λ and $\underline{C}(\lambda^{ci}) = 0$. Hence, when λ increases above λ^N, deterrence is reduced in both regimes, all the more so as the enforcer becomes more optimistic.

Building on these insights, Proposition 6 provides sufficient conditions ensuring that expected welfare remains higher in the absence of commitments. Since the participation in the benchmark regime depends, through $\tilde{\imath}(\lambda)$, also on the distribution of private costs $f(c)$, we can identify such conditions in terms of the hazard rate $h(.)$. Specifically, as $\mathcal{W}_N(\lambda)$ is decreasing in λ in the range $\lambda \in \left(\lambda^N, \lambda^{ci}\right)$, a "global" sufficient condition is for $\mathcal{W}_C^P(\lambda)$ to be instead increasing in that range, that is:

$$\frac{d\mathcal{W}_C^P}{d\lambda}(\lambda) = \frac{dP^P}{d\lambda}(\lambda) W^i(\lambda) + P^P(\lambda) \frac{dW^i}{d\lambda}(\lambda) > 0,$$

where $P^P(\lambda) = F(1 - \underline{C}(\lambda))$ and $dP^P(\lambda)/d\lambda$ is therefore proportional to $f(1 - \underline{C}(\lambda))$; it follows that this condition amounts to imposing an upper bound on the hazard rate $h(1 - \underline{C}(\lambda)) = f(1 - \underline{C}(\lambda))/F(1 - \underline{C}(\lambda))$, namely:[18]

$$h(1 - \underline{C}(\lambda)) \leq g(\lambda) \equiv \frac{(\lambda^c - \lambda)^2}{\rho\lambda^c(\lambda^c - \lambda^{ci})\left(\lambda^i - \lambda\right)}.$$

As both sides of this inequality are decreasing in λ, it holds throughout the range $\lambda \in \left(\lambda^N, \lambda^{ci}\right)$ whenever $h\left(1 - \underline{C}\left(\lambda^N\right)\right) \leq \underline{h} \equiv g\left(\lambda^{ci}\right)$.

Similarly, the local condition

$$\frac{d\mathcal{W}_C^P}{d\lambda}\left(\lambda^{ci}\right) > \frac{d\mathcal{W}_N}{d\lambda}\left(\lambda^{ci}\right) = \frac{dP^N}{d\lambda}\left(\lambda^{ci}\right) W^i\left(\lambda^{ci}\right)$$

ensures that expected welfare is higher in the absence of commitments for λ close to λ^{ci}. Using (7) and (9), $dP^N(\lambda)/d\lambda$ can be shown to be proportional to $F(1 - \underline{C}(\lambda))$,

[18] See the proof of Proposition 6.

implying that the above condition amounts again to imposing a (weaker) upper bound on the hazard rate, namely, $h(1) < \overline{h}$.

5 Reformed Regulation

As already mentioned, the existing regulation not only introduces commitments, but also gives the initiative to the firms, which enables them to appropriate a substantial share of the resulting benefits. To disentangle the two aspects, we now consider a reformed regulation giving the initiative to the enforcer. That is, we now assume that it is the enforcer who can propose (and design) a commitment. We adjust accordingly the timing of the game as follows (interpreting $C = 0$ as not opening the case):

Timing (reformed regulation):

Stage 0**:** Nature draws the type $\theta \in \{G, B\}$ and the cost $c \in [0, 1]$, which are privately observed by the firm; the firm then decides whether to undertake the action. If the firm does not undertake the action, the game is over.

Stage 1**:** If an action is undertaken, the enforcer then decides whether to open a case. If it does not (i.e., if it offers $C = 0$), the game is over; otherwise, the enforcer offers a positive commitment $C \in (0, 1]$.

Stage 2**:** If a case is opened, the firm either accepts the commitment (A) or rejects it (R), in which case the enforcer proceeds to a full investigation.

There is a key difference between the current and reformed regulatory regimes. In the former setting, the commitment offered by the firm acts as a signalling device. The enforcer updates its beliefs accordingly, before deciding whether to accept the commitment or investigate. In the latter setting, the commitment proposed by the enforcer acts instead as a screening device: the firm then either accepts it, ending the game, or rejects it, triggering an investigation.

We now analyze the equilibria, starting with the last stage.

5.1 Stage 2

In stage 2, the firm accepts the offered commitment C whenever it lies below its maximal acceptable commitment, $\bar{C}_\theta$. The continuation equilibria can therefore be of three types: separating, pooling or semi-separating. Specifically, we have:

Lemma 3 (continuation equilibria – reformed regulation) *For any C selected by the enforcer in stage 1, the continuation equilibria and the associated expected welfare, given the enforcer's interim belief λ_1, are as follows:*

- *if $C = 0$, then the case is not opened and the expected welfare is $W^c(\lambda_1)$;*
- *if $C \in \left(0, \bar{C}_G\right)$, then there exists a unique continuation equilibrium, in which both types pool on accepting the commitment, yielding an expected welfare equal to $W^a(\lambda_1, C) = (1 - C)W^c(\lambda_1)$*
- *if $C > \bar{C}_B$, then there exists a unique continuation equilibrium, in which both types pool on rejecting the commitment, yielding $W^i(\lambda_1)$;*
- *if $C \in \left(\bar{C}_G, \bar{C}_B\right)$, then there exists a unique continuation equilibrium, in which the two types separate, with B accepting and G rejecting the commitment, yielding*

$$W^s(\lambda_1, C) \equiv \lambda_1(W - k) - (1 - \lambda_1)(1 - C)L;$$

- *if $C = \bar{C}_G$, then the pooling equilibrium on acceptance and the separating equilibrium co-exist with a continuum of semi-separating equilibria in which B accepts the commitment with probability 1 whereas G rejects it with any probability $v_G \in (0, 1)$. The pooling equilibrium on acceptance Pareto-dominates all other equilibria.*
- *if $C = \bar{C}_B$, then the pooling equilibrium on rejection and the separating equilibrium co-exist with a continuum of semi-separating equilibria in which G rejects the commitment with probability 1 whereas B rejects it with probability $v_B \in (0, 1)$. The separating equilibrium Pareto-dominates all other equilibria.*

Proof. See Appendix M. ■

5.2 Stage 1

In stage 1 the enforcer, given its interim beliefs λ_1, chooses whether to open a case by offering a positive commitment $C > 0$, triggering a continuation equilibrium described in Lemma 3 – focusing as before on the Pareto-dominant one in case of multiplicity – or not to open it at all, which is equivalent to offering a zero commitment. The associated expected welfare is given by (using $W^c(\lambda_1) = W^a(\lambda_1, 0)$):

$$W(\lambda_1, C) = \begin{cases} W^a(\lambda_1, C) & \text{if } C \in \left[0, \bar{C}_G\right] \\ W^s(\lambda_1, C) & \text{if } C \in \left(\bar{C}_G, \bar{C}_B\right] \\ W^i(\lambda_1) & \text{if } C > \bar{C}_B. \end{cases} \tag{20}$$

We first note that, $W^s(\lambda_1, C)$ being strictly increasing in C, the relevant option in the range $C \in \left(\bar{C}_G, \bar{C}_B\right]$ is $C = \bar{C}_B$. Furthermore, this option also dominates any higher C; indeed, we have:

$$W^s(\lambda_1, \bar{C}_B) - W^i(\lambda_1) = (1 - \lambda_1)[k + (sk + \rho S)L] > 0. \tag{21}$$

Both options lead G to reject the offered commitment and trigger an investigation; by contrast, $\bar{C}_B$ is accepted by B, whereas any higher commitment is rejected. Intuitively, the commitment option is more efficient, as it saves on investigation costs. As the enforcer can moreover impose the maximal commitment acceptable by B, the resulting expected welfare is higher.

It follows that the relevant options are $C \in \left[0, \bar{C}_G\right]$ and $C = \bar{C}_B$. Furthermore, in the first range, the best option is $C = \bar{C}_G$ if $\lambda_1 < \lambda^c$, and $C = 0$ if $\lambda_1 > \lambda^c$. Building on this, the following lemma establishes the optimal response of the enforcer:

Lemma 4 (enforcer's response – reformed regulation) *There exists $\hat{\lambda}_1 \in (\lambda^{ci}, \lambda^c)$ such that, when the action is undertaken by the firm, the enforcer's optimal response is as follows:*

(i) if $\lambda_1 \in (0, \hat{\lambda}_1)$, the enforcer offers $\bar{C}_B$, triggering separation;

(ii) if $\lambda_1 \in (\hat{\lambda}_1, \lambda^c)$, the enforcer offers $\bar{C}_G$, triggering acceptance;

(iii) if $\lambda_1 > \lambda^c$, the enforcer does not open a case (i.e., selects $C = 0$).

In the limit case where $\lambda_1 = \hat{\lambda}_1$, the enforcer is indifferent between $\bar{C}_B$ (triggering separation) and $\bar{C}_G$ (triggering acceptance); in the limit case where $\lambda_1 = \lambda^c$ the enforcer is indifferent between any $C \in (0, \bar{C}_G]$ (triggering acceptance) and not opening the case (i.e., selecting $C = 0$).

Proof. See Appendix N. ■

5.3 Stage 0

In stage 0 the firm, having observed its type, decides whether to undertake the action or not, given the optimal response of the enforcer described in Lemma 4.

If the enforcer's prior is sufficiently pessimistic, its interim belief will remain highly pessimistic (namely, $\lambda_1 < \hat{\lambda}_1$) even will full (deterrence and) screening; the enforcer will therefore always opt for the larger relevant commitment, $\bar{C}_B$. This, in turn, leads both types to adopt the practice as if doing so triggers an investigation with probability 1 – specifically, G will rejects the commitment and face an investigation, whereas B will accept it but obtain the same profit as under investigation. The participation of type θ is thus given by

$$\bar{P}_{\theta}^{S} \equiv F(1-\bar{C}_{\theta}),$$

and the enforcer's resulting interim belief is equal to

$$\bar{\lambda}_1^S(\lambda) \equiv \frac{\lambda \bar{P}_G^S}{\lambda \bar{P}_G^S + (1-\lambda)\bar{P}_B^S} (\geq \lambda). \tag{22}$$

By construction, $\bar{P}_G^S > \bar{P}_B^S$, implying $\bar{\lambda}_1^S(\lambda) > \lambda$ for $\lambda \in (0,1)$, and $\bar{\lambda}_1^S(\lambda)$ strictly increases from 0 to 1 as λ increases 0 to 1. The above strategies thus constitute the unique equilibrium strategies as long as $\lambda < \hat{\lambda} \equiv (\bar{\lambda}_1^S)^{-1}(\hat{\lambda}_1)$.

If instead the enforcer is sufficiently optimistic, its interim belief will remain highly optimistic (namely, $\lambda_1 > \lambda^c$) even in the absence of any screening. The enforcer will therefore not open the case (i.e., it will select $C = 0$). This, in turn, leads both types to adopt the practice with the same probability, $F(1)$, and the enforcer thus does not revise its belief. Hence, for $\lambda > \lambda^c$, there is no enforcement.

Finally, in the intermediate range $\lambda_1 \in (\hat{\lambda}, \lambda^c)$, two cases can be distinguished. For $\lambda > \hat{\lambda}_1$, even in the absence of screening the enforcer remains sufficiently optimistic for opting for the lower relevant commitment, $\bar{C}_G$ (which is accepted by both types, thus generating no screening indeed). For $\lambda < \hat{\lambda}_1$, in the absence of screening the enforcer would be too pessimistic for offering $\bar{C}_G$, but offering the higher commitment, $\bar{C}_B$, would generate full screening and make the enforcer too optimistic for offering it. As a result, in equilibrium, the enforcer randomizes between the two commitments – and B participates in such a way that the enforcer's interim belief is equal to $\hat{\lambda}_1$, making the enforcer indeed indifferent between the two commitments.

Building on this, the following proposition describes the equilibria for any given λ.

Proposition 7 (reformed regulation) *The equilibrium $\mathcal{E}_R$ is as follows:*

- *For $\lambda \leq \hat{\lambda}$, where $\hat{\lambda}(< \hat{\lambda}_1)$ is such that $\bar{\lambda}_1^S(\hat{\lambda}) = \hat{\lambda}_1$, there exists a unique equilibrium, in which type $\theta = G, B$ undertakes the action with probability $\bar{P}_{\theta}^S$, the enforcer updates the beliefs to $\bar{\lambda}_1^S(\lambda)(\leq \hat{\lambda}_1)$ and offers $\bar{C}_B$ (which G rejects and B accepts).*

- *For $\lambda \in (\hat{\lambda}, \hat{\lambda}_1)$, there exists a unique equilibrium, in which (i) G undertakes the action with probability $\bar{P}_G^S$ and B undertakes it with probability*

$$\bar{P}_B^M(\lambda) = F\left(1 - \hat{\sigma}(\lambda)\bar{C}_G - (1-\hat{\sigma}(\lambda))\bar{C}_B\right),$$

where the probability $\hat{\sigma}(\lambda)$ increases with λ from $\hat{\sigma}(\hat{\lambda}) = 0$ to $\hat{\sigma}(\hat{\lambda}_1) = 1$, and (ii) the enforcer updates its belief to $\lambda_1 = \hat{\lambda}_1$ and offers $\bar{C}_G$ (which both types accept) with *probability $\hat{\sigma}(\lambda) \in (0,1)$, and $\bar{C}_B$ (which G rejects and B accepts) with complementary probability $1 - \hat{\sigma}(\lambda)$.*

- *For $\lambda \in [\hat{\lambda}_1, \lambda^c)$, there exists an equilibrium in which both types undertake the action with probability $\bar{P}_G^S$, the enforcer maintains its beliefs (i.e., $\lambda_1 = \lambda$) and offers $\bar{C}_G$ (which both types accept). This equilibrium is the only Pareto-dominant equilibrium if $\lambda = \hat{\lambda}_1$ and is the unique equilibrium for $\lambda \in (\hat{\lambda}_1, \lambda^c)$.*

- *For $\lambda \geq \lambda^c$, there exists an equilibrium in which both types undertake the action with probability $F(1)$, the enforcer maintains its beliefs (i.e., $\lambda_1 = \lambda$) and does not open the case. This equilibrium is the only Pareto-dominant equilibrium if $\lambda = \lambda^c$ and is the unique equilibrium for $\lambda > \lambda^c$.*

Proof. See Appendix O. ■

The following Corollary characterizes the expected welfare generated by the reformed regulation:

Corollary 5 (expected welfare – reformed regulation) *Expected welfare is given by:*

$$\mathcal{W}_R(\lambda) = \begin{cases} \mathcal{W}_R^S(\lambda) = \bar{P}^S(\lambda) W^S(\lambda) & \text{for } \lambda \leq \hat{\lambda}, \\ \mathcal{W}_R^M(\lambda) = \bar{P}^M(\lambda) W^M\left(\hat{\lambda}_1\right) & \text{for } \lambda \in (\hat{\lambda}, \hat{\lambda}_1), \\ \mathcal{W}_R^A(\lambda) = \bar{P}^A(\lambda) W^A(\lambda) & \text{for } \lambda \in [\hat{\lambda}_1, \lambda^c), \end{cases} \tag{23}$$

where total participation is equal to:

$$\bar{P}^S(\lambda) = \lambda \bar{P}_G^S + (1-\lambda)\bar{P}_B^S,$$
$$\bar{P}^M(\lambda) = \lambda \bar{P}_G^S + (1-\lambda)\bar{P}_B^M(\lambda),$$
$$\bar{P}^A(\lambda) = \bar{P}_G^S,$$

and expected welfare for a given participation is:

$$W^S(\lambda) = W^s(\bar{\lambda}_1^S(\lambda), \bar{C}_B),$$
$$W^M\left(\hat{\lambda}_1\right) = W^s(\hat{\lambda}_1, \bar{C}_B) = W^a(\hat{\lambda}_1, \bar{C}_G),$$
$$W^A(\lambda) = W^a(\lambda, \bar{C}_G).$$

5.4 Comparison of policy regimes

We now show that the reformed regulation regime (R) dominates both the benchmark regime (N) and the current regulation (C):

Proposition 8 (reformed regulation dominates) *For any $\lambda < \lambda^c$, the reformed regulatory regime yields a higher expected welfare than the benchmark and current regulatory regimes.*

Proof. See Appendix P. ■

From Proposition 5, among the three relevant equilibria generated by the current regulatory regime, the pooling equilibrium $\mathcal{E}_N$ replicates the outcome of the benchmark regime and the semi-equilibrium $\mathcal{E}_C^S$ is dominated by that outcome. Hence, we only need to compare the reformed regime to the benchmark regime and to the pooling equilibrium $\mathcal{E}_C^P$ of the current regime, in which the firm offers the minimum acceptable commitment $\underline{C}(\lambda)$.

It will be useful to distinguish four cases based on how pessimistic the enforcer's prior beliefs are. [Recall that, from Lemma 4, $\hat{\lambda}_1 > \lambda^{ci}$, which in turn implies $\hat{\lambda} = (\bar{\lambda}_1^S)^{-1}(\hat{\lambda}_1) > \lambda^N = (\bar{\lambda}_1^S)^{-1}(\lambda^{ci})$.[19] Hence, the four ranges identified below are all non-empty.]

- Case a: $\lambda \in \left(0, \lambda^N\right]$. In this range, the benchmark regime dominates the current regime and induces the enforcer to investigate with probability 1. Furthermore, as $\lambda \leq \lambda^N < \hat{\lambda}$, the reformed regime gives rise to the separating equilibrium in which G faces an investigation wih probability 1, whereas B accepts $\bar{C}_B$. It follows that, compared with the benchmark regime, the reformed regime (i) performs equally well in terms of deterrence and screening (in both cases, participation rates are the same as when investigating with probability 1) and (ii) is also equally ex-post effective when facing G (investigation with probability 1 in both regimes). However, when facing B, the reformed regime is more ex-post effective, as it replaces the investigation with the commitment $\bar{C}_B$; doing so enables the enforcer to save on the investigation cost k and to appropriate B's cost savings (which include B's share of the investigation cost, sk, as well as the expected pecuniary sanction ρS) – conditional on facing a bad type, this amounts to increasing *de facto* the commitment by $\bar{C}_B - \hat{C} = k/L + sk + \rho S > 0$.

- Case b: $\lambda \in (\lambda^N, \hat{\lambda}]$. In this range, the reformed regime still gives rise to the separating equilibrium and thus dominates the other two regimes in terms of deterrence and screening: it maintains full deterrence and screening (as when investigating with probability 1), whereas (i) the investigation probability falls down in the benchmark regime, which reduces both deterrence and screening, and (ii) the pooling equilibrium $\mathcal{E}_C^P$ of the current regime leads to the commitment $\underline{C}(\lambda)$, implying no screening and lower deterrence, as $\underline{C}(\lambda) \leq \bar{C}_G < \bar{C}_B$. In addition, the reformed regime is again more ex-post effective. Indeed, conditional on the enforcer's interim beliefs, the benchmark and current regimes deliver the same expected welfare as an investigation, $W^i(\cdot)$;[20] by contrast, as in the previous case, the reformed regulation provides ex-post an equally effective policy when facing G and a more effective one when facing B.

[19] From Proposition 7, $\hat{\lambda}$ is such that $\bar{\lambda}_1^S(\hat{\lambda}) = \hat{\lambda}_1$. And from (6), λ^N is such that $\lambda_1^e(\lambda^N, 1) = \lambda^{ci}$, where by construction, $\lambda_1^e(\cdot, 1) = \bar{\lambda}_1^S(\cdot)$, which is strictly increasing (see (2) and (22)).

[20] Specifically, $\mathcal{E}_N$ yields $\lambda_1^N(\lambda) = \lambda^{ci}$ and $W_1^N(\lambda^{ci}) = W^i(\lambda^{ci})$, whereas $\mathcal{E}_C^P$ yields $\lambda_1^C(\lambda) = \lambda$ and $W_1^C(\lambda) = W^i(\lambda)$.

• Case c: $\lambda \in (\hat{\lambda}, \lambda^{ci})$. In this range, the reformed regime (which yields the equilibrium M) dominates again the other two regimes in terms of participation, screening and policy effectiveness. To see this, recall that the reformed regulation generates participation rates based on commitments equal to $\bar{C}_G$ for G, and to either $\bar{C}_G$ or $\bar{C}_B$ for B. By contrast, under the current regulation, the relevant pooling equilibrium – when it exists – generates participation rates based on $\underline{C}(\lambda) \leq \bar{C}_G(< \bar{C}_B)$, and is thus dominated by the reformed regime in terms of deterrence and screening. Furthermore, the reformed regulation yields a more optimistic interim belief than the benchmark regime (namely, $\hat{\lambda}_1 > \lambda^{ci}$); hence, it performs better in terms of screening. Moreover, G's participation is based on facing an investigation with probability strictly lower than 1 in the benchmark regime, and on facing an investigation with certainty in the reformed regime. Hence, G's participation is lower under the reformed regulation. As there is more screening, B's participation must be even lower, implying that the reformed regime thus performs better than the benchmark regime in terms of deterrence as well as screening. Finally, the reformed regime dominates the others also in terms of ex-post effectiveness, as the enforcer is indifferent between investigating or not in the other two regimes, and could but prefers not to do so in the reformed regime.

• Case d: $\lambda \in \left[\lambda^{ci}, \lambda^{c}\right)$. In this range, the benchmark regime and the current regulation entail no enforcement, whereas the reformed regulation (leading to equilibrium M, or to pooling on acceptance) achieves some enforcement, in a range of λ in which enforcement is moreover strictly desirable, since $W^{c}(\lambda) < 0$. It follows that the reformed regime performs at least as well as the other two in terms of screening (and strictly better in the range $\lambda \in \left[\lambda^{ci}, \hat{\lambda}_1\right)$ in which it generates the equilibrium M), and performs strictly better than the other two in terms of deterrence ($P_R(\lambda) < P_N(\lambda) = P_C(\lambda)$) and ex-post effectiveness (the benchmark and current regulatory regimes yield $\lambda_1 = \lambda$ and $W_1^N(\lambda) = W_1^C(\lambda) = W^c(\lambda)$, whereas the reformed regime yields $W_1^R\left(\lambda_1^R(\lambda)\right) > W^c\left(\lambda_1^R(\lambda)\right)$, where $\lambda_1^R(\lambda)$ are the enforcer's interim belief under the reformed regulation).

Summing-up, the reformed regulatory regime does at least as well as the other two regimes in all dimensions (deterrence, screening and ex-post effectiveness), and always strictly outperforms the other two regimes in at least one dimension. Specifically, the reformed regulatory regime:

- always strictly outperforms the other two regimes in terms of ex-post effectiveness;
- weakly dominates the other two regimes in terms of screening – it moreover strictly outperforms the current regulation for $\lambda \in \left(0, \lambda^N\right]$, and both other regimes for $\lambda \in (\lambda^N, \hat{\lambda}_1)$;
- weakly dominates the other two regimes in terms of deterrence – it moreover strictly outperforms the current regulation for $\lambda \in \left(0, \lambda^N\right]$, and both other regimes for

$\lambda \in (\lambda^N, \lambda^c)$.

6 Conclusion

Since the reform in the enforcement of art. 101 and 102 TFEU in 2003, known as "the modernization", the commitment procedure has become widely used in European antitrust cases. This instrument enables a firm under investigation to offer measures intended to limit the anticompetitive effects. If accepted by the enforcer, these remedies become binding commitments, but there is no fine or finding of infringement.

To study the impact of this reform, we consider a setting in which a firm has the opportunity to undertake a practice that may (exogenously) be pro- or anti-competitive. The firm knows the nature of the practice, whereas the enforcer only has prior beliefs about it. We first consider the equilibrium outcomes generated by two regimes: a benchmark regime in which the enforcer can only rely on investigations, which are costly but bring evidence with some probability; and a regime reflecting the current regulation, in which the firm can propose a commitment that, if accepted, reduces both social effects and private profits.

We show that the success of a policy regime relies on its performance over three dimensions: deterrence (discouraging the adoption of the practice), screening (selectively targeting bad actions), and ex-post effectiveness (improving ex-post value and/or reducing enforcement costs). We find that, compared with the benchmark regime, the current regulation never improves ex-post effectiveness, and reduces screening (at least weakly). Furthermore, while it may sometimes improve deterrence, it appears unlikely to generate a higher expected welfare.

The poor performance of the current regulation led us to explore a reform giving the enforcer the initiative to propose a commitment. We find that the reformed regulation dominates the benchmark regime and the current regulation (and actually widens the scope of enforcement): it performs weakly better over all three dimensions (deterrence, screening and ex-post effectiveness), and strictly so over at least one dimension.

The debate surrounding the introduction of the commitments procedure centers on a trade-off between ex-post effectiveness, on the one hand, and deterrence or screening, on the other hand. The above insights suggest that the current regulation performs poorly on the latter front, without providing much gains on the former. They also suggest that the proposed reform could do better on both fronts.

The underlying intuition relies on a key feature, namely (following the literature and reflecting enforcement practice), that the regulator is not in a position to pre-commit itself to a given investigation policy; it reacts instead to the action when it is undertaken – based on its beliefs, appropriately updated for the occurrence of the action. By giving the firm control over commitment design, the current regulation enables it to offer a minimally acceptable commitment, appropriating in this way the lion's share of the gains stemming from this new tool. As a result, ex-post effectiveness is not much improved, whereas deterrence and screening are much hampered.

By contrast, by giving the regulator control over commitment design, the proposed reform could enable the enforcer to appropriate a larger share of these gains. This would not only improve ex-post effectiveness, but, by so doing, it would also foster ex-post intervention (and, indeed, widens the scope for enforcement), which in turn would enhance ex-ante deterrence and screening.

Our analysis assumes that the enforcer and the firm have symmetric information about relevant commitments. In practice, the firm may have private information about feasible remedies, or about their costs or benefits. The enforcer may however elicit such information from other interested industry participants, such as rivals or customers (industrial customers in the case of intermediate goods, or possibly consumer associations in the case of final goods). Taking this informational problem into account, and studying the impact of control over commitment design in such environment, constitutes an interesting avenue for further research.

References

References

[1] Bebchuk L. (1984), Litigation and Settlement under Imperfect Information, *Rand Journal of Economics*, 15: 404-15.

[2] Choné P., Souam S. and Vialfont A. (2014), On the Optimal Use of Commitment Decisions under European Competition Law, *International Review of Law and Economics,* 37: 169-79

[3] Daughety A., Reinganum J. (2011), Settlement, in Sanchirico C. (ed), *Encyclopedia of Law and Economics*, vol. 7, Edward Elgar Pub.

[4] Gautier A., Petit N. (2018), Optimal Enforcement of Competition Policy: The Commitment Procedure under Uncertainty, *European Journal of Law and Economics,* 45: 195-224.

[5] Mariniello M. (2014), Commitment or Prohibition? The EU Antitrust Dilemma., *Bruegel Policy Brief*, issue 1/2014.

[6] Polinsky M., Rubinfeld D. (1988), The Deterrent Effect of Settlements and Trials, *International Review of Law and Economics,* 8: 109-116.

[7] Reinganum J., Wilde L. (1986), Settlement, Litigation and the Allocation of Litigation Cost, *Rand Journal of Economics*, 17: 557-66.

[8] Shavell S. (1982), Suit, Settlement and Trial: a Theoretical Analysis under Alternative Methods for the Allocation of Legal Costs, *Journal of Legal Studies,* 11: 55-81.

[9] Wils W. (2006), Settlement of EU Antitrust Investigations: Commitment Decisions under Art. 9 of Regulation n. 1/2003, *World Competition*, 29: 345-366.

[10] Wils W. (2008), The Use of Settlements in Public Antitrust Enforcement: Objectives and Principles, *World Competition*, 319: 335-352.

Appendix

A Proof of Proposition 1

Suppose that, in equilibrium, undertaking the action induces the enforcer to investigate with probability ι. Anticipating this, the firm undertakes an action of type θ with probability $P_\theta(\iota)$ given by (1); the enforcer's updated beliefs, given by (2), can be written as

$$\lambda_1^e(\lambda,\iota) \equiv \frac{1}{1+\frac{1-\lambda}{\lambda}\Phi(\iota)}$$

where

$$\Phi(\iota) \equiv \frac{P_B(\iota)}{P_G(\iota)}(>0)$$

satisfies $\Phi(0) = 1$ and is strictly decreasing in ι:

$$\Phi'(\iota) = -\Phi(\iota)\left[\bar{C}_B h_B(\iota) - \bar{C}_G h_G(\iota)\right] < 0, \tag{24}$$

where

$$h_\theta(\iota) \equiv h(1-\iota\bar{C}_\theta),$$

and the inequality stems from the hazard rate $h(\cdot) = f(\cdot)/F(\cdot)$ being strictly decreasing and $\bar{C}_B > \bar{C}_G$. It follows that $\lambda_1^e(\lambda,\iota)$ strictly increases with ι as well as with λ, and moreover satisfies $\lambda_1^e(\lambda,0) = \lambda$ and $\lambda_1^e(\lambda,1) = \lambda^{ci}$ for $\lambda = \lambda^N(<\lambda^{ci})$, defined in (6). Therefore, there are three possible cases:

- If $\lambda \geq \lambda^{ci}$, then any $\iota > 0$ would induce $\lambda_1^e(\lambda,\iota) > \lambda^{ci}$, implying that the enforcer would be unwilling to investigate; hence, the enforcer closes the case (or never opens it): $\iota(\lambda) = 0$.

- If instead $\lambda \leq \lambda^N$, then any $\iota < 1$ would induce $\lambda_1^e(\lambda,\iota) < \lambda^{ci}$, implying that the enforcer would be unwilling to close the case; hence, the enforcer investigates whenever the firm undertakes the action: $\iota(\lambda) = 1$.

- Finally, if $\lambda^N < \lambda < \lambda^{ci}$, then the enforcer must be indifferent between closing the case or proceeding with an investigation: $\iota = 0$ would lead to $\lambda_1^e(\lambda,0) < \lambda^{ci}$, in which case the enforcer would rather investigate ($\iota = 1$), a contradiction; likewise, $\iota = 1$ would lead to $\lambda_1^e(\lambda,\iota) > \lambda^{ci}$, in which case the enforcer would rather close the case ($\iota = 0$), another contradiction. It follows that the enforcer's posterior must satisfy $\lambda_1^e(\lambda,\iota) = \lambda^{ci}$, which in turn requires the enforcer to investigate with probability $\tilde{\iota}(\lambda) \in (0,1)$, where $\tilde{\iota}(\lambda)$ is the unique solution in ι to $\lambda_1^e(\lambda,\iota) = \lambda^{ci}$.

and is implicitly defined by

$$\Phi(\tilde{\iota}(\lambda)) = \frac{\lambda}{1-\lambda}\frac{1-\lambda^{ci}}{\lambda^{ci}}. \tag{25}$$

The LHS of (25) is continuously differentiable in ι, whereas the RHS is continuously differentiable in λ. It follows that $\tilde{\iota}(\lambda)$ is continuously differentiable in λ. Using (24) and (25) yields:

$$\tilde{\iota}'(\lambda) = \frac{-1}{\lambda(1-\lambda)\left[\bar{C}_B h_B(\tilde{\iota}(\lambda)) - \bar{C}_G h_G(\tilde{\iota}(\lambda))\right]} < 0, \tag{26}$$

where the inequality follows from the monotonicity of $h(\cdot)$ and $\bar{C}_B > \bar{C}_G$. Hence, $\tilde{\iota}(\lambda) \in [0,1]$ is decreasing in λ for $\lambda \in \left[\lambda^N, \lambda^{ci}\right]$.

Finally, the participation of the good and bad types in the no-commitment regime (superscript "N") can be expressed as

$$P_\theta^N(\lambda) \equiv F(1-\iota^N(\lambda)\bar{C}_\theta).$$

B Proof of Corollary 1

The properties of $P_G^N(\lambda)$ and $P_B^N(\lambda)$ directly follow from $\bar{C}_G < \bar{C}_B$ and the properties of $\iota^N(\lambda)$. Moreover, for $\lambda \in (0, \lambda^N]$:

$$\frac{dP^N(\lambda)}{d\lambda} = P_G^N(\lambda) - P_B^N(\lambda) > 0,$$

and for $\lambda \in (\lambda^N, \lambda^{ci})$, using (26) and $p_\theta^N(\lambda) \equiv f\left(1-\iota^N(\lambda)\bar{C}_\theta\right)$:

$$\begin{aligned}\frac{dP^N(\lambda)}{d\lambda} &= P_G^N(\lambda) - P_B^N(\lambda) + \lambda\frac{dP_G^N(\lambda)}{d\lambda} + (1-\lambda)\frac{dP_B^N(\lambda)}{d\lambda} \\ &= P_G^N(\lambda) - P_B^N(\lambda) + \frac{\lambda\bar{C}_G p_G^N(\lambda) + (1-\lambda)\bar{C}_B p_B^N(\lambda)}{\lambda(1-\lambda)\left[\bar{C}_B h_B^N(\lambda) - \bar{C}_G h_G^N(\lambda)\right]} > 0. \end{aligned} \tag{27}$$

C Proof of Corollary 2

The posterior of λ is

$$\lambda_1^N(\lambda) = \lambda\frac{P_G^N(\lambda)}{P^N(\lambda)} > \lambda.$$

Expected welfare is given by:

$$\mathcal{W}_N(\lambda) = P^N(\lambda)\left\{[1-\iota(\lambda)]\,W^c(\lambda_1^N(\lambda)) + \iota(\lambda)W^i(\lambda_1^N(\lambda))\right\},$$

where from Proposition 1: for $\lambda \leq \lambda^N$, $\iota^N(\lambda) = 1$; and for $\lambda \in (\lambda^N, \lambda^{ci})$, $\lambda_1^N(\lambda) = \lambda^{ci}$, implying that the enforcer is indifferent between investigating or not, and so $W^c(\lambda_1^N(\lambda)) = W^i(\lambda_1^N(\lambda))\ (= W^i(\lambda^{ci}))$. Hence, in both cases, $\mathcal{W}_N(\lambda) = P^N(\lambda)W^i(\lambda_1^N(\lambda))$.

Moreover, for $\lambda \in (0, \lambda^N)$ the enforcer investigates with probability 1 and we have

$$\frac{d\mathcal{W}_N(\lambda)}{d\lambda} = \frac{dP^N(\lambda)}{d\lambda} W^i\left(\lambda_1^N(\lambda)\right) + P^N(\lambda)\frac{dW^i(\lambda_1)}{d\lambda_1}\frac{d\lambda_1^N(\lambda)}{d\lambda} = P_G^N W + P_B^N(1-\rho)L > 0$$

since $\frac{dP^N(\lambda)}{d\lambda} = P_G^N - P_B^N$, $\frac{dW^i(\lambda_1)}{d\lambda_1} = W + (1-\rho)L$ and $\frac{d\lambda_1^N(\lambda)}{d\lambda} = \frac{P_G^N P_B^N}{(P^N)^2}$.

For $\lambda \in (\lambda^N, \lambda^{ci})$, where $\mathcal{W}_N(\lambda) = P^N(\lambda)W^i(\lambda^{ci})$, we have:

$$\frac{d\mathcal{W}_N(\lambda)}{d\lambda} = \frac{dP^N(\lambda)}{d\lambda}W^i(\lambda^{ci}) < 0,$$

as P^N is strictly increasing in λ from Corollary 1, and $W^i(\lambda^{ci}) < 0$.

D Proof of Lemma 1

Fix C and let $\lambda_2 = \lambda_2(C)$ denote the enforcer's revised belief when offered C. Accepting C thus generates an expected welfare equal to $W^a(\lambda_2, C) = (1-C)W^c(\lambda_2)$, whereas running an investigation yields $W^i(\lambda_2)$. Let

$$\Delta(\lambda_2, C) \equiv W^a(\lambda_2, C) - W^i(\lambda_2)$$

denote the welfare differential between the two options. Using (3) and (5), it can be expressed as:

$$\begin{aligned}
\Delta(\lambda_2, C) &= -C(W+L)(\lambda_2 - \lambda^c) - \rho L\left(\lambda^{ci} - \lambda_2\right) \\
&= (W+L)\left[(\rho\lambda^c - C)\lambda_2 - \lambda^c\left(\rho\lambda^{ci} - C\right)\right] \\
&= (W+L)(\rho\lambda^c - C)\left[\lambda_2 - \underline{\lambda}(C)\right],
\end{aligned}$$

where the last equality stems from (12). Therefore:

- If $C < \rho\lambda^c$, $\Delta(\lambda_2, C) \gtreqless 0$ if and only if $\lambda_2 \gtreqless \underline{\lambda}(C)$.
- If instead $C \geq \rho\lambda^c$, $\Delta(\lambda_2, C)$ is (weakly) decreasing in λ_2; hence,

$$\Delta(\lambda_2, C) \geq \Delta(1, C) = (1-C)W - (W-k) = k - CW \geq k - W \geq 0,$$

  where the second inequality stems from $C \leq 1$ and the last one from Assumption 3. Hence, the enforcer accepts the commitment C for any $\lambda_2(C)$.

E Proof of Proposition 2

Part (i). Suppose that $\lambda \geq \lambda^{ci}$, and consider a candidate equilibrium in which both types undertake the action with probability 1 in stage 0, and offer no commitment in stage 1. We then have $\lambda_2 = \lambda_1 = \lambda \geq \lambda^{ci}$, implying that, in stage 2, the enforcer is willing to close the case. This, in turn, gives both types the maximal profit of 1, implying that they have no incentive to deviate in stages 0 and 1. This establishes the existence of an equilibrium in which both types obtain a payoff of 1, which therefore Pareto-dominates any other equilibrium outcome.

Part (ii). Consider now the case $\lambda < \lambda^{ci}$ and consider a candidate equilibrium in which $\Pi_\theta = 1$ for some type $\theta \in \{B, G\}$. To obtain this profit, type θ must offer zero commitment, and doing so must induce the enforcer to close the case. As the other type could mimic θ, it follows that, in this candidate equilibrium, both types obtain the maximal profit of 1. Hence, in stage 0 they both undertake the action with probability 1, implying that $\lambda_1 = \lambda$. As $E[\lambda_2(C)]_{C \in \mathcal{C}_G \cup \mathcal{C}_B} = \lambda_1 < \lambda^{ci}$, there exists $C \in \mathcal{C}_G \cup \mathcal{C}_B$ such that $\lambda_2 (C) \leq \lambda_1 < \lambda^{ci}$. From Lemma 1, the enforcer then either investigates or accepts C, and the latter case arises only if $C \geq \underline{C} (\lambda_2 (C)) > \underline{C} \left(\lambda^{ci}\right) = 0$. It follows that the type offering C obtains a profit strictly lower than 1, a contradiction.

Therefore, in any equilibrium, both types must obtain a profit strictly lower than 1. Furthermore, as G could mimic B, and G's ex-post payoffs are always weakly higher than B's, in equilibrium G must obtain a weakly higher expected profit than B.

F Proof of Lemma 2

Fix $\lambda < \lambda^{ci}$ and suppose that there exists $C_G \in \mathcal{C}_G \setminus \mathcal{C}_B$. Upon observing C_G, the enforcer updates its belief to $\lambda_2(C_G) = 1$, and thus accepts C_G. Hence, in equilibrium G obtains $\Pi_G = \Pi_G(C_G) = 1 - C_G$. Furthermore, by offering C_G, B could secure $\Pi_B(C_G) = 1 - C_G = \Pi_G$; we thus have $\Pi_B \geq (\Pi_B(C_G) =) \, \Pi_G \geq \Pi_B$, where the last inequality stems from Proposition 2. It follows that $\Pi_B = \Pi_G$. In addition, G can secure a payoff of $1 - \bar{C}_G$ by offering $\bar{C}_G$; indeed, it would then obtain $1 - \bar{C}_G$ regardless of whether the enforcer accepts this commitment or runs an investigation. Hence, $\Pi_B = (\Pi_G =) \, 1 - C_G \geq 1 - \bar{C}_G$.

Consider now a commitment $C_B \in \mathcal{C}_B$. If C_B is rejected with probability 1 by the enforcer, then B obtains $1 - \bar{C}_B < \Pi_B$, a contradiction. Hence, C_B must be accepted with positive probability. Furthermore, if it were accepted with probability 1, we should have $C_B = C_G$, otherwise the type with the higher offered commitment would have an incentive to deviate and mimic the other types. Hence, it must be the case that the enforcer randomizes between acceptance and investigation, which according to the regulation is possible only for $C_B = 0$. To maintain $\Pi_B = 1 - C_G$, it must be the case that the enforcer investigates with probability $\iota = C_G / \bar{C}_B$; but this gives G an incentive to deviate and mimic B (and obtain in this way $1 - \iota\bar{C}_G = 1 - C_G\bar{C}_G/\bar{C}_B > 1 - C_G = \Pi_G$), a contradiction. Therefore, in equilibrium, either $\mathcal{C}_G = \mathcal{C}_B$ or $\mathcal{C}_G \subset \mathcal{C}_B$.

G Proof of Proposition 3

The proof is structured in four steps. We first characterize a unique candidate equilibrium by establishing that $\mathcal{C}_B \setminus \mathcal{C}_G = \left\{\hat{C}\right\}$ (step 1), $\mathcal{C}_G = \{0\}$ (step 2) and $\iota(0) = \hat{C}/\bar{C}_B$ (step 3). We then establish existence for $\lambda < \lambda^S$ (step 4).

• **Step 1: $\mathcal{C}_B \backslash \mathcal{C}_G = \left\{\hat{C}\right\}$.** According to Lemma 2, there exists $C_B \in \mathcal{C}_B \setminus \mathcal{C}_G$; as B reveals itself by offering such commitment, it follows that $\lambda_2(C_B) = 0$. To identify the possible values for C_B, recall that B can secure $\Pi_B^S \equiv 1 - \hat{C}$ by offering $C_B = \hat{C}$, which the enforcer always accepts, regardless of its revised beliefs. Furthermore:

- If $C_B > \hat{C}$, then the enforcer accepts C_B and $\Pi_B(C_B) = 1 - C_B < \Pi_B^S$, a contradiction.
- Likewise, if $C_B < \hat{C}$ then the enforcer rejects C_B (as $\lambda_2(C_B) = 0$) and $\Pi_B(C_B) = 1 - \bar{C}_B < \Pi_B^S$, another contradiction.

Hence, a semi-separating equilibrium (superscript "S") satisfies $\mathcal{C}_B \setminus \mathcal{C}_G = \{\hat{C}\}$. Furthermore, as the enforcer accepts $\hat{C}$ whenever offered, B's payoff is equal to Π_B^S.

• **Step 2: $\mathcal{C}_G = \{0\}$.** Offering any $C_G \in \mathcal{C}_G (= \mathcal{C}_G \cap \mathcal{C}_B)$ must give B its equilibrium payoff Π_B^S. Hence, this cannot induce the enforcer to accept it with probability 1 (B's indifference would require $C_G = \hat{C}$) or to investigate with probability 1 (B would obtain $1 - \bar{C}_B < \Pi_B^S$). It follows that the enforcer must be randomizing, which, according to the policy rule, is feasible only if $C_G = 0$. Furthermore, we must have $\lambda_2(0) = \lambda^{ci}$ to induce the enforcer to randomize.

• **Step 3: $\iota(0) = \iota^S \equiv \hat{C}/\bar{C}_B$.** B's indifference condition then requires $\Pi_B(C_G) = 1 - \iota(C_G)\bar{C}_B = 1 - \hat{C}$ or:

$$\iota(C_G) = \frac{\hat{C}}{\bar{C}_B} = \iota^S.$$

• **Step 4: existence.** Thus, in equilibrium, $\mathcal{C}_G = \{0\}$ and $\mathcal{C} = \mathcal{C}_B = \{0, \hat{C}\}$; furthermore, (i) the enforcer accepts $\hat{C}$ and investigates with probability ι^S when offered no commitment, and (ii) the equilibrium payoffs satisfy

$$\Pi_B^S = 1 - \hat{C} < \Pi_G^S = 1 - \iota^S\bar{C}_G = 1 - \frac{\bar{C}_G}{\bar{C}_B}\hat{C}.$$

By construction, B is indifferent between offering 0 or $\hat{C}$, whereas G strictly prefers offering 0 (as offering $\hat{C}$ would yield a payoff of $\Pi_B^S < \Pi_G^S$). Furthermore, any deviation to $C \notin \mathcal{C} = \{0, \hat{C}\}$ can easily be deterred by interpreting it as signalling a bad type.[21] Any $C \in (0, \hat{C}$ is then rejected, giving type θ a payoff $1 - \widetilde{C} < \Pi_B^S < \Pi_G^S$, unless it exceeds $\hat{C}$, in which case it is accepted but gives both types $1 - \bar{C}_\theta < 1 - \iota^S \bar{C}_\theta = \Pi_\theta^S$.

Turning to stage 0, type θ undertakes the action with probability $P_\theta^S = F(\Pi_\theta^S)$. Hence:

$$\lambda_1(\lambda) = \frac{\lambda P_G^S}{\lambda P_G^S + (1-\lambda) P_B^S}.$$

To induce $\lambda_2(0) = \lambda^{ci}$, the probability ν_B with which the bad type offers no commitment must be large enough. Specifically, using

$$\lambda_2(0) = \frac{\lambda_1(\lambda)}{\lambda_1(\lambda) + [1 - \lambda_1(\lambda)]\nu_B},$$

we must have:

$$\nu_B = \frac{\lambda}{1-\lambda} \frac{1-\lambda^{ci}}{\lambda^{ci}} \frac{P_G^S}{P_B^S},$$

where the right-hand side is always positive for $\lambda > 0$, but is lower than 1 only if $\lambda < \lambda^S$, where λ^S is defined in (17). Hence, the semi-separating equilibrium $\mathcal{E}_C^S$ exists only for $\lambda \in (0, \lambda^S)$. Conversely, for any $\lambda \in (0, \lambda^S)$, the equilibrium strategies described above do constitute a semi-separating equilibrium.

H Proof of Corollary 3

For $\lambda \in \left[0, \lambda^S\right)$, the semi-separating equilibrium yields participation rates given by (15), which satisfy $P_G^S > P_B^S$ and do not depend on the prior λ. Furthermore, G offers no commitment with probability 1, whereas B does so with probability ν_B, designed to induce a posterior equal to λ^{ci} (to leave the enforcer indifferent between investigating or not), and otherwise offers $\hat{C}$, inducing a posterior equal to 0. Letting

$$P_0^S(\lambda) \equiv \lambda P_G^S + (1-\lambda)\nu_B P_B^S$$

denote the overall probability of no commitment being offered, the enforcer's interim belief, upon observing that the action has been undertaken, thus satisfies

$$\lambda_1^S(\lambda) = \frac{P_0^S(\lambda)}{P^S(\lambda)} \times \lambda^{ci} + \left[1 - \frac{P_0^S(\lambda)}{P^S(\lambda)}\right] \times 0 = \frac{P_0^S(\lambda)}{P^S(\lambda)} \lambda^{ci} \in \left(0, \lambda^{ci}\right).$$

[21] The belief $\lambda_2(C) = 0$ for $C \notin \mathcal{C}$ is moreover "reasonable" (e.g., consistent with the intuitive criterion), as B would indeed have more to gain if the deviant offer were instead accepted.

Furthermore, expected welfare can be expressed as:

$$\begin{aligned}
W^S(\lambda) &= P^S(\lambda)\left\{\frac{P_0^S(\lambda)}{P^S(\lambda)}W^i\left(\lambda^{ci}\right)+\left[1-\frac{P_0^S(\lambda)}{P^S(\lambda)}\right]W^i(0)\right\} \\
&= P^S(\lambda)\,W^i\left(\frac{P_0^S(\lambda)}{P^S(\lambda)}\lambda^{ci}\right) \\
&= P^S(\lambda)\,W^i\left(\lambda_1^S(\lambda)\right),
\end{aligned}$$

where the second equality relies on the linearity of $W^i(\lambda_1)$ in λ_1 (namely, $W^i(\lambda_1) = W^i(0) + \lambda_1[W + (1-\rho)L]$).

I Proof of Proposition 4

Fix $\lambda < \lambda^{ci}$ and suppose there that exists a pooling equilibrium; that is, $\mathcal{C}_G = \mathcal{C}_B = \mathcal{C}$. Let λ_1 denote the enforcer's interim belief, upon observing that the action has been undertaken. We consider three cases, depending on the value of this belief.

Case 1: $\lambda_1 > \lambda^{ci}$. By construction, there then exists $C \in \mathcal{C}$ such that $\lambda_2(C) \geq \lambda_1 > \lambda^{ci}$. From Lemma 1, the enforcer then accepts C, implying that both types obtain the same payoff. As the firm must be indifferent between all equilibrium offers, it follows that both types always obtain the same expected payoff, and thus undertake the action with the same probability; hence, $\lambda^{ci} < \lambda_1 = \lambda$, contradicting the working condition $\lambda < \lambda^{ci}$.

Case 2: $\lambda_1 = \lambda^{ci}$. By construction, there then exists $C \in \mathcal{C}$ such that $\lambda_2(C) \geq \lambda_1 (= \lambda^{ci})$. If $C > 0$, then from Lemma 1 the enforcer accepts it; hence, both types obtain again the same payoff, implying $\lambda = \lambda_1 \geq \lambda^{ci}$, a contradiction. Hence, $C = 0$. Likewise, if the enforcer then closes the case with probability 1, again both types obtain the same payoff, implying $\lambda = \lambda_1 = \lambda^{ci}$, a contradiction. Hence, the enforcer must open an investigation with positive probability: $\iota(0) > 0$, implying that B obtains $1 - \iota(0)\bar{C}_B$. As B could secure a payoff of $1 - \hat{C}$ by offering the commitment $\hat{C}$, the probability of investigation cannot be too large, namely:

$$\iota(0) \leq \frac{\hat{C}}{\bar{C}_B} = \iota^S.$$

Suppose now that there exists another offered commitment, $C' \in \mathcal{C} \setminus \{C\}$. As offering $C = 0$ (and being investigated with probability $\iota(0)$) gives different payoffs to the two types, to ensure that both of them are indifferent between offering C or C', it must be the case that C' induces as well the enforcer to open an investigation with positive probability; furthermore, as just seen, to deter B from deviating this probability must be lower than $\iota^S < 1$. It follows that the enforcer must again be indifferent between opening an investigation or closing the case, which requires ($\lambda_2(C') = \lambda^{ci}$, and) $C' = 0$, a contradiction.

Hence, $\mathcal{C} = \{0\}$ and the enforcer opens an investigation with probability $\iota(0) \leq \iota^S$. It follows that the participations are $P_\theta(\iota(0)$, given by (1), and $\lambda_1 = \lambda_1^e(\lambda, \iota(0))$, given by (2). The working condition $\lambda_1 = \lambda^{ci}$ therefore implies that $\iota(0) = \tilde{\iota}(\lambda)$, given by (7), which is decreasing in λ. That is, the candidate equilibrium coincides with the equilibrium that arises when commitments are not available. As $\tilde{\iota}(\lambda)$ decreases from 1 to 0 as λ increases from λ^N to λ^{ci} and $\iota(0) \leq \iota^S$, the prior cannot be too pessimistic, namely, λ must exceed the level for which $\tilde{\iota}(\lambda) = \iota^S$, which, from (17), amounts to $\lambda \geq \lambda^S$.

Conversely, for any $\lambda \in [\lambda^S, \lambda^{ci})$, there exists a pooling equilibrium that coincides with the equilibrium $\mathcal{E}_N$ arising when commitments are not available. Indeed, in this equilibrium a firm of type θ obtains a payoff $\Pi_\theta^N(\lambda) = 1 - \tilde{\iota}(\lambda)\bar{C}_\theta$, where $\tilde{\iota}(\lambda) \leq \iota^S$, implying that B obtains

$$\Pi_B^N(\lambda) \geq 1 - \hat{C},$$

whereas G obtains

$$\Pi_G^N(\lambda) \geq 1 - \iota^S \bar{C}_G = 1 - \frac{\bar{C}_G}{\bar{C}_B}\hat{C} > \max\{1 - \bar{C}_G, 1 - \hat{C}\},$$

where the inequality stems from $\hat{C} < \bar{C}_B$ and $\bar{C}_G < \bar{C}_B$. To establish existence, it suffices to interpret any deviant offer $\tilde{C}$ as signalling a bad type. The deviant offer $\tilde{C}$ would therefore be accepted only if it exceeds $\hat{C}$, implying that type θ would obtain $1 - \tilde{C} \leq 1 - \hat{C} \leq \Pi_\theta^N(\lambda)$; any lower $\tilde{C}$ would instead induce the enforcer to open an investigation, implying that type θ would obtain $1 - \bar{C}_\theta < \Pi_\theta^N(\lambda)$, where the inequality stems again from $\tilde{\iota}(\lambda) \leq \hat{C}/\bar{C}_B < 1$. Hence, there is no profitable deviation.

Case 3: $\lambda_1 < \lambda^{ci}$. By construction, there exists $C \in \mathcal{C}$ such that $\lambda_2(C) \leq \lambda_1$. If $\underline{\lambda}(C) > \lambda_2(C)$, then from Lemma 1 the enforcer would open an investigation and B would thus obtain $\Pi_B = 1 - \bar{C}_B$; however, B could secure $1 - \hat{C}\left(> 1 - \bar{C}_B\right)$ by offering $\hat{C}$, a contradiction. Hence, we must have $\underline{\lambda}(C) \leq \lambda_2(C) < \lambda^{ci}$, implying $C > 0$; from Lemma 1, the enforcer then accepts C, implying that both types obtain the same payoff, $1 - C\,(< 1)$.

Suppose now that there exists another offered commitment, $C' \in \mathcal{C} \setminus \{C\}$. By construction, offering C' must also give the same payoff to both types. This, in turn, implies that C' must be accepted with probability 1 (as any positive probability of investigation would generate different payoffs for the two types). But then, the firm cannot be indifferent between offering C or C'. Hence, $\mathcal{C} = \{C\}$; that is, both types offer C with probability 1 and, as they obtain the same payoff, participate for the same cost realizations. Hence, $\lambda_2(C) = \lambda_1 = \lambda$, implying $\lambda \geq \underline{\lambda}(C)$ or, equivalently, $C \geq \underline{C}(\lambda)$. Furthermore, as both types obtain $1 - C$, we must have $C \leq \min\left\{\hat{C}, \bar{C}_G\right\}$, as both types can secure $1 - \hat{C}$ by offering $\hat{C}$, and G can also secure $1 - \bar{C}_G$ by offering $\bar{C}_G$.

The candidate equilibria are therefore such that $\mathcal{C} = \{C\}$, where $C \in [\underline{C}(\lambda), \min\{\hat{C}, \bar{C}_G\})$, and $\lambda_2(C) = \lambda_1 = \lambda \geq \underline{\lambda}(C) \geq \underline{\lambda}(\bar{C}_G)$. Conversely, if $\lambda \geq \underline{\lambda}(\bar{C}_G)$, then for any $C \in [\underline{C}(\lambda), \min\left\{\hat{C}, \bar{C}_G\right\}]$ there exists a pooling equilibrium in which both types offer C, which is accepted, and thus obtain $1 - C$. To rule out profitable deviations, it suffices to interpret any deviant offer $\tilde{C}$ as signalling a bad type. $\tilde{C}$ would then be accepted only if it exceeds $\hat{C}$, in which case both types would obtain $1 - \tilde{C} \leq 1 - \hat{C} \leq 1 - C$; and if instead the enforcer were to open an investigation, each type θ would again obtain $1 - \bar{C}_\theta \leq 1 - \bar{C}_G \leq 1 - C$.

Summing-up, such pooling equilibria exist if and only if $\lambda \in \left[\underline{\lambda}(\bar{C}_G), \lambda^{ci}\right)$, and whenever they exist, the Pareto-efficient one is for the lowest admissible commitment, $\underline{C}(\lambda)$, which corresponds to $\mathcal{E}_C^P$, in which the firm offers the minimum acceptable commitment $\underline{C}(\lambda)$ and obtains $\Pi^P(\lambda) \equiv 1 - \underline{C}(\lambda)$.

J Proof of Corollary 4

In the Pareto-efficient pooling equilibrium $\mathcal{E}_C^P$, the firm offers $\underline{C}(\lambda)$, which is accepted, and thus obtains $\Pi^P(\lambda) \equiv 1 - \underline{C}(\lambda)$. Hence, the expected participation is

$$P^P(\lambda) \equiv F(1 - \underline{C}(\lambda)).$$

As by construction $W^a(\lambda, \underline{C}(\lambda)) = W^i(\lambda)$, expected welfare is then given by:

$$\mathcal{W}_C^P(\lambda) = P^P(\lambda)\left[\lambda W - (1-\lambda)L\right](1 - \underline{C}(\lambda)) = P^P(\lambda) W^i(\lambda).$$

K Proof of Proposition 5

The expected welfare generated by the equilibria $\mathcal{E}_N$ and $\mathcal{E}_C^S$, for $\lambda \in \left[0, \lambda^S\right)$, are respectively given by (10) and (18). Their difference can be expressed as:

$$\begin{aligned} \Delta\mathcal{W}(\lambda) &= \mathcal{W}_C^S(\lambda) - \mathcal{W}_N(\lambda) \\ &= \left[P^S(\lambda) - P^N(\lambda)\right] W^i(\lambda_1^S(\lambda)) + P^N(\lambda)\left[W^i(\lambda_1^S(\lambda)) - W^i(\lambda_1^N(\lambda))\right], \end{aligned}$$

where $P^N(\lambda)$, $\lambda_1^N(\lambda)$ are given by (9), (11), whereas $P^S(\lambda)$ are $\lambda_1^S(\lambda)$ given by (19).

As noted at the beginning of Section 4.2.2, $\tilde{\iota}\left(\lambda^S\right) = \iota^S$. As $\iota^N(\lambda)$ is (weakly) decreasing in λ, and strictly so for $\lambda \in \left(\lambda^N, \lambda^{ci}\right)$, it follows that, in the relevant range $\lambda < \lambda^S$, we have:

$$\iota^N(\lambda) > \iota^N\left(\lambda^S\right) = \tilde{\iota}\left(\lambda^S\right) = \iota^S.$$

Participation is therefore lower in $\mathcal{E}_N$ than in $\mathcal{E}_C^S$: $P_\theta^N(\lambda) < P_\theta^S$ for $\theta = G, B$, and so $P^N(\lambda) < P^S(\lambda)$. Furthermore, the enforcer's interim belief is more optimistic in $\mathcal{E}_N$ than in $\mathcal{E}_C^S$: given the participation rates, these beliefs are respectively given by $\lambda_1^N(\lambda) = \lambda_1^e(\lambda, \iota^N(\lambda))$ and $\lambda_1^S(\lambda) = \lambda_1^e(\lambda, \iota^S)$, and $\iota^N(\lambda) > \iota^S$ thus implies $\lambda_1^N(\lambda) > \lambda_1^S(\lambda)$. In addition, $W^i(.)$ is negative (from Assumption 3), and increasing in λ. Hence, $W^i(\lambda_1^S(\lambda)) < W^i(\lambda_1^N(\lambda)) < 0$. It then follows from the expression of $\Delta W(\lambda)$ that $\mathcal{W}_C^S(\lambda) < \mathcal{W}_N(\lambda)$.

L Proof of Proposition 6

• Part (i). The equilibria $\mathcal{E}_N$ and $\mathcal{E}_C^P$ co-exist in the interval $\lambda \in \left[\underline{\lambda}(\bar{C}_G), \lambda^{ci}\right)$, where we have $\mathcal{W}_N(\lambda) = P^N(\lambda) W^i(\lambda^{ci})$ and $\mathcal{W}_C^P(\lambda) = P^P(\lambda) W^i(\lambda)$, where $W^i(\lambda) < W^i(\lambda^{ci}) < 0$. Furthermore, for $\lambda \leq \lambda^N$, $\iota^N(\lambda) = 1$ and so $P^P(\lambda) = F(1 - \underline{C}(\lambda)) > F(1 - \bar{C}_\theta) = F(1 - \iota^N(\lambda)\bar{C}_\theta) = P_\theta^N(\lambda)$ for $\theta = G, B$, implying that $P^P(\lambda) > P^N(\lambda)$; it follows that:

$$
\begin{aligned}
\Delta\mathcal{W}(\lambda) &\equiv \mathcal{W}_C^P(\lambda) - \mathcal{W}_N(\lambda) \\
&= P^P(\lambda)\left[W^i(\lambda) - W^i(\lambda^{ci})\right] + \left[P^P(\lambda) - P^N(\lambda)\right] W^i(\lambda^{ci}) \\
&< 0,
\end{aligned}
$$

where the inequality stems from both terms being negative, as (i) $P^P(\lambda) > 0$ and $W^i(\lambda) < W^i(\lambda^{ci})$ and (ii) $P^N(\lambda) > P^N(\lambda)$ and $W^i(\lambda^{ci}) < 0$.

• Part (ii). We now focus on the case $\lambda > \lambda^N$, where $\lambda_1^N(\lambda) = \lambda^{ci}$ and, from Corollary 2, $\mathcal{W}_N(\lambda)$ is strictly decreasing in λ. As $\lim_{\lambda \to \lambda^{ci}} \mathcal{W}_N(\lambda) = \lim_{\lambda \to \lambda^{ci}} \mathcal{W}_C^P(\lambda) = W^i(\lambda^{ci})$, it follows that commitments are never desirable if $\mathcal{W}_C^P(\lambda)$ is (weakly) increasing in λ. Using $W^i(\lambda) = [W + (1-\rho)L](\lambda - \lambda^i)$, where

$$
\lambda^i \equiv \frac{(1-\rho)L + k}{W + (1-\rho)L}
$$

is such that $W^i(\lambda^i) = 0$, and

$$
\begin{aligned}
P^P(\lambda) &= F(1 - \underline{C}(\lambda)) = F\left(1 - \frac{\rho\lambda^c(\lambda^{ci} - \lambda)}{\lambda^c - \lambda}\right), \\
\frac{dP^P}{d\lambda}(\lambda) &= f(1 - \underline{C}(\lambda)) \frac{\rho\lambda^c(\lambda^c - \lambda^{ci})}{(\lambda^c - \lambda)^2} > 0,
\end{aligned}
$$

the derivative of $\mathcal{W}_C^P(\lambda)$ can be expressed as:

$$
\begin{aligned}
\frac{d\mathcal{W}_C^P}{d\lambda}(\lambda) &= P^P(\lambda)\frac{dW^i}{d\lambda}(\lambda) + \frac{dP^P}{d\lambda}(\lambda)W^i(\lambda) \\
&= [W + (1-\rho)L] F(1 - \underline{C}(\lambda)) \left[1 - \frac{h(1 - \underline{C}(\lambda))}{g(\lambda)}\right],
\end{aligned} \tag{28}
$$

where $h\left(1-\underline{C}\left(\lambda\right)\right)$ and

$$g\left(\lambda\right)\equiv\frac{\left(\lambda^{c}-\lambda\right)^{2}}{\rho\lambda^{c}(\lambda^{c}-\lambda^{ci})\left(\lambda^{i}-\lambda\right)}$$

are both strictly decreasing in λ:

$$\frac{d}{d\lambda}\left(h\left(1-\underline{C}\left(\lambda\right)\right)\right)=-h'\left(1-\underline{C}\left(\lambda\right)\right)\underline{C}'\left(\lambda\right)<0,$$

where the inequality stems from $h\left(\cdot\right)$ and $\underline{C}\left(\cdot\right)$ being both decreasing functions, and:

$$g'\left(\lambda\right)=-\frac{\left(\lambda^{c}-\lambda\right)\left(2\lambda^{i}-\lambda^{c}-\lambda\right)}{\rho\lambda^{c}(\lambda^{c}-\lambda^{ci})\left(\lambda^{i}-\lambda\right)^{2}}<0,$$

where the inequality follows from $\lambda<\lambda^{ci}<\lambda^{c}<\lambda^{i}$. It follows that $\mathcal{W}_{C}^{P}\left(\lambda\right)$ is (weakly) increasing in λ in the range $[\lambda^{N},\lambda^{ci})$ if

$$h\left(1-\underline{C}\left(\lambda^{N}\right)\right)\leq\underline{h}\equiv g\left(\lambda^{ci}\right)=\frac{\lambda^{c}-\lambda^{ci}}{\rho\lambda^{c}\left(\lambda^{i}-\lambda^{ci}\right)}.$$

• Part (iii). As already noted, $\mathcal{W}_{C}^{P}\left(\lambda^{ci}\right)=\mathcal{W}_{N}\left(\lambda^{ci}\right)$. Furthermore, using (10), (28), (27) and $W^{i}\left(\lambda\right)=\left[W+\left(1-\rho\right)L\right]\left(\lambda-\lambda^{i}\right)$, the condition:

$$\frac{d\mathcal{W}_{C}^{P}}{d\lambda}\left(\lambda^{ci}\right)>\frac{d\mathcal{W}_{N}}{d\lambda}\left(\lambda^{ci}\right).$$

is equivalent to:

$$\left[W+\left(1-\rho\right)L\right]F\left(1\right)\left[1-\frac{h\left(1\right)}{g\left(\lambda^{ci}\right)}\right]>-\left[W+\left(1-\rho\right)L\right]F\left(1\right)\frac{\lambda^{i}-\lambda^{ci}}{1-\lambda^{ci}}\left[\frac{\bar{C}_{B}}{\lambda^{ci}\left(\bar{C}_{B}-\bar{C}_{G}\right)}-1\right],$$

which amounts to $h\left(1\right)<\overline{h}$, where:

$$\begin{aligned}\overline{h}&\equiv\underline{h}\left\{1+\frac{\lambda^{i}-\lambda^{ci}}{1-\lambda^{ci}}\left[\frac{\bar{C}_{B}}{\lambda^{ci}\left(\bar{C}_{B}-\bar{C}_{G}\right)}-1\right]\right\}\\&=\underline{h}\left\{\frac{1-\lambda^{i}}{1-\lambda^{ci}}+\frac{\lambda^{i}-\lambda^{ci}}{\lambda^{ci}\left(1-\lambda^{ci}\right)}\frac{\bar{C}_{B}}{\left(\bar{C}_{B}-\bar{C}_{G}\right)}\right\}\left(>\underline{h}\right).\end{aligned}$$

It follows that, whenever $h\left(1\right)<\overline{h}$, commitments are undesirable for λ close to λ^{ci}; that is, there then exists $\overline{\lambda}<\lambda^{ci}$ such that:

$$\frac{d\mathcal{W}_{C}^{P}}{d\lambda}\left(\lambda\right)>\frac{d\mathcal{W}_{N}}{d\lambda}\left(\lambda\right)$$

for $\lambda\in(\overline{\lambda}\lambda^{ci})$, implying that $\mathcal{W}_{C}^{P}\left(\lambda\right)<\mathcal{W}_{N}\left(\lambda\right)$ in that range.

M Proof of Lemma 3

A firm of type θ accepts the commitment if $C < \bar{C}_\theta$, rejects it if $C > \bar{C}_\theta$ and is indifferent between the two options if $C = \bar{C}_\theta$, where $\bar{C}_G < \bar{C}_B$. Hence, both types accept if $C < \bar{C}_G$, both types reject if $C > \bar{C}_B$, and G accepts whereas type B rejects if $C \in (\bar{C}_G, \bar{C}_B)$.

If $C = \bar{C}_G$ type B accepts and type G, being indifferent, is willing to reject with any probability $\nu_G \in [0, 1]$. Expected welfare is equal to:

$$W(\lambda_1, \bar{C}_G, \nu_G) = \lambda_1 \left[\nu_G (W - k) + (1 - \nu_G)(1 - \bar{C}_G)W\right] - (1 - \lambda_1)(1 - \bar{C}_G)L.$$

Taking the derivative with respect to ν_G yields:

$$\frac{\partial W(\cdot)}{\partial \nu_G} = \lambda_1 \left[W - k - (1 - \bar{C}_G)W\right] = \lambda_1 k (sW - 1) < 0,$$

as $sW < sk = \bar{C}_G < \bar{C}_B < 1$, where the first inequality stems from Assumption 3 and the last one from Assumption 1. As by construction the two types of firms obtain the same payoffs in all equilibria, the pooling equilibrium on acceptance Pareto-dominates all the other equilibria.

Finally, if $C = \bar{C}_B$ type G rejects and type B, being indifferent, is willing to reject with any probability $\nu_B \in [0, 1]$. Expected welfare is equal to:

$$W(\lambda_1, \bar{C}_B, \nu_B) = \lambda_1 (W - k) - (1 - \lambda_1) \left\{\nu_B[(1 - \rho)L + k] + (1 - \nu_B)(1 - \bar{C}_B)L\right\}.$$

Taking the derivative with respect to ν_B and using $\bar{C}_B = sk + \rho(1 + S)$ yields:

$$\frac{\partial W(\cdot)}{\partial \nu_B} = -(1 - \lambda_1) [k + (sk + \rho S)L] < 0.$$

Hence, the separating equilibrium Pareto-dominates all the other continuation equilibria.

N Proof of Lemma 4

The welfare differential $\Delta(\lambda_1) \equiv W^s(\lambda_1, \bar{C}_B) - W^a(\lambda_1, \bar{C}_G)$ satisfies:

$$\Delta'(\lambda_1) = -(k - \bar{C}_G W) - (\bar{C}_B - \bar{C}_G)L < 0,$$

where the inequality stems from Assumptions 1 and 3, which together imply $k > \bar{C}_B W (> \bar{C}_G W)$, and $\Delta(\lambda_1) = 0$ for

$$\hat{\lambda}_1 \equiv \frac{\left(\bar{C}_B - \bar{C}_G\right) L}{\left(\bar{C}_B - \bar{C}_G\right) L + k - \bar{C}_G W}. \tag{29}$$

It follows that $W^s(\lambda_1, \bar{C}_B) \gtreqqless W^a(\lambda_1, \bar{C}_G)$ if and only if $\lambda_1 \lesseqqgtr \hat{\lambda}_1$.Furthermore, the condition $k > \bar{C}_B W (> \bar{C}_G W)$ ensures that the denominator in (29) is strictly larger than $\left(\bar{C}_B - \bar{C}_G\right)(W + L)\,(> 0)$, implying $\hat{\lambda}_1 < L/(W + L) = \lambda^c$. In addition:

$$\begin{aligned} \Delta\left(\lambda^{ci}\right) &= W^i(\lambda^{ci}) + (1 - \lambda^{ci})[k + (sk + \rho S)L] - \left(1 - \bar{C}_G\right) W^c(\lambda^{ci}) \\ &= \bar{C}_G W^c(\lambda^{ci}) + (1 - \lambda^{ci})[k + (sk + \rho S)L] \\ &= \lambda^{ci} skW + (1 - \lambda^{ci})\,(k + \rho SL)\,, \end{aligned}$$

where the first equality stems from (21) and the definition of $W^a(.)$, the second one from $W^c(\lambda^{ci}) = W^i(\lambda^{ci})$ (by the definition of λ^{ci}), and the last one from $\bar{C}_G = sk$ and the expression of $W^c(\cdot)$. As the last expression is positive, it follows that $\Delta(\lambda^{ci}) > 0$, implying $\hat{\lambda}_1 > \lambda^{ci}$.

Building on this, we have:

- If $\lambda_1 < \lambda^c$, implying $W^c(\lambda_1) < 0$, then the best choice in the range $\left[0, \bar{C}_G\right]$ is $\bar{C}_G$. Furthermore:
  - if $0 < \lambda_1 < \hat{\lambda}_1\,(< \lambda^c)$, then $\bar{C}_G$ is dominated by $\bar{C}_B$ and the associated separating continuation equilibrium; hence, the enforcer's optimal response is $\bar{C}_B$;
  - if instead $\lambda_1 = \hat{\lambda}_1\,(< \lambda^c)$, then the enforcer is now indifferent between $\bar{C}_G$, prompting acceptance, and $\bar{C}_B$, prompting separation; hence, any lottery over $\left\{\bar{C}_B, \bar{C}_G\right\}$ constitutes an optimal response for the enforcer.
  - finally, if $\hat{\lambda}_1 < \lambda_1 < \lambda^c$, then the enforcer now prefers $\bar{C}_G$, prompting acceptance, to $\bar{C}_B$, prompting separation; hence, the enforcer's optimal response is $\bar{C}_G$.
- If instead $\lambda_1 = \lambda^c \left(> \hat{\lambda}_1\right)$, then the enforcer is indifferent between any C in the range $\left[0, \bar{C}_G\right]$ (as $W^c(\lambda_1) = 0$) and prefers any such commitment to $\bar{C}_B$ (as $\lambda_1 > \hat{\lambda}_1$); hence, any lottery over $\left[0, \bar{C}_G\right]$ constitutes an optimal response for the enforcer.
- Finally, if $\lambda_1 > \lambda^c \left(> \hat{\lambda}_1\right)$, then the best choice in the range $\left[0, \bar{C}_G\right]$ is 0 (as $W^c(\lambda_1) > 0$) and this choice dominates $\bar{C}_B$ (as $\lambda_1 > \hat{\lambda}_1$); hence, the enforcer's optimal response is 0 (i.e., not opening a case).

O Proof of Proposition 7

From Lemma 4, the enforcer offers $C \in [0, \bar{C}_G] \cup \{\bar{C}_B\}$. Furthermore, offering any $C \in [0, \bar{C}_G]$ gives the same payoff to both types of firm, whereas offering $C = \bar{C}_B$ gives a strictly lower payoff to B than to G. It follows that, in equilibrium, G undertakes the action weakly more often than B. This, in turn, implies that the enforcer can only become (weakly) more optimistic. Specifically, $\lambda_1 = \lambda$ if the enforcer is not expected to offer $\bar{C}_B$, and $\lambda_1 \in (\lambda, \bar{\lambda}_1^S(\lambda)]$ otherwise, the upper bound corresponding to the case where $\bar{C}_B$ is offered with probability 1.

Consider first the case $\lambda < \hat{\lambda}$. We then have $\lambda_1 \leq \bar{\lambda}_1^S(\lambda) < \hat{\lambda}_1$, implying from Lemma 4 that the enforcer is expected to offer $C = \bar{C}_B$ with probability 1. This induces participation rates equal to $\bar{P}_\theta^S$, which in turn yields $\lambda_1 = \bar{\lambda}_1^S(\lambda)(\leq \hat{\lambda}_1)$. Hence, there exists a unique equilibrium, in which the participation rates are indeed $\bar{P}_\theta^S$ and the enforcer offers $C = \bar{C}_B$, which is accepted by B and rejected by G.

Consider next the case $\lambda \geq \hat{\lambda}_1$. If the firm were to expect the enforcer to offer $\bar{C}_B$ with positive probability, we would have $\lambda_1 > (\lambda \geq)\hat{\lambda}_1$; but then, from Lemma 4 the enforcer would not offer $\bar{C}_B$, a contradiction. Hence, the firm must expect the enforcer to offer $C \in [0, \bar{C}_G]$, and both types thus participate with the same probability; it follows that $\lambda_1 = \lambda(\geq \hat{\lambda}_1)$. If $\lambda \in (\hat{\lambda}_1, \lambda^c)$, then from Lemma 4 the firm expects the enforcer to offer $C = \bar{C}_G$, which gives the same participation rate $\bar{P}_G^S$ for both types. If instead $\lambda > \lambda^c$, then from Lemma 4 the firm expects the case not to be opened and participation rates are both equal to $F(1)$. In both instances, the described strategies constitute the unique equilibrium. By contrast, in the boundary case $\lambda = \hat{\lambda}_1$, the enforcer is indifferent between offering $\bar{C}_G$ or $\bar{C}_B$. G would accept $\bar{C}_G$ and reject $\bar{C}_B$, and is thus also indifferent between these two offers. B would instead accept both and prefers the lower commitment. It follows that offering $\bar{C}_G$ constitutes the unique Pareto-dominant equilibrium. Last, in the boundary case $\lambda = \lambda^c$, the enforcer is indifferent between any $C \in [0, \bar{C}_G]$, whereas both types of firm prefer the case not to be opened; hence, the unique Pareto-dominant equilibrium is the one in which, with probability 1, the enforcer does not open the case.

Finally, consider the case $\lambda \in (\hat{\lambda}, \hat{\lambda}_1)$. If the firm were to expect the enforcer to offer $\bar{C}_B$ with probability 1, we would have $\lambda_1(= \bar{\lambda}_1^S(\lambda)) > \hat{\lambda}_1$; but then, from Lemma 4 the enforcer would not offer $\bar{C}_B$, a contradiction. Conversely, if the firm were to expect the enforcer to offer $\bar{C}_B$ with probability 0, we would have $\lambda_1(= \lambda) < \hat{\lambda}_1$; but then, from Lemma 4 the enforcer would offer $\bar{C}_B$, a contradiction. Hence, the firm must expect the enforcer to offer $C = \bar{C}_B$ with a probability lying strictly between 0 and 1; from Lemma 4, it follows that the enforcer's updated beliefs must be equal to $\hat{\lambda}_1$, and that the enforcer must randomize between $\bar{C}_B$ and $\bar{C}_G$. Denoting by σ the probability of offering $\bar{C}_G$, the participation rates are then $\bar{P}_G^S$ for G and

$$\bar{P}_B(\sigma) \equiv F(1 - \sigma\bar{C}_G - (1 - \sigma)\bar{C}_\theta)$$

for B. These participation rates induce the enforcer to update its beliefs to:

$$\bar{\lambda}_1(\lambda, \sigma) = \frac{\lambda\bar{P}_G}{\lambda\bar{P}_G + (1 - \lambda)\bar{P}_B(\sigma)},$$

which is increasing in λ and decreasing in σ, from $\bar{\lambda}_1(\lambda, 0) = \bar{\lambda}_1^S(\lambda)(> \hat{\lambda}_1)$ to $\bar{\lambda}_1(\lambda, 1) = \lambda(< \hat{\lambda}_1)$. Hence, there exists a unique $\bar{\sigma}(\lambda)$, which increases with λ from $\bar{\sigma}(\hat{\lambda}) = 0$ to $\bar{\sigma}(\hat{\lambda}_1) = 1$, for which $\bar{\lambda}_1(\lambda, \bar{\sigma}(\lambda)) = \hat{\lambda}_1$. Summing-up, in equilibrium, G must undertake the action with probability $\bar{P}_G^S$ and B must do so with probability $\bar{P}_B^M(\lambda) \equiv \bar{P}_\theta(\bar{\sigma}(\lambda))$, which induces the enforcer to update its beliefs to $\lambda_1 = \hat{\lambda}_1$, and in response the enforcer offers $\bar{C}_G$ (which is accepted by both types) with probability $\bar{\sigma}(\lambda)$, otherwise it offers $\bar{C}_B$ (which G rejects and B accepts); conversely, these strategies do constitute an equilibrium (and, thus, the unique one).

P Proof of Proposition 8

Expected welfare under the reformed regulation is given by (23). From Corollary 2, in the benchmark regime it is instead equal to

$$\mathcal{W}_N(\lambda) = P^N(\lambda) W^i(\lambda_1^N(\lambda)),$$

where the participation $P^N(\lambda)$ and the interim belief $\lambda_1^N(\lambda)$ are respectively given by (9) and (11). Finally, under the current regulation, the pooling equilibrium is the only one that can generate a higher expected welfare than in the benchmark regime; hence, the relevant welfare expression is

$$\mathcal{W}_C(\lambda) \equiv \mathcal{W}_C^P(\lambda) = P^P(\lambda) W^i(\lambda),$$

where $P^P(\lambda) \equiv F(1 - \underline{C}(\lambda))$.

We can distinguish three cases, depending on the enforcer's pessimism.

• **Case 1:** $\lambda \leq \lambda^N$. In this case, we know from Proposition 1 that the enforcer investigates with probability 1 in the benchmark regime, and from Propositions 5 and 6 that the current regulation does not perform better. Furthermore, from Corollary 1, $P_\theta^N\left(\lambda^N\right) = F(1 - \bar{C}_\theta) = \bar{P}_\theta^S$, implying

$$\lambda_1^N(\lambda^N) = \bar{\lambda}_1^S(\lambda^N), \tag{30}$$

Moreover, from Corollary 2:

$$\lambda^N < \lambda_1^N(\lambda^N) = \lambda^{ci} \tag{31}$$

and from Proposition 7:

$$\hat{\lambda}_1 = \bar{\lambda}_1^S(\hat{\lambda}) \tag{32}$$

Together with Lemma 4, conditions (30)-(32) yield $\bar{\lambda}_1^S(\lambda^N) = \lambda^{ci} < \hat{\lambda}_1 = \bar{\lambda}_1^S(\hat{\lambda})$; as $\bar{\lambda}_1^S(\cdot)$ is strictly increasing, it follows that $\lambda^N < \hat{\lambda}$. From Proposition 7, under the reformed regulation, the enforcer therefore offers $\bar{C}_B$, which is accepted by B and rejected by G. As a result, compared with the benchmark regime, the reformed regulation generates the same participation rates but, when the action is bad, it enables the enforcer to save on the cost of investigation and to impose the maximal acceptable commitment. Specifically, when facing a good type, the enforcer investigates with probability 1 in both regimes (because it has pessimistic interim beliefs in the benchmark regime, and because G rejects the offered commitment under the reformed regulation); hence, both regimes then yield the same outcome. By contrast, conditional on facing a bad type, expected welfare is $W^i(0) = -L + (\rho L - k)$ in the benchmark regime and

$$W^S(0) = -(1 - \bar{C}_B)L = -L + [sk + \rho(1 + S)]L = W^i(0) + k + (sk + \rho S)L > W^i(0).$$

under the reformed regulation; it follows that the latter performs strictly better.

- **Case 2:** $\lambda \in (\lambda^N, \lambda^{ci})$. From Lemma 4, $\lambda^{ci} < \hat{\lambda}_1$. It follows that, under the reformed regulation, expected welfare is of the form $P_R(\lambda)W_R(\lambda)$, where:

$$P_R(\lambda) = \begin{cases} \bar{P}^S(\lambda) & \text{for } \lambda \leq \hat{\lambda}, \\ \bar{P}^M(\lambda) & \text{for } \lambda \in (\hat{\lambda}, \hat{\lambda}_1), \end{cases} \tag{33}$$

and

$$W_R(\lambda) = \begin{cases} W^S(\lambda) & \text{for } \lambda \leq \hat{\lambda}, \\ W^M & \text{for } \lambda \in (\hat{\lambda}, \hat{\lambda}_1). \end{cases} \tag{34}$$

In each of the other two regimes, that is, for $\tau = N, C$, expected welfare is instead of the form

$$\mathcal{W}_\tau(\lambda) \equiv P_\tau(\lambda)W^i(\lambda_1^\tau(\lambda)),$$

where:

$$P_\tau(\lambda) = \begin{cases} P^N(\lambda) & \text{for } \tau = N, \\ P^P(\lambda) & \text{for } \tau = C, \end{cases} \tag{35}$$

and

$$\lambda_1^\tau(\lambda) = \begin{cases} \lambda^{ci} & \text{for } \tau = N, \\ \lambda & \text{for } \tau = C. \end{cases} \tag{36}$$

The welfare differential between the reformed regulation and the regime $\tau = N, C$ can thus be expressed as:

$$\mathcal{W}_R(\lambda) - \mathcal{W}_\tau(\lambda) = [P_R(\lambda) - P_\tau(\lambda)]\, W^i(\lambda_1^\tau(\lambda)) + P_R(\lambda)\left[W_R(\lambda) - W^i(\lambda_1^\tau(\lambda))\right]. \tag{37}$$

We first note that the current regulation induces a lower participation than the other two regimes:

Claim 1 (participation) $P_R(\lambda) < P_\tau(\lambda)$ *for* $\lambda \in (\lambda^N, \lambda^{ci})$ *and* $\tau = N, C$.

Proof. We first compare $P_R(\lambda)$ to $P_C(\lambda) = P^P(\lambda)$. From (33), $P_R(\lambda) \leq \max\{\bar{P}^S(\lambda), \bar{P}^M(\lambda)\} < \bar{P}_G^S$. By contrast, $P^P(\lambda) \geq \bar{P}_G^S$ from Proposition 4 (recalling that, under the current regulation, the relevant pooling equilibrium, in which the firm offers $\underline{C}(\lambda)$, exists only when $\underline{C}(\lambda) \leq \bar{C}_G$). Hence, $P_R(\lambda) < P^P(\lambda)$.

We now compare $P_R(\lambda)$ to $P_N(\lambda) = P^N(\lambda)$. For $\lambda^N < \lambda \leq \hat{\lambda}$, it follows from Proposition 1 and Corollary 1 that $P^N(\lambda) > \bar{P}^S(\lambda) = P_R(\lambda)$, because the participation $P^N(\lambda)$ is then based on an investigation rate $\iota(\lambda) < 1$, whereas the participation rate $\bar{P}^S(\lambda)$ is based on profits equivalent to those obtained with an investigation rate equal to 1.

For $\hat{\lambda} < \lambda < \lambda^{ci}$ (which can arise when $\hat{\lambda} < \lambda^{ci}$), $P_R(\lambda) = \bar{P}^M(\lambda)$ from Corollary 5. Furthermore, it follows from Corollary 2, Proposition 7 and $\lambda^{ci} < \hat{\lambda}_1$ that:

$$\frac{\lambda P_G^N(\lambda)}{P^N(\lambda)} = \lambda^{ci} < \hat{\lambda}_1 = \frac{\lambda \bar{P}_G^S}{P^M(\lambda)},$$

implying

$$P^N(\lambda) > \frac{P_G^N(\lambda)}{\bar{P}_G^S} P^M(\lambda) \geq P^M(\lambda),$$

where the last inequality stems from $P_G^N(\lambda) > \bar{P}_G^S$, as the participation $P_G^N(\lambda)$ is based on an investigation rate $\iota(\lambda) < 1$, whereas the participation rate P_G^S is based instead on an investigation rate equal to 1. ■

From (36), $\lambda_1^\tau(\lambda) \leq \lambda^{ci} < \lambda^c$ for $\lambda \in (\lambda^N, \lambda^{ci})$; together with Claim 1, this implies that the first term in (37) is positive.

Next, we show that the enforcer's interim belief is higher under the reformed regulation; that is, letting $\lambda_1^R(\lambda)$ denote the enforcer's interim belief under the reformed regulation, we have:

Claim 2 (interim belief) $\lambda_1^R(\lambda) > \lambda_1^\tau(\lambda)$ *for* $\lambda \in (\lambda^N, \lambda^{ci})$ *and* $\tau = N, C$.

Proof. From Proposition 7 and (30)-(31), in the range $\lambda \in (\lambda^N, \lambda^{ci})$ we have $\lambda_1^R(\lambda) > \lambda_1^R(\lambda^N) = \lambda^{ci}$, whereas from (36), $\lambda_1^C(\lambda) < \lambda_1^B(\lambda) = \lambda^{ci}$. Hence, $\lambda_1^R(\lambda) > \lambda_1^\tau(\lambda)$ for $\tau = N, C$. ■

To conclude the argument, it suffices to note that, under the reformed regulation, the enforcer could choose to offer an unacceptable commitment; therefore, we have $W_R(\lambda) \geq W^i\left(\lambda_1^R(\lambda)\right) > W^i(\lambda_1^\tau(\lambda))$ for $\tau = N, C$, implying that the second term in (37) is also positive since $P_R(\lambda) > 0$.

- **Case 3:** $\lambda \in (\lambda^{ci}, \lambda^c)$.

In regime $\tau = N, C$ for $\lambda \in [\lambda^{ci}, \lambda^{c})$ there is no enforcement; hence:

$$\mathcal{W}_\tau(\lambda) = P_\tau(\lambda) W_\tau(\lambda) = F(1) W^c(\lambda).$$

The welfare differential between the reformed regulation and the regime $\tau = N, C$ can thus be expressed as

$$\mathcal{W}_R(\lambda) - \mathcal{W}_\tau(\lambda) = \left[P_R(\lambda) - P_\tau(\lambda)\right] W^c(\lambda) + P_R(\lambda) \left[W_R(\lambda) - W^c(\lambda)\right],$$

where:

- for $\lambda \leq \hat{\lambda}$, $\bar{P}^S(\lambda) = P_R(\lambda)$ and $W_R(\lambda) = W^S(\lambda)$;
- for $\lambda \in (\hat{\lambda}, \hat{\lambda}_1)$, $P_R(\lambda) = \bar{P}^M(\lambda)$ and $W_R(\lambda) = W^M$;
- for $\lambda \in [\hat{\lambda}_1, \lambda^c)$, $P_R(\lambda) = \bar{P}^A(\lambda)$ and $W_R(\lambda) = W^A(\lambda)$.

Participation is obviously higher in the absence of enforcement: $F(1) > P_R(\lambda)$. Moreover, $W^c(\lambda) < 0$ for $\lambda < \lambda^c$. Hence, the first term is positive. To conclude the argument, it suffices to note that, under the reformed regulation, at stage 1 the enforcer could choose to close the case (i.e., offer $C = 0$); therefore, we have $W_R(\lambda) \geq W^c\left(\lambda_1^R(\lambda)\right) \geq W^c(\lambda)$, where the second equality stems from the fact that any enforcement deters bad types weakly more than good types. As $P_R(\lambda) > 0$, it follows that the second term is also positive.

Online Appendix

A Separating equilibria

We study here the scope for separating equilibria in the policy regime considered in Section 4. For this purpose, we (maintain Assumptions 1 and 2, but) relax somewhat Assumption 3. Specifically, we will maintain the following Assumptions:

Assumption 1: $S < (1 - sk - \rho)/\rho$.

Assumption 2: $k < \rho L$.

Assumption 3′: $k > \underline{k} \equiv (1 - \lambda^c)\rho L$.

As already noted, Assumption 1 amounts to $\bar{C}_B < 1$ and thus ensures that the firm always undertakes the action when it is costless, whereas Assumption 2 amounts to $W^i(0) > W^c(0)$, and thus implies $\lambda^{ci} > 0$.

Assumption 3′ amounts instead to $W^c\left(\lambda^{ci}\right) < 0$, implying $\lambda^{ci} < \lambda^c$; using (5) then yields

$$0 > W^i\left(\lambda^c\right) = W^c\left(\lambda^c\right) + \rho L\left(\lambda^{ci} - \lambda^c\right) = \rho L\left(\lambda^{ci} - \lambda^c\right),$$

implying $\lambda^c > \lambda^{ci}$. Under Assumptions 2 and 3′ we thus have:

$$0 < \lambda^{ci} < \lambda^c < \min\left\{\lambda^i, 1\right\}, \tag{A.1}$$

where

$$\lambda^i \equiv \frac{k + (1-\rho)L}{W + (1-\rho)L}$$

denotes the threshold for which $W^i\left(\lambda^i\right) = 0$, which satisfies:

$$\lambda^i \lesseqgtr 1 \Longleftrightarrow k \lesseqgtr W. \tag{A.2}$$

A.1 Enforcer's response

We first characterize the enforcer's response under Assumptions 1, 2 and 3′. Fix the offered commitment C and the enforcer's revised belief $\lambda_2 = \lambda_2(C)$, and let

$$\underline{\lambda}(C) \equiv \frac{\rho\lambda^{ci} - C}{\rho\lambda^c - C}\lambda^c$$

denote as before the unique solution in λ_2 to $\Delta(\lambda_2, C) = 0$, where

$$\Delta(\lambda_2, C) \equiv W^a(\lambda_2, C) - W^i(\lambda_2) = (W + L)\left[C\left(\lambda^c - \lambda_2\right) - \rho\lambda^c(\lambda^{ci} - \lambda_2)\right].$$

We have:

$$\underline{\lambda}'(C) = -\frac{\left(\lambda^c - \lambda^{ci}\right)\rho\lambda^c}{\left(\lambda^c\rho - C\right)^2} < 0,$$

where the inequality stems from (A.1). Furthermore, $\underline{\lambda}\,(C)$ tends to $-\infty$ as C tends to $\rho\lambda^c_-$, tends instead to $+\infty$ as C tends to $\rho\lambda^c_+$, and satisfies $\underline{\lambda}(\hat{C}) = 0$ and $\underline{\lambda}(\check{C}) = 1$, where $\hat{C} = \rho\lambda^{ci}\,(< \rho\lambda^c)$ and

$$\check{C} = \frac{1-\lambda^{ci}}{1-\lambda^c}\rho\lambda^c\,(> \rho\lambda^c)\,,$$

where $\check{C} \leq 1$ if and only if $\lambda^i \leq 1$.[22]

Finally, $\Delta(0, C) \gtreqless 0$ if and only if $C \gtreqless \hat{C} = \rho\lambda^{ci}$. Therefore:

- for $C \in \left[0, \hat{C}\right]$, we have $\Delta(0, C) \leq 0$ and $\underline{\lambda}(C) \geq 0$; it follows that $\Delta(\lambda_2, C) \geq 0$ if and only if $\lambda_2 \geq \underline{\lambda}(C)$;

- for $C \in \left(\hat{C}, \check{C}\right)$, we have $\Delta(0, C) > 0$ and $\underline{\lambda}(C) \notin [0, 1]$; it follows that $\Delta(\lambda_2, C) > 0$ for any $\lambda_2 \in [0, 1]$;

- for $C \in \left[\check{C}, 1\right]$, we have $\Delta(0, C) > 0$ and $\underline{\lambda}(C) > 0$; it follows that $\Delta(\lambda_2, C) \geq 0$ if and only if $\lambda_2 \leq \underline{\lambda}(C)$.

Summing-up, we have:

Lemma A.5 (enforcer's response) *The enforcer's response to an offered commitment C is given by:*

- if $C = 0$, the enforcer closes the case if $\lambda_2 \geq \lambda^{ci}\,(= \underline{\lambda}(0))$, runs an investigation if $\lambda_2 < \lambda^{ci}$, and is indifferent between the two options if $\lambda_2 = \lambda^{ci}$;

- if $C \in \left(0, \hat{C}\right)$, the enforcer accepts C if $\lambda_2 \geq \underline{\lambda}(C)$, and otherwise runs an investigation;

- if $C \in \left[\hat{C}, \check{C}\right]$, the enforcer accepts C regardless of its revised belief;

- if $C \in (\check{C}, 1]$, the enforcer accepts C if $\lambda_2 \leq \underline{\lambda}(C)$, and otherwise runs an investigation.

In other words, if $k \geq W$, implying ($\lambda^i \geq 1$, and) $\check{C} \geq 1$, the enforcer's response is to accept any $C \geq \hat{C}$, and to accept any lower C only if $\lambda_2\,(C)$ exceeds $\underline{\lambda}(C)$, as described in Lemma 1. If instead $k < W$, implying $\check{C} < 1$, then the enforcer's response is to accept any $C \in \left(\hat{C}, \check{C}\right)$, to accept any $C < \hat{C}$ only if $\lambda_2\,(C)$ exceeds $\underline{\lambda}(C)$, and to accept any $C > \check{C}$ only if instead $\lambda_2\,(C)$ lies *below* $\underline{\lambda}(C)$.

[22] To see this, it suffices to note that $\underline{\lambda}(1) = \lambda^i$ as, by construction, $\Delta(\lambda_2, 1) = -W^i\,(\lambda_2)$.

A.2 Separating equilibria

For any commitment $C \in [0,1]$, let $\lambda_2(C)$ denote the enforcer's revised belief when offered C, and $\alpha(C)$ its probability of acceptance. In addition, let $\mathcal{C}_\theta \neq \varnothing$ denote the set of commitments offered by type $\theta \in \{B, G\}$. Separating equilibria are such that the offered commitments reveal the type of the firm; that is, $\mathcal{C}_B \cap \mathcal{C}_B = \varnothing$, implying $\lambda_2(C) = 0$ for any $C \in \mathcal{C}_B$ and $\lambda_2(C) = 1$ for any $C \in \mathcal{C}_B$. The following proposition characterizes these equilibria (where the superscript S stands for separation).

Proposition 9 (separating equilibria) *Separating equilibria exist if only if $k > W$ and $\bar{C}_G \leq \hat{C}$. Furthermore, in any separating equilibrium, $\mathcal{C}_B = \{\hat{C}\}$ and $\alpha(\hat{C}) = 1$, whereas $\mathcal{C}_G \geq \check{C}$ and $\alpha(\mathcal{C}_G) = 0$ for every $C_G \in \mathcal{C}_G$; B thus obtains $\Pi_B = \Pi_B^S \equiv 1 - \hat{C}$, whereas G obtains $\Pi_G = \Pi_G^S \equiv 1 - \bar{C}_G$.*

Proof. The proof is decomposed into three steps.

- *Step 1: In a separating equilibrium, $\mathcal{C}_B = \{\hat{C}\}$, $\alpha(\hat{C}) = 1$ and $\Pi_B = \Pi_B^S$.* Recall that B can secure $\Pi_B^S = 1 - \hat{C}$ by offering $\hat{C}$, which is always accepted by the enforcer. Furthermore, for any $C_B \in \mathcal{C}_B$, $\lambda_2(C_B) = 0$; the enforcer thus accepts C_B if $C_B \geq \hat{C}$, and rejects it otherwise. Therefore, if it were the case that $C_B > \hat{C}$, the enforcer would accept C_B and B would obtain $1 - C_B < 1 - \hat{C} = \Pi_B^S$, a contradiction. Conversely, if it were the case that $C_B < \hat{C}$, the enforcer would reject C_B and B would obtain $1 - \bar{C}_B < 1 - \hat{C} = \Pi_B^S$, another contradiction. Hence, $\mathcal{C}_B = \{\hat{C}\}$, $\alpha(\hat{C}) = 1$ and $\Pi_B = \Pi_B^S$.

- *Step 2: In a separating equilibrium, $\Pi_G = \Pi_G^S$ and, for every $C_G \in \mathcal{C}_G$, $C_G \geq \check{C}$ and $\alpha(C_G) = 0$.* If it were the case that $\alpha(C_G) = 1$ for some $C_G \in \mathcal{C}_G$, preventing both types from mimicking each other would require $C_G = C_B (= \hat{C})$, a contradiction. Likewise, if it were the case that $\alpha(C_G) \in (0,1)$ for some $C_G \in \mathcal{C}_G$, then the rule would imply $C_G = 0$ (i.e., no commitment); as $\lambda_2(C) = 1$ for any $C \in \mathcal{C}_G$, the enforcer would then close the case (i.e., accept $C_G = 0$), giving B an incentive to offer $C_G = 0$ as well, a contradiction. Hence, $\alpha(C_G) = 0$ for every $C_G \in \mathcal{C}_G$, implying $\Pi_G = 1 - \bar{C}_G = \Pi_G^S$.

  By construction, any $C_G \in \mathcal{C}_G$ induces the enforcer to revise its belief to $\lambda_2(C_G) = 1$. From Lemma A.5, the enforcer would therefore accept C_G unless it strictly exceeds $\check{C}$. Hence, $C_G > \check{C}$ for every $C_G \in \mathcal{C}_G$.

- *Step 3: A separating equilibrium exists if and only if $\lambda^i \leq 1$ and $\bar{C}_G \leq \hat{C}$.* From step 2, in a separating equilibrium, any $C_G \in \mathcal{C}_G$ satisfies $C_G > \check{C}$. As by construction $C_G \leq 1$, it follows that $\check{C} < 1$, which requires $k > W$. Furthermore, from steps 1 and 2, G obtains $\Pi_G^S = 1 - \bar{C}_G$ and could obtain $1 - \hat{C}$ by mimicking B; preventing such a deviation thus requires $\bar{C}_G \leq \hat{C}$. Hence, a separating equilibrium exists only if $\lambda^i \leq 1$ and $\bar{C}_G \leq \hat{C}$.

Conversely, suppose that $\lambda^i \leq 1$ and $\bar{C}_G \leq \hat{C}$, and consider a candidate equilibrium satisfying $\mathcal{C}_B = \{\hat{C}\}$, $\alpha(\hat{C}) = 1$, $C_G \geq \check{C}$ and $\alpha(C_G) = 0$ for every $C_G \in \mathcal{C}_G$, implying $\Pi_B = \Pi_B^S$ and $\Pi_G = \Pi_G^S$.

As just noted, the condition $\bar{C}_G \leq \hat{C}$ ensures that G has no incentive to mimic B. Furthermore, B has no incentive to mimic G, as doing so would give B a payoff $1 - \bar{C}_B < 1 - \hat{C} = \Pi_B^S$ (as G's offers are rejected). Likewise, whenever a deviant offer would trigger an investigation, B would obtain again $1 - \bar{C}_B < 1 - \hat{C} = \Pi_B^S$, whereas G would obtain its equilibrium payoff. To establish existence, it then suffices to assume that any unexpected offer $C \notin \mathcal{C} \equiv \mathcal{C}_B \cup \mathcal{C}_G$ is interpreted by the enforcer as signaling a bad type. The enforcer would therefore accept the offer only if $C > \hat{C}$, making the deviation unprofitable for both types. ■