Strategic Experimentation, Enforcement Transparency, and Reputation *

Yi Chen† Kai Du‡ Phillip Stocken§ Zhe Wang¶

October 10, 2022

Abstract

We consider a two-period model featuring two self-interested agents that test a sequentially rational principal's enforcement propensity through their misconduct and a principal that disciplines them to build a reputation for strict enforcement. We show that a transparent enforcement setting, which allows the principal to "make an example" of a nefarious agent, may result in more misconduct than that of an opaque enforcement setting. This is due to the endogenous enforcement externalities that may heighten misconduct. We also show that a principal with a longer decision horizon may actually induce higher misconduct. Our results have implications for enforcement transparency and a principal's decision horizon, and apply to various settings, including relations between headquarters and division managers, common owners and portfolio firm managers, and regulators and firms.

Keywords: Experimentation; Peer Learning; Reputation; Enforcement Externality; Enforcement Transparency

JEL Classification: D82, D83, G18, G31, G34, M41, M48

*We thank Mark Bagnoli, Tim Baldenius, Marco Battaglini, Jeremy Bertomeu, Oliver Boguth, Sam Bonsall, Dongkyu Chang, Jeffrey Coles, Jess Cornaggia, Kimberly Cornaggia, Jonathan Glover, Matthew Gustafson, Moritz Hiemann, Johannes Hörner, Steve Huddart, Peter Iliev, Yuhta Ishii, Justin Johnson, SP Kothari, Nicholas Lambert, Jorge Lemus, Stefan Lewellen, Fei Li, Lizhi Liu, Qingmin Liu, Christian Lundblad, Nadya Malenko, Andrew McClennan, Christine Parlour, Heikki Rantakari, Haresh Sapra, Timothy Simin, Joel Sobel, Bruno Strulovici, Shyam Sunder, Juuso Välimäki, Mihail Velikov, Shuyang Wang, Mao Ye, Yao Zeng, Amir Ziv, Alexey Zhdanov, and workshop or conference participants at Penn State University, Columbia Business School, Aalto University, 2022 Asian Meeting of Economic society, 2022 Conference on Mechanism and Institution Design, 2022 Queen's Organizational Economics Conference, and the Securities and Exchange Commission (SEC) for comments and suggestions. We benefited from discussions with several SEC staff, especially David Estabrook, Robert Fisher, Chauncey Martin, and Ryan Wolfe. We acknowledge financial support from our respective institutions.

†Johnson College of Business, Cornell University; Email: yc2535@cornell.edu.

‡Smeal College of Business, Penn State University; Email: kxd30@psu.edu.

§Tuck School of Business, Dartmouth College; Email: phillip.c.stocken@tuck.dartmouth.edu.

¶Smeal College of Business, Penn State University; Email: zxw192@psu.edu.

1 Introduction

We study a two-period game with a sequentially rational principal and self-interested agents whose misconduct imposes costs on the principal. For example, consider a capital budgeting setting featuring a CEO, who cares about firm value, and two self-interested division managers, who, given their empire building incentives, prefer to adopt business plans with negative net present value. The CEO can intervene and force a divisional manager to drop the plan and avoid the inefficient investment, albeit at a cost.[1] Given these intervention costs, the CEO may choose to intervene in a division's capital budgeting decision only if it is sufficiently detrimental to shareholder value. The division managers, who are uncertain about the CEO's intervention cost, may test the CEO's forbearance. The CEO, on the other hand, attempts to build a reputation for aggressive intervention. In this environment, would the CEO prefer to publicize the intervention in one division's capital budgeting to "make an example" of the nefarious manager or to keep the intervention secret? In addition, would the CEO be more successful in deterring divisional empire-building when the CEO has a longer decision horizon (e.g., when more likely to remain in the position)?

Analogous questions can arise in other settings. For example, consider firms that engage in value-destroying activities and an institutional shareholder (e.g., a venture capital fund, a private equity fund, or a mutual fund) that is the common owner of the firms. In this scenario, how would a requirement to publicize an intervention affect firm misconduct? Would it benefit or harm the common owner? In addition, would a common owner with a longer-term investment strategy be more effective than one with a short-term strategy at deterring individual firm managers' value-destroying behavior? As another example, consider firms that maximize their own profits at the expense of the social welfare and a regulator who is concerned about this misconduct (e.g., violations of securities and data protection regulations). Would a more transparent enforcement regime reduce or increase firm miscon-

[1] The intervention cost may include the administrative effort to obtain evidence of misconduct, the communication cost, or the damage to the relationship between the headquarters and the divisions.

duct?[2] Would a regulator with a longer term (or similarly, a higher likelihood of remaining in office) be more capable of deterring regulated firms' misconduct?

To this end, consider a two-period game with a sequentially rational and forward-looking principal and two self-interested agents who may misbehave at the principal's expense. There are two periods. In each period, the agents simultaneously choose their levels of misconduct. After observing their misconduct, the principal decides whether to discipline them. At the end of the first period, the agents' misconduct levels are publicly observed. If the principal tolerates an agent's misconduct, then the agent enjoys a private benefit equal to the misconduct and the principal bears damages depending on its sensitivity to misconduct. Alternatively, if the principal disciplines an agent, then the agent's misconduct is corrected, eliminating any private benefit to the agent and damages to the principal. Disciplining an agent, however, is costly, as it imposes an enforcement cost on the principal and a penalty on the agent. The principal's willingness to discipline the agents depends on the enforcement cost and the misconduct damage sensitivity, and this propensity is privately known to the principal while the agents only hold a common prior belief.

We consider two scenarios about the principal's enforcement choices. In the benchmark, the principal's enforcement choices are only known to the disciplined agent, and in the main model, they are publicized. The driving forces in the opaque enforcement benchmark are equivalent to those in a one-agent-one-principal model. An agent's optimal misconduct decision depends on the intra-period trade-off between the potential penalty from enforcement and the private benefit from the misconduct if it is tolerated. Moreover, while more egregious misconduct increases the likelihood of enforcement and the agent's private benefit, it provides an option value to the agent because the agent can condition its second-period misconduct on the information revealed by the first period's enforcement outcome. If the principal tolerates

[2]The transparency of enforcement and adjudication has had a long legislative history and the focus of a recent executive order requiring greater transparency. In some regulatory domains, enforcement transparency has also been the subject of lawmaking and debates. For example, the Public Company Accounting Oversight Board (PCAOB) Enforcement Transparency Act of 2017 stipulates the transparency of a principal's enforcement actions, making PCAOB disciplinary proceedings open to the public unless the board, on its own motion or after considering the motion of a party, orders otherwise. See https://www.congress.gov/bill/115th-congress/senate-bill/610. In bank regulation, the 1989 Financial Institutions Reform, Recovery, and Enforcement Act requires that enforcement decisions be publicly disclosed. In contrast, transparency in the setting of bank stress testing has been debated in academia, among principals, and among practitioners (e.g., Goldstein and Sapra, 2013; Goldstein and Leitner, 2020).

high first-period misconduct, then its enforcement propensity is revealed to be weak. The agent will then choose a high misconduct level in the second period. Motivated by the option value of learning, each agent experiments with more aggressive first-period misconduct.

In the main model where enforcement choices are publicly observed, the two agents' intra-period trade-offs and the experimentation motives still exist. Moreover, they interact because anticipating learning from its peer's misconduct and enforcement outcome, an agent will free ride on its peer's experimentation and reduce its own misconduct. This is the canonical *positive information externality* explored in the extant collective experimentation literature.

Simultaneously, the principal, who is *strategic*, has a reputation concern. In the first period, the principal takes into account the costs and benefits of enforcement within that period, but it also cares about its reputation that the current enforcement signals, which affects future misconduct. This reputation concern leads to excessively aggressive enforcement in the first period, which discourages misconduct.

This is not the end of story, however, as the strategic principal is aware that its enforcement choices are publicly observed. A strategic principal together with transparent enforcement results in the first-period enforcement strategy against one agent depending on the misconduct of *both* agents. The interdependence of the enforcement strategy creates an endogenous *enforcement externality*, as one agent's misconduct affects its peer's chance of being disciplined.

The reason the enforcement strategy is interdependent across agents is that the principal's cost to mimic being intolerant of misconduct depends on the *proximity* of the agents' misconduct levels in the first period. If they are close, then it is more costly for the principal to mimic being intolerant as it must punish *both* agents. Disciplining only the agent with slightly more egregious misconduct while tolerating the other's results in an insufficient reputation boost, as the agents will infer that the principal's enforcement capacity is such that it will correct the more serious misconduct but not enough that it will correct the other, which is only slightly lower. In contrast, if one agent's misconduct is much higher than that of the other, then disciplining only the agent with the much higher misconduct still enables the principal to pool with stronger types and mimic being tough.

The enforcement externality that results from the principal being strategic and the enforcement being transparent can be positive or negative. It is *positive* if an agent's higher misconduct benefits its peer by reducing the peer's enforcement risk; this effect *discourages* an agent from engaging in misconduct as it shields behind its peer's heightened misconduct. Alternatively, it is *negative* if an agent's higher misconduct harms its peer by exposing the peer to higher enforcement risk; this effect *encourages* an agent to increase its misconduct because part of its heightened enforcement risk is shared with its peer.

We solve for the unique pure-strategy, perfect Bayesian equilibrium. In the opaque enforcement benchmark, the principal punishes an agent if and only if the agent's misconduct is sufficiently high. That is, the enforcement decision on one agent only depends on the agent's own misconduct. Each agent chooses a strictly positive and identical misconduct level in the first period. The first-period misconduct leads to an even higher second-period misconduct level when it is tolerated and zero second-period misconduct otherwise.

In the main model with transparent enforcement, the equilibria are starkly different depending on whether the principal is *impatient* (i.e., discount factor less than 1/2) or *patient* (i.e., discount factor greater than or equal to 1/2).

- For an impatient principal, the equilibrium enforcement strategy in the first period is to discipline both agents if both agents' misconduct levels are high, to discipline only one agent if the agent's misconduct is much higher than that of the other, and to discipline neither if both agents' misconduct levels are low. In equilibrium, the agents' misconduct levels fall in a region where the enforcement externality is *positive*. Together with the positive information externality, the equilibrium misconduct is squashed, and it may even be eliminated in both periods.

- For a patient principal, the equilibrium enforcement strategy in the first period is to discipline either both agents or neither, but never only one agent. In equilibrium, the agents' misconduct levels fall in a region where enforcement externality is *negative*, which works against the positive information externality, and it yields heightened misconduct.

The analysis has two key sets of implications. First, we show the effect of enforcement transparency (and the resultant peer learning) on the agents' expected misconduct levels and the principal's expected total cost. The effect depends on the principal's decision horizon.

- When the principal is impatient, enforcement transparency discourages misconduct and reduces the principal's cost, because it induces the agents to free ride on each other's misconduct. In another word, enforcement transparency deters misconduct as it allows the principal to "make an example" of a nefarious agent. This result echos the conclusions of the traditional collective experimentation literature (Bolton and Harris, 1999; Keller et al., 2005) that finds information spillovers reduce or slow down experimentation.

- When the principal is patient, however, the result is reversed. The enforcement externality turns negative, and it dominates the information externality. Accordingly, enforcement transparency encourages misconduct, and it increases the principal's cost.

Second, we study the effect of the principal's decision horizon on the expected misconduct level. We show that the effect depends on whether enforcement outcomes are publicly observed.

- When enforcement is opaque, misconduct levels in both periods decrease as the principal's discount factor increases.

- When enforcement is transparent, the misconduct levels are non-monotonic in the principal's decision horizon. Within certain parameter ranges, the principal's longer decision horizon actually may lead to higher misconduct levels in both periods.

Intuitively, a longer decision horizon, or equivalently, a stronger incentive to build a reputation for being intolerant of misconduct, has two countervailing effects on the agents' misconduct. On one hand, the principal is more intolerant in the first period to develop a reputation for being tough. This effect, which is the only effect present when enforcement is opaque, deters misconduct. On the other hand, a patient principal disciplines two agents simultaneously and never only one agent. As an agent cannot shield behind its peer's misconduct, the agent engages in more misconduct.

Related Literature There is a large literature on experimentation (Rothschild (1974); Aghion, Bolton, Harris, and Jullien (1991); Manso (2011); Bergemann and Hege (2005); Nanda and Rhodes-Kropf (2017)). See Bergemann and Välimäki (2006) and Hörner and Skrzypacz (2016) for surveys of this literature. Among them, particularly related is the collective experimentation literature where players learn from each other. The literature typically assumes that the object about which the players learn, or the "bandit arm," is *non-strategic*, and it finds that information externalities among players cause free-riding and thereby reduce or slow down learning (e.g., Bolton and Harris (1999); Keller, Rady, and Cripps (2005); Bonatti and Hörner (2011); Chen (2020)). In our paper, in contrast, agents learn about a *strategic* principal, who in turn optimally manages its reputation. We find that peer learning among agents not only results in free-riding, but also creates an endogenous enforcement externality. For a sufficiently patient principal, we show that this new effect dominates the free-riding effect and raises the equilibrium experimentation.

Bergemann and Välimäki (2000) is among the few papers that study collective learning against strategic players. They consider a *two-sided* learning game, where both sellers and buyers seek to learn the unknown quality of a new product. The buyers' purchases of the product yield information about its true quality, and the sellers strategically set prices to influence this learning. They show that the combination of the informational externalities among the buyers and the strategic pricing by the sellers results in excessive experimentation. In contrast, our paper considers a *one-sided* learning model, where agents experiment with misconduct levels to test the principal's tolerance and the principal interferes with the agents' learning through strategic enforcement. We find that the principal's decision horizon crucially determines whether peer learning increases experimentation.

Our results on the effect of enforcement transparency and principal's decision horizon shed light on questions in various institutional settings. Our paper is related to the literature on capital budgeting (Harris and Raviv (1998); Bernardo, Cai, and Luo (2001); Malenko (2019)), institutional common ownership (Edmans, Levit, and Reilly (2019); Gilje, Gormley, and Levit (2020)), short-termism of institutional investors (Dasgupta and Piacentino (2015); Song (2019); Burkart and Dasgupta (2021)), shareholder activism spillovers and the threat of shareholder activism (Lee and Park (2009); Gantchev, Gredil, and Jotikasthira (2019); Lakkis

(2021)), regulatory reputation management (Boot and Thakor (1993); Morrison and White (2013); Shapiro and Skeie (2015); Huang (2017)), regulatory enforcement transparency (see Goldstein and Sapra (2013) and Goldstein and Leitner (2020) for surveys of the literature that focuses on banking regulations), and crime (Bond and Hagerty (2010)). See Section 6 for more detailed discussions of our connections to these papers. Our paper differs from this antecedent literature, as it explores a novel channel through which the reputation-building motive of a strategic principal interacts with peer-learning among agents.

The paper proceeds as follows. Section 2 introduces the model. Section 3 analyzes a benchmark environment where enforcement is opaque. Section 4 characterizes the equilibrium and derives implications when enforcement is transparent. Section 5 extends the analysis to wider parameter ranges. Section 6 discusses the implications for various institutional settings. Section 7 concludes. All proofs are relegated to the Appendix.

2 Model

We study a model of collective experimentation by two self-interested agents facing a strategic principal with private information. The game has two periods. In each period $t = 1, 2$, agents $i = A, B$ simultaneously choose the levels of *misconduct*, $x_t^i \geqslant 0$. After observing x_t^A and x_t^B, the principal makes *enforcement* decisions $e_t^i \in \{0, 1\}$ with respect to agent $i = A, B$, such that $e_t^i = 1$ denotes the decision to discipline agent i in period t and $e_t^i = 0$ indicates otherwise. If agent i's misconduct x_t^i is tolerated ($e_t^i = 0$), the agent derives a private benefit of x_t^i while the principal internalizes a misconduct damage of kx_t^i, where $k > 0$ is interpreted as the principal's misconduct damage sensitivity. Instead, if the agent is disciplined ($e_t^i = 1$), the agent's private benefit is eliminated and the principal avoids the misconduct damage. Additionally, the agent suffers a *penalty* $L > 0$ and the principal incurs an *enforcement cost* $c \geqslant 0$.

The enforcement cost c and misconduct damage sensitivity k are the principal's private information, which are fixed across periods. At the beginning of period 1, agents hold a common prior belief about the joint distribution of (c, k). After period 1, both agents' misconduct levels (x_1^A, x_1^B) are observable. In the benchmark section (Section 3), the principal's

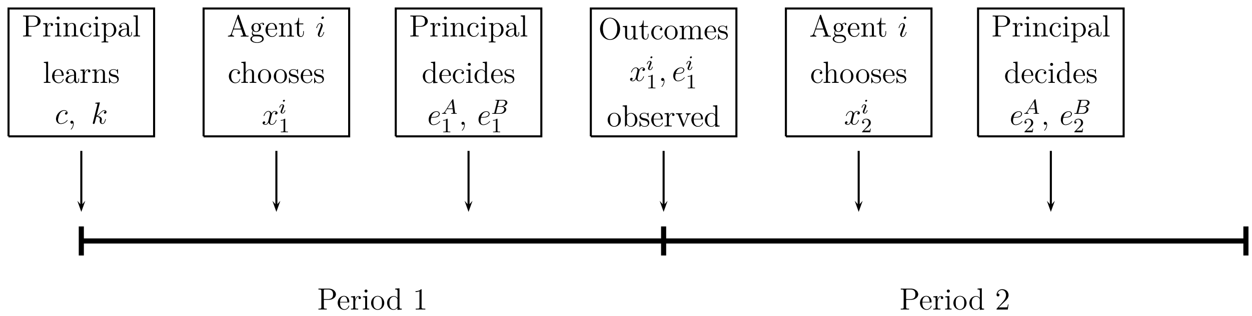


Figure 1: Timeline of the two-period game

enforcement decisions on one agent is not observed by the other agent, while in the main model (Section 4), we allow the principal's enforcement decisions (e_1^A, e_1^B) to be publicly observable. After observing the misconduct levels and possibly the principal's enforcement decisions, the agents enter period 2 with updated beliefs using Bayes' rule. The timeline is summarized in Figure 1.

All players are risk neutral. Given the realized misconduct levels x_t^i and enforcement decisions e_t^i, $t = 1, 2$, agent i enjoys a *total discounted payoff* across both periods of

$$U^i \equiv \left[x_1^i(1 - e_1^i) - Le_1^i\right] + \delta\left[x_2^i(1 - e_2^i) - Le_2^i\right],$$

where $\delta \in (0, 1)$ is the agents' common discount factor. The principal incurs a *total discounted cost* across both periods and from both agents of

$$\begin{aligned} C &\equiv \sum_{i=A,B}\Big(\left[kx_1^i(1 - e_1^i) + ce_1^i\right] + \delta_P\left[kx_2^i(1 - e_2^i) + ce_2^i\right]\Big) \\ &= k\sum_{i=A,B}\Big(\left[x_1^i(1 - e_1^i) + \theta e_1^i\right] + \delta_P\left[x_2^i(1 - e_2^i) + \theta e_2^i\right]\Big), \end{aligned}$$

where $\delta_P \in (0, 1)$ is the principal's discount factor, and

$$\theta \equiv \frac{c}{k} \geqslant 0$$

is interpreted as the *normalized enforcement cost*. This is because the misconduct damage sensitivity k enters the principal's cost only as a global multiplier, and hence the parameters are strategically equivalent to a normalized one with $k = 1$ and $c = \theta$, where the variable θ summarizes the principal's private information relevant for decision-making. Hereafter, we let $k = 1$ for notational convenience, and refer to θ as the principal's *type*. A higher θ corresponding to a lower enforcement propensity. Agents form common prior in the beginning of period 1 that θ is continuously distributed with c.d.f. F and p.d.f. f with support $[0, \infty)$. A higher discount factor δ or δ_P is interpreted as the agents or the principal, respectively, being patient or having a longer decision horizon. We require the parameters to satisfy the following conditions.

Assumption 1 **(i)** *The hazard rate* $h(\theta) \equiv \frac{f(\theta)}{1-F(\theta)}$ *is continuously increasing in* θ.
(ii) *The penalty is moderate:* $1 \leqslant Lh(0) < 1 + \delta - \delta_P$.

Part (i) is a mildly restrictive monotone hazard rate property. Part (ii) requires the penalty to be large enough (i.e., $L \geqslant 1/h(0)$), such that in the last period, if the posterior belief is the same as the prior belief, then an agent i will choose zero misconduct to avoid any enforcement risk.[3] On the other hand, the penalty is sufficiently low (i.e., $L < (1 + \delta - \delta_P)/h(0)$), such that the agents will engage in some positive misconduct in period 1.[4]

With a slight abuse of notation, the strategy of agent i is a pair (x_1^i, x_2^i), where $x_1^i \in [0, \infty)$ and x_2^i maps $\{x_1^j, e_1^j\}_{j=A,B}$ into $[0, \infty)$. The principal's strategy is a mapping from $\theta \geqslant 0$ to a quadruple $\{e_t^j\}_{j=A,B,t=1,2}$, where e_1^j maps $\{x_1^j\}_{j=A,B}$ into $\{0, 1\}$ and e_2^j maps $\{x_1^j, e_1^j, x_2^j\}_{j=A,B}$ into $\{0, 1\}$. We impose the following technical constraint on the principal's strategy:

Assumption 2 *For any principal type* θ, $t = 1, 2$ *and* $j = A, B$, *enforcement decision* e_t^j *is lower semi-continuous in* (x_t^A, x_t^B).

Thus, if the principal is indifferent between disciplining an agent and not, it breaks the tie by not disciplining. This ensures that agents' best responses are always well-defined.[5]

[3] Part (i) and (ii) together imply that in the second period, each agent chooses the highest misconduct level known to avoid any enforcement risk, as is established in Lemma 1. We show numerically in Section 5.2 that the key driving forces of the model remain unchanged if this assumption is relaxed such that agents may choose a period-2 misconduct level that entails some enforcement risk.

[4] Note that (ii) implies $\delta_P < \delta$, that is, the principal is more impatient than the agents. Otherwise, the principal always punishes any agent with positive misconduct in period 1 to signal its type, and foreseeing this, agents are deterred from any misconduct in period 1.

[5] This assumption is without loss of generality because an agent can always reduce its misconduct by ε to avoid enforcement.

We characterize the perfect Bayesian equilibrium (PBE). A PBE is a strategy profile $\{x_t^j, e_t^j\}_{j=A,B,t=1,2}$ along with agent i's posterior belief $q^i(x_1^A, x_1^B, e_1^A, e_1^B) \in \Delta([0,\infty))$ at the beginning of period 2, $i = A, B$, such that: (a) agent i's strategy (x_1^i, x_2^i) maximizes its expected total discounted payoff $\mathbb{E}(U^i)$ given the strategies of agent $-i$ and the principal; (b) the principal's strategy $\{e_t^j\}_{j=A,B,t=1,2}$ minimizes its expected total discounted cost $\mathbb{E}(C)$ given the agents' strategies; (c) the agent i's posterior belief $q^i(x_1^A, x_1^B, e_1^A, e_1^B)$ is consistent with the strategies and Bayes' rule whenever possible.

3 Benchmark: Opaque Enforcement

Consider a benchmark setting in which agent i cannot observe the enforcement outcome e_1^{-i} of its peer agent $-i$. In other words, we digress to an *opaque* setting in this benchmark to shut down peer learning between the agents. Thus, while each agent can still observe its peer's misconduct in period 1, without observing the enforcement outcome of its peer, the agent learns no payoff-relevant information from its peer. Therefore, each agent only relies on its own period-1 misconduct and its own enforcement outcome to learn the principal's type θ. As a result, the game with opaque enforcement is separable between the agents, and hence, it is equivalent to two replications of a *one-agent-one-principal* game.

Using backward induction, we begin by understanding the continuation game in period 2. The following lemma characterizes the equilibrium strategies of the three parties.

Lemma 1 (Period-2 Equilibrium)

In period 2, for $i = A, B$:

(i) *The principal disciplines agent i ($e_2^i = 1$) if and only if $x_2^i > \theta$.*

(ii) *Agent i chooses x_2^i to maximize $x_2^i \Pr(x_2^i \leqslant \theta; q^i) - L \Pr(x_2^i > \theta; q^i)$, where q^i is agent i's posterior belief at the beginning of period 2. In particular, if q^i is the posterior belief with truncated support $[\theta_L, \theta_H] \in [0,\infty)$, then $x_2^i = \theta_L$.*

As period 2 is the final period, all players behave myopically. Part (i) of Lemma 1 states that the principal compares the misconduct damage x_2^i to the enforcement cost θ, without considering its reputation. Analogously, part (ii) indicates that an agent derives no further option value from learning, and therefore, it maximizes its within-period payoff. Further, part (ii) states that an agent chooses the lower bound of the support of its posterior belief about the principal's type, thereby avoiding any enforcement risk. That is, the agent chooses the highest misconduct that is known to be "safe." This property is ensured by Assumption 1, which constrains the penalty L to be sufficiently large.

Next, we turn to period 1. Suppose agent $i \in \{A, B\}$ chooses misconduct x_1^i. If the principal disciplines this agent, its total discounted cost associated with agent i alone is

$$\theta + \delta_P \min\{\theta, x_2^i(x_1^i, e_1^i = 1)\}, \tag{1}$$

that is, it incurs enforcement cost θ in period 1 and the continuation cost due to enforcement or misconduct in period 2, whichever is lower. If, instead, it does not discipline agent i, its total cost associated with agent i is

$$x_1^i + \delta_P \min\{\theta, x_2^i(x_1^i, e_1^i = 0)\}, \tag{2}$$

that is, it incurs misconduct damage x_1^i in period 1 and the ensuing continuation cost in period 2. The principal disciplines agent i in period 1 if and only if (1) is less than (2).

Observe that the principal's total cost displays the "single-crossing" property, regardless of the agent's period-2 strategy. That is, the principal is less willing to discipline an agent with misconduct x_1^i in period 1 if its normalized enforcement cost θ is higher.[6] Consequently, the principal follows a *cutoff* strategy such that an agent is disciplined if and only if θ is strictly lower than a cutoff type, denoted as $\theta^\dagger$.

Given agent i's period-1 misconduct x_1^i, we determine the equilibrium cutoff type $\theta^\dagger$ that is indifferent between disciplining the agent and not. When the type-$\theta^\dagger$ principal disciplines the agent, (1) reduces to $\theta^\dagger + \delta_P \cdot 0$. This follows because the agent chooses $x_2^i = 0$ to avoid any enforcement risk when the principal's type is inferred to be in $[0, \theta^\dagger)$, according to Lemma 1 (ii). When the agent is not disciplined, (2) simplifies to $x_1^i + \delta_P \cdot \theta^\dagger$, because with

[6] To see this, observe that the derivative of (1) with respect to θ is no lower than 1, while the derivative of (2) with respect to θ is no higher than δ_P. Thus, the former derivative is larger than the latter, implying that the difference between (1) and (2) increases in θ, i.e., the "single-crossing" property holds. This argument is true for any $x_2^i(\cdot, \cdot)$.

the principal's type inferred to be in $[\theta^{\dagger}, \infty)$, the agent chooses the highest "safe" level of misconduct $x_2^i = \theta^{\dagger}$. The cutoff type $\theta^{\dagger}$ can be solved as a function of x_1^i from the indifference between (1) and (2), yielding

$$\theta^{\dagger} = \phi(x_1^i) \equiv \frac{x_1^i}{1 - \delta_P}. \tag{3}$$

The principal disciplines agent i in period 1 if and only if $\theta < \theta^{\dagger}$, or equivalently,

$$x_1^i > \theta(1 - \delta_P). \tag{4}$$

Understanding the principal's strategy, agent i foresees the risk associated with each level of period-1 misconduct and optimally chooses x_1^i to maximize the expected payoff

$$\max_{x_1^i} \left(x_1^i + \delta \cdot \phi\left(x_1^i\right)\right)\left(1 - F\left(\phi\left(x_1^i\right)\right)\right) + (-L) \cdot F\left(\phi\left(x_1^i\right)\right). \tag{5}$$

Observe, with probability $1 - F\left(\phi\left(x_1^i\right)\right)$, the principal's type is higher than the cutoff type $\theta^{\dagger} = \phi(x_1^i)$, and the agent's period-1 misconduct is tolerated. The agent enjoys a private benefit of x_1^i in period 1, while choosing period-2 misconduct x_2^i to be the highest safe choice, $\phi(x_1^i)$. With the complementary probability, the principal's type is lower than the cutoff type $\phi(x_1^i)$, and the principal punishes the agent in period 1. The agent suffers the penalty L in period 1, while choosing $x_2^i = 0$ in period 2 to avoid the possibility of enforcement. The first-order condition with respect to x_1^i gives the equilibrium period-1 misconduct. The next proposition characterizes the equilibrium when enforcement is opaque.

Proposition 1 (Equilibrium: Opaque Enforcement)
Suppose agent $-i$, $i = A, B$, $t = 1, 2$, does not observe e_t^i. Let ϕ be defined as in (3). In the unique perfect Bayesian equilibrium, the type-θ principal disciplines an agent with x_1^i in period 1 if and only if $\theta < \phi(x_1^i)$ (or equivalently, $x_1^i > \theta(1 - \delta_P)$). Agent i's period-1 misconduct is $x_1^i = x^*$, where $x^* > 0$ uniquely solves

$$\left(L + (1 + \delta - \delta_P)\phi(x^*)\right) h\left(\phi(x^*)\right) = 1 + \delta - \delta_P. \tag{6}$$

In addition, $x_2^i = 0$ *if* $e_1^i = 1$*, and* $x_2^i = \phi(x^*)$ *if* $e_1^i = 0$*.*

Each agent chooses a strictly positive and identical action x^* in the first period. The first-period misconduct leads to an even higher second-period misconduct level $\phi(x_1^i)$ if the principal tolerates it and zero second-period misconduct otherwise. The trade-off an agent faces when choosing the period-1 misconduct is as follows. With higher period-1 misconduct, the agent faces an intra-period trade-off between a higher private benefit conditional on the principal's forbearance and a higher likelihood of being disciplined. In addition, higher period-1 misconduct enhances the option value of learning in that if the higher misconduct is tolerated in period 1, the principal's type θ is revealed to be even higher, allowing the agent to raise its period-2 misconduct without being disciplined. As (5) reflects, a higher x_1^i leads to a higher cutoff type $\phi(x_1^i)$, and a higher period-2 misconduct when x_1^i is tolerated. Thus, period-1 misconduct serves as an *experiment* to probe the principal's enforcement propensity.

The principal, on the other hand, when anticipating an agent's experimentation, disciplines an agent aggressively in period 1 to build a reputation for stringent enforcement. This tendency is reflected in (4) as the principal of type θ disciplines agent i with misconduct as low as $\theta(1-\delta_P)$, which is below the misconduct level that a myopic principal with $\delta_P = 0$ would discipline.

The principal's period-1 enforcement strategy is *independent* across agents, that is, the decision whether to discipline one agent depends only on that agent's period-1 misconduct and not on its peer agent's misconduct. In other words, the increase in an agent's misconduct does not affect the peer agent's risk of being disciplined.

Finally, we investigate whether a principal with a longer decision horizon is better at deterring misconduct and reducing the costs that are imposed on the principal. Intuitively, when the principal is more patient, it cares more about its period-2 payoff. Accordingly, it adopts a more stringent enforcement criterion in the first period to build a reputation for tough enforcement. Anticipating this more aggressive enforcement, the agents reduce their misconduct. The next corollary formalizes this result.

Corollary 1 (Effect of δ_P: Opaque Enforcement)

A higher δ_P strictly reduces the agents' misconduct in both periods.

4 Main Model: Transparent Enforcement

We turn to consider the main model where enforcement is transparent. That is, the enforcement outcomes (e_t^A, e_t^B) are observable to both agents after period $t = 1, 2$. Like the benchmark, an agent learns about the principal's type from its own misconduct and enforcement outcome, and the principal tries to build a reputation for stringent enforcement. Unlike the benchmark, however, enforcement transparency introduces *externalities* between the two agents. First, an agent can free ride on the information revealed by its peer's experimentation. This peer learning leads to a positive *information externality* between the two agents. Second, the principal's punishment of one agent may depend on the other agent's misconduct, which generates *enforcement externalities* between the two agents.

Using backward induction, we start with period 2. The equilibrium characterization in Lemma 1 continues to apply to the main model because all three players act myopically in period 2, regardless of the enforcement transparency.

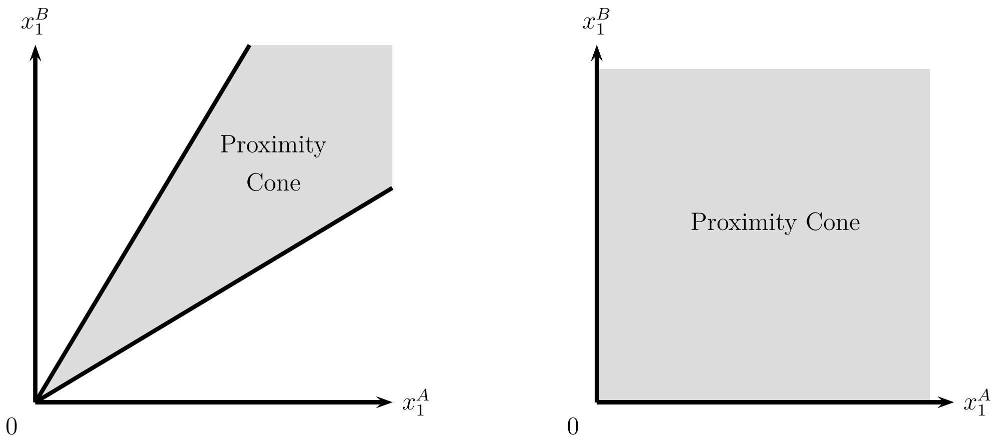
4.1 Period 1: Principal's Enforcement

Next, we consider the first period. We start with the equilibrium strategy of the principal and characterize properties of the principal's strategy that necessarily arise in equilibrium.

Lemma 2 (i) (Monotonicity) *In period 1, if $x_1^i > x_1^j$ and $e_1^j = 1$, then $e_1^i = 1$.*
(ii) (Cutoff Strategy) *Given any realized misconduct profile (x_1^A, x_1^B), there exist cutoff types $0 \leqslant \theta^{**} \leqslant \theta^* < \infty$ such that the principal disciplines both agents when $\theta < \theta^{**}$, disciplines only one agent when $\theta^{**} \leqslant \theta < \theta^*$, and disciplines neither when $\theta \geqslant \theta^*$.*

Part (i) implies that if the principal disciplines only one agent, then it must be the one with the higher level of misconduct. Part (ii) shows that the principal follows a *cutoff* enforcement strategy that depends on its type θ, echoing the counterpart in the benchmark.[7]

[7]As in the benchmark, the fact that the principal follows a cutoff enforcement strategy does not depend on the agents' second period strategy.



(a) $\delta_P < \frac{1}{2}$ (b) $\delta_P \geqslant \frac{1}{2}$

Figure 2: Proximity cone for various δ_P.

As the principal's type θ increases, its enforcement strategy becomes increasingly tolerant: it disciplines both agents when θ is low, one agent when θ is moderate, and neither agent when θ is high.

Although the principal's cutoff strategy defines three cases, some cases may disappear as (x_1^A, x_1^B) varies. For example, it is possible that $\theta^{**} = \theta^*$, in which case a principal never finds it optimal to discipline only one agent. In this degenerate case, the principal disciplines both agents if its type is low and disciplines neither otherwise. This degenerate case arises when the agents' period-1 misconduct x_1^A and x_1^B are *close* together. Formally, we define a subset of the $x_1^A - x_1^B$ plane as the *proximity cone* such that when the realized levels of misconduct (x_1^A, x_1^B) fall inside this cone, the principal never finds it optimal to punish only one agent.

Definition 1 (Proximity Cone)

For $\delta_P \in (0, 1)$*, the* proximity cone *is the set*

$$P(\delta_P) \equiv \left\{(x_1^A, x_1^B) \in [0, \infty) \times [0, \infty) : \ x_1^B \geqslant x_1^A(1 - 2\delta_P), \ x_1^A \geqslant x_1^B(1 - 2\delta_P)\right\}.$$

Figure 2(a) illustrates the proximity cone for $\delta_P < \frac{1}{2}$, which is the area sandwiched between the two rays $x_1^B = x_1^A(1 - 2\delta_P)$ and $x_1^B = \frac{x_1^A}{1-2\delta_P}$. The levels of misconduct are close to each other if (x_1^A, x_1^B) is in the proximity cone. Figure 2(b) shows that the proximity cone for $\delta_P \geqslant \frac{1}{2}$ expands to the entire quadrant.

In this light, we characterize the period-1 strategy for a relatively impatient principal with $\delta_P < \frac{1}{2}$ and then for a relatively patient principal with $\delta_P \geqslant \frac{1}{2}$. The principal's strategies use the following definitions:

$$\phi_{1|2}(x_1^A, x_1^B) = \frac{\min\{x_1^A, x_1^B\}}{1 - 2\delta_P}, \tag{7}$$

$$\phi_{0|1}(x_1^A, x_1^B) = \frac{\max\{x_1^A, x_1^B\}(1 - 2\delta_P) - \min\{x_1^A, x_1^B\}(2\delta_P)}{(1 - 2\delta_P)^2}, \tag{8}$$

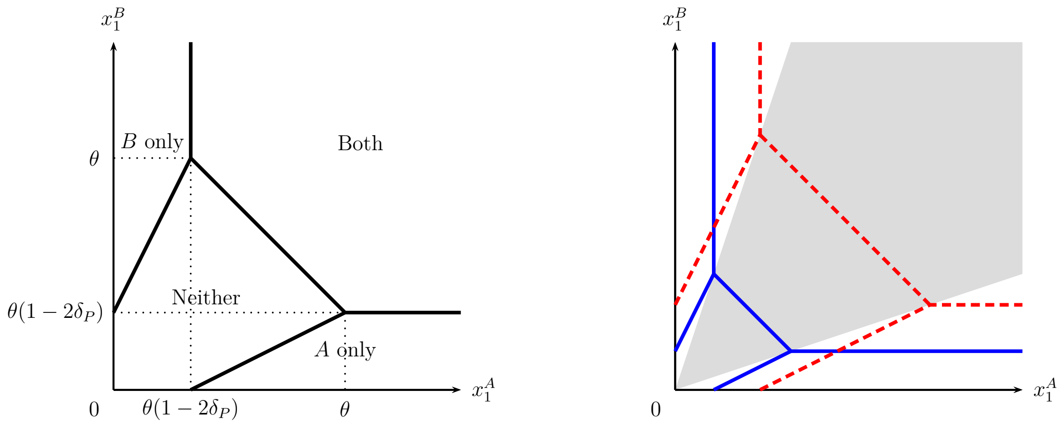
$$\phi_{0|2}(x_1^A, x_1^B) = \frac{x_1^A + x_1^B}{2(1 - \delta_P)}. \tag{9}$$

Given (x_1^A, x_1^B), $\theta^{**} = \phi_{1|2}(x_1^A, x_1^B)$ is the cutoff type of the principal that is indifferent about disciplining one or two agents, and $\theta^* = \phi_{0|1}(x_1^A, x_1^B)$ is cutoff type indifferent about disciplining neither or only one agent, provided that $\theta^* > \theta^{**}$. When the principal never disciplines only one agent, we have $\theta^{**} = \theta^* = \phi_{0|2}(x_1^A, x_1^B)$ as the cutoff type of the principal that is indifferent between disciplining neither or both agents.

4.1.1 Impatient Principal

The next proposition characterizes the principal's period-1 enforcement strategy when the principal is *impatient*, i.e., $\delta_P < \frac{1}{2}$. It shows that when (x_1^A, x_1^B) is within the proximity cone, the principal chooses only between disciplining both agents and disciplining neither. When (x_1^A, x_1^B) is outside of the proximity cone, however, the principal chooses between disciplining both agents, only one agent, or neither agent.

Proposition 2 (Principal Strategy in Period 1: $\delta_P < \frac{1}{2}$)
Suppose $\delta_P < \frac{1}{2}$. Let $\phi_{1|2}$, $\phi_{0|1}$, and $\phi_{0|2}$ be defined as in (7)-(9)*, respectively.*
(i) *If (x_1^A, x_1^B) is in the proximity cone, then the principal disciplines both agents when $\theta < \phi_{0|2}(x_1^A, x_1^B)$ and disciplines neither otherwise.*
(ii) *If (x_1^A, x_1^B) is not in the proximity cone, then the principal disciplines both agents when $\theta < \phi_{1|2}(x_1^A, x_1^B)$, disciplines one agent when $\phi_{1|2}(x_1^A, x_1^B) \leqslant \theta < \phi_{0|1}(x_1^A, x_1^B)$, and disciplines neither when $\theta \geqslant \phi_{0|1}(x_1^A, x_1^B)$.*



(a) Enforcement zones: fixed θ

(b) Enforcement zones: varying θ

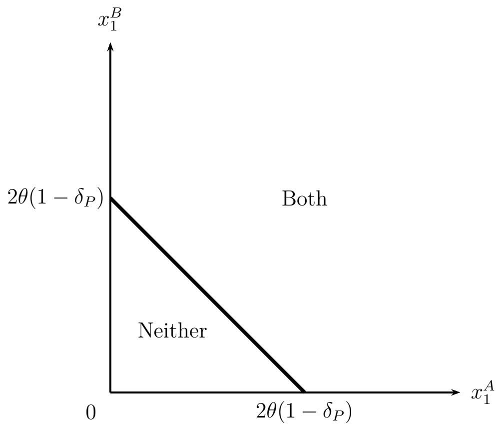
Figure 3: Principal's enforcement zones in period 1, for $\delta_P < \frac{1}{2}$. (a): four distinct enforcement zones for a fixed θ. (b): solid (resp. dashed) lines are the boundaries of the enforcement zones for a low (resp. high) type, and the shaded area is the proximity cone.

Proposition 2 describes the principal's enforcement decision when fixing (x_1^A, x_1^B). Alternatively, Corollary 2 characterizes the enforcement decision when fixing type $\theta \geqslant 0$.

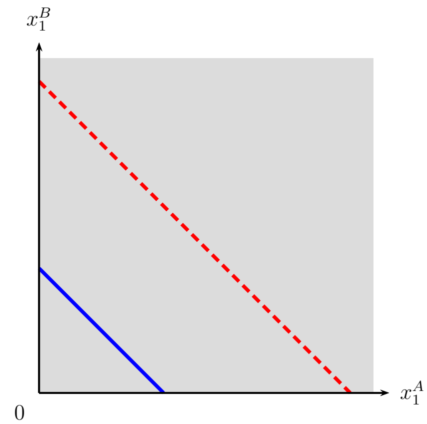
Corollary 2 (Enforcement Zones of a Fixed Type)
Fix $\delta_P < \frac{1}{2}$ and principal type $\theta \geqslant 0$. In period 1, the principal disciplines both agents if the profile (x_1^A, x_1^B) is such that $\phi_{0|2}(x_1^A, x_1^B) > \theta$ and $\phi_{1|2}(x_1^A, x_1^B) > \theta$; it disciplines only one agent if $\phi_{1|2}(x_1^A, x_1^B) \leqslant \theta$ and $\phi_{0|1}(x_1^A, x_1^B) > \theta$; it disciplines neither agent if $\phi_{0|2}(x_1^A, x_1^B) \leqslant \theta$ and $\phi_{0|1}(x_1^A, x_1^B) \leqslant \theta$.

Figure 3(a) illustrates the four distinct enforcement zones. The solid lines are the boundaries between adjoining zones, and the principal is indifferent between adjacent decisions at these boundaries. The type-θ principal disciplines both agents if both misconduct levels and their sum are all sufficiently high. It disciplines only one agent if its misconduct is much higher than its peer's, and disciplines neither agent if both levels are low.

The enforcement zones vary with the principal's type θ. In Figure 3(b), the solid lines are the boundaries for a lower type, while the dashed lines are those for a higher type. The corresponding boundaries for different types are parallel to each other. The higher type is more "lenient" in terms of enforcement zones than the lower type because for any profile (x_1^A, x_1^B), the higher type disciplines weakly fewer agents than the lower type. This



(a) Enforcement zones: fixed θ



(b) Enforcement zones: varying θ

Figure 4: Principal's enforcement zones in Period 1 for $\delta_P \geqslant \frac{1}{2}$. (a): two distinct enforcement zones for a fixed θ. (b): the solid (resp. dashed) line is the boundary of the enforcement zones for a low (resp. high) type, and the shaded area is the proximity cone.

is consistent with the cutoff structure in Lemma 2. Notably, by varying θ, the boundary between punishing neither and both firms traces out the shaded area, which is the proximity zone. When (x_1^A, x_1^B) is in the proximity cone, either both agents are disciplined (by lower types), or neither agent is disciplined (by higher types), but no principal type θ will discipline one agent only.

4.1.2 Patient Principal

We now turn to a *patient* principal, defined by $\delta_P \geqslant \frac{1}{2}$. In this case, the proximity cone expands to the entire space (see Figure 2(b)), reflecting that the principal disciplines both agents, if any at all, in period 1.[8] The next proposition characterizes the principal's period-1 strategy for a patient principal.

Proposition 3 (Principal Strategy in Period 1: $\delta_P \geqslant \frac{1}{2}$)
Suppose $\delta_P \geqslant \frac{1}{2}$. Let $\phi_{0|2}$ be defined as in (9). The principal disciplines both agents when $\theta < \phi_{0|2}(x_1^A, x_1^B)$ and disciplines neither agent otherwise.

[8] In this case, punishing only one agent is never optimal because it leads to either a strictly higher total discounted cost than punishing neither agent or a higher cost than punishing both agents. See the proof of Proposition 3 for the formal argument.

Fixing the principal's type, Figure 4(a) illustrates the two enforcement zones of the principal's period-1 strategy. Proposition 3 shows that the type-θ principal disciplines both agents when $\phi_{0|2}(x_1^A, x_1^B) > \theta$ (i.e., the sum $x_1^A + x_1^B$ is sufficiently large), and disciplines neither agent otherwise. Figure 4(b) compares the enforcement zones for a lower and a higher type of principal. Again, the principal becomes more lenient as its type increases.

4.1.3 Discussion

The optimal enforcement strategies depicted in Figure 3 and 4 have the following three key features. First, the principal attempts to build a *tough reputation* by disciplining agents aggressively in period 1, as the optimal enforcement strategy of a principal with $\delta_P > 0$ is more stringent than that of a *myopic* principal with $\delta_P = 0$. A myopic principal only cares about the within-period trade-off between the salvaged misconduct damage x_1^i and the enforcement cost θ. As a result, the myopic principal disciplines an agent if and only if $x_1^i > \theta$. In contrast, a principal with $\delta_P > 0$ uses a stricter enforcement strategy and punishes weakly more agents than its myopic counterpart with the same θ for any profile (x_1^A, x_1^B).

Second, the cost of building a tough reputation depends on whether the agents' misconduct levels are *close* to each other or *far apart*. First consider the case where two agents' misconduct levels are similar, that is, (x_1^A, x_1^B) is in the proximity cone. Without loss of generality, suppose x_1^A is *slightly* higher than x_1^B. If the principal wants to mimic being tough, it must punish both agents. Otherwise, the punishment of only agent A results in a relatively small improvement to its reputation, as it will be inferred that the principal's enforcement capacity is such that it will punish agent A, the one with more egregious misconduct, but not enough to punish agent B, whose misconduct is only slightly less severe. In other words, punishing only one agent reveals θ to be in a narrow range, leaving little information rent to the principal. In this case, the principal had better punish both agents for a significant reputation gain.

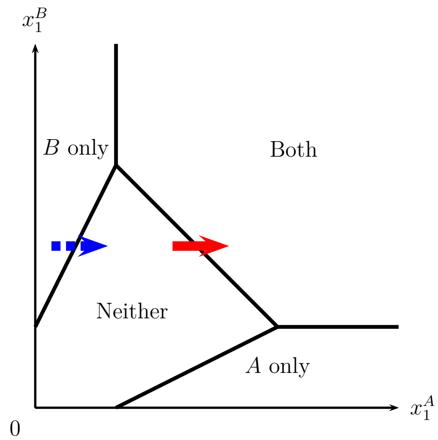
Now compare this outcome with the case where agent A's misconduct is *much* higher than agent B's, that is, (x_1^A, x_1^B) is not in the proximity cone. Now, punishing only agent A does not hurt substantially the principal's reputation, as agent B engages in only mild misconduct. The large gap in the agents' misconduct levels enables the principal to pool with much lower types, even though it punishes only one agent.

Third, the principal's optimal first-period disciplining of the two agents creates endogenous *enforcement externalities* between the two agents. The enforcement externalities may be positive or negative, as defined next.

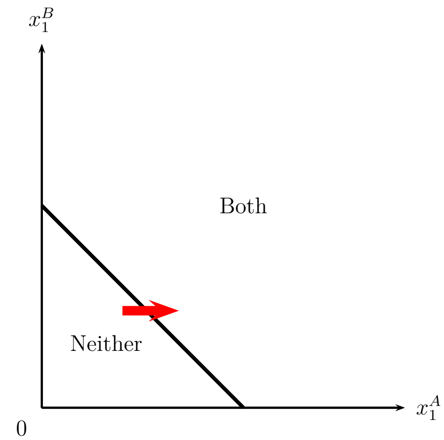
Definition 2 (Enforcement Externality) *For $i = A, B$,*
(i) A positive *enforcement externality arises when agent i's higher period-1 misconduct benefits peer agent $-i$ by reducing the probability of its peer being punished in period 1;*
(ii) A negative *enforcement externality arises when agent i's higher period-1 misconduct harms peer agent $-i$ by raising the probability of its peer being punished in period 1.*

Intuitively, agent i imposes a *positive* enforcement externality on the peer agent $-i$ and attenuates agent $-i$'s enforcement risk when agent i raises its misconduct from a much lower level than agent $-i$'s misconduct x_1^{-i} to a level close to x_1^{-i}. Leveraging our discussion in the second point, the principal tends to discipline only the peer agent $-i$ if its misconduct differs substantially from agent i's, as the principal pays the enforcement cost only once without sacrificing its reputation greatly. As x_1^i increases to a level close to x_1^{-i}, the principal now must discipline both to successfully mimic being tough, which is twice as costly. Therefore, the principal is inclined to tolerate both agents' misconduct, despite agent i's misconduct being more egregious. Higher misconduct by agent i, therefore, protects peer agent $-i$ by reducing its enforcement risk. The positive externality can be seen in Figure 5(a), represented by the dashed arrow for $\delta_P < \frac{1}{2}$. As x_1^A increases and crosses the upward-sloping boundary, the principal switches from disciplining only agent B to disciplining neither agent, reducing agent B's enforcement risk.

Alternatively, agent i imposes a *negative* enforcement externality on peer agent $-i$ and elevates agent $-i$'s enforcement risk when agent i raises its misconduct above a level close to x_1^{-i}. In this case, the two agents' misconduct levels are sufficiently close that the principal either punishes both agents or neither, depending on the *aggregate* misconduct $x_1^i + x_1^{-i}$. As the agent i increases its misconduct x_1^i, the sum increases, exposing peer agent $-i$ to a greater risk of being disciplined. The negative externality can be seen in both panels of



(a) $\delta_P < \frac{1}{2}$



(b) $\delta_P \geqslant \frac{1}{2}$

Figure 5: Enforcement externalities. Solid (resp. dashed) arrow shows the negative (resp. positive) externalities. (a): both externalities exist for $\delta_P < \frac{1}{2}$. (b): only negative externality exists for $\delta_P \geqslant \frac{1}{2}$.

Figure 5. Consider the solid arrows. As x_1^A increases and crosses the downward-sloping boundary, the principal switches from disciplining neither agent to disciplining both agents.

The enforcement externalities arise from the *interaction* between a non-myopic principal and peer learning. In contrast, one agent's misconduct has no bearing on its peer's enforcement when enforcement is opaque (Section 3) or when the principal is myopic.

4.2 Period 1: Agents' Misconduct

We next solve for the agents' equilibrium misconduct in period 1 given the principal's period-1 strategy. We simply write $\phi_{0|1}$, $\phi_{1|2}$, and $\phi_{0|2}$ while keeping in mind that they are functions of (x_1^A, x_1^B). For $\delta_P \geqslant \frac{1}{2}$, the misconduct levels (x_1^A, x_1^B) are always in the proximity cone (see Figure 2(b)). Agent i's total payoff is

$$U^i = \left(x_1^i + \delta \cdot \phi_{0|2}\right)\left(1 - F(\phi_{0|2})\right) + (-L)F(\phi_{0|2}). \tag{10}$$

To interpret, $\phi_{0|2}$ is the cutoff type of the principal facing misconduct profile (x_1^A, x_1^B). With probability $1 - F(\phi_{0|2})$, the principal disciplines neither agent in period 1, and agent i enjoys private benefit x_1^i and infers the principal's type to be at least $\phi_{0|2}$. Applying Lemma 1, the agent chooses $x_2^i = \phi_{0|2}$ to avoid the possibility of punishment. With the complementary probability, the principal disciplines both agents in period 1, and agent i suffers the penalty L and refrains from any misconduct in period 2.

When $\delta_P < \frac{1}{2}$, agent i's payoff is also characterized by (10) if (x_1^A, x_1^B) is in the proximity cone. Alternatively, if (x_1^A, x_1^B) is outside the proximity cone, there are two additional cases depending on whether x_1^A or x_1^B is larger. Agent i's payoff is

$$U^i = \left(x_1^i + \delta \cdot \phi_{0|1}\right)\left(1 - F(\phi_{0|1})\right) + \left(x_1^i + \delta \cdot \phi_{1|2}\right)\left(F(\phi_{0|1}) - F(\phi_{1|2})\right) + (-L)F(\phi_{1|2}) \quad \text{if } x_1^i < x_1^{-i}(1 - 2\delta_P), \tag{11}$$

$$U^i = \left(x_1^i + \delta \cdot \phi_{0|1}\right)\left(1 - F(\phi_{0|1})\right) + \left(-L + \delta \cdot \phi_{1|2}\right)\left(F(\phi_{0|1}) - F(\phi_{1|2})\right) + (-L)F(\phi_{1|2}) \quad \text{if } x_1^{-i} < x_1^i(1 - 2\delta_P). \tag{12}$$

In both expressions, the payoff comes from three events: both agents are disciplined (if $\theta < \phi_{1|2}$), only one agent is disciplined (if $\phi_{0|1} \leqslant \theta < \phi_{1|2}$), and neither agent is disciplined (if $\theta \geqslant \phi_{0|1}$). There is only one difference between (11) and (12): when only one agent is disciplined, agent i still receives x_1^i in period 1 in (11) because agent $-i$ was the one being disciplined, but agent i suffers a penalty L in (12) when it is disciplined.[9]

When enforcement is transparent, peer learning generates an *information externality* and an *enforcement externality* simultaneously. The information externality is well-documented in the collective experimentation literature. This effect arises when players experiment to learn about the properties of a non-strategic object, and a player learns from other players' outcomes and uses the information learned in future decision making. In our setting, the positive information externality still exists even though the agents face a strategic principal instead of a non-strategic object. Intuitively, if agent $-i$ was not disciplined

[9] To interpret (11) and (12), suppose $i = A$. Consider first the case with $x_1^A < x_1^B(1 - 2\delta_P)$, that is, agent B's period-1 misconduct is the higher of the two. We interpret (11), which is agent A's payoff. With probability $1 - F(\phi_{0|1})$, we have $\theta \geqslant \phi_{0|1}$ and neither agent is disciplined. Agent A enjoys a private benefit of x_1^A in period 1, and the agents' posterior belief at the end of period 1 has the support $[\phi_{0|1}, \infty)$. Agent A then chooses $x_2^A = \phi_{0|1}$ to avoid punishment in period 2. Hence, agent A's total payoff conditional on no enforcement is $x_1^A + \delta\phi_{0|1}$. With probability $F(\phi_{0|1}) - F(\phi_{1|2})$, only agent B, the one with higher misconduct, is disciplined. Agent A enjoys private benefit x_1^A in period 1 and chooses $x_2^A = \phi_{1|2}$, the lower bound of the support of the updated belief. Agent A's total payoff conditional on disciplining only agent B is $x_1^A + \delta\phi_{1|2}$. Finally, with probability $F(\phi_{1|2})$, both agents are disciplined. Agent A suffers the penalty L in period 1, and it chooses $x_2^A = 0$. Agent A's total payoff conditional on both agents being disciplined is $-L$. The case in which $x_1^B < x_1^A(1 - 2\delta_P)$ can be interpreted analogously, with the only difference being that when only one agent is disciplined, agent A instead of agent B is disciplined.

after engaging in aggressive misconduct in period 1, then agent i learns from this outcome and sets its period-2 misconduct x_2^i to depend on agent $-i$'s period-1 misconduct x_1^{-i}. This is illustrated by (10)-(12), as x_2^i is set to $\phi_{0|2}(x_1^i, x_1^{-i})$ in (10), $\phi_{0|1}(x_1^i, x_1^{-i})$ or $\phi_{1|2}(x_1^i, x_1^{-i})$ in (12), and $\phi_{0|1}(x_1^i, x_1^{-i})$ in (11).[10]

In contrast, the enforcement externality is a novel effect that only arises when peer learning interacts with a non-myopic strategic principal. As is discussed in Section 4.1, fixing a principal's type θ, agent i's misconduct may affect the peer agent's enforcement outcome favorably (i.e., positive enforcement externality) or unfavorably (i.e., negative enforcement externality). When taking expectations across all possible principal's types, the enforcement externality is reflected in (10) and (12) because the probability of agent i being disciplined depends on agent $-i$'s misconduct x_1^{-i}.[11] That is, agent i's enforcement risk in (10), $F(\phi_{0|2}) = F(\frac{x_1^i + x_1^{-i}}{2(1-\delta_P)})$, increases as x_1^{-i} increases (negative externality); agent i's enforcement risk in (12), $F(\phi_{0|1}) - F(\phi_{1|2}) + F(\phi_{1|2}) = F(\phi_{0|1}) = F(\frac{x_1^i(1-2\delta_P) - x_1^{-i}(2\delta_P)}{(1-2\delta_P)^2})$, decreases as x_1^{-i} increases (positive externality).[12]

Against this background, we conduct the following analysis. First, we solve for the agents' equilibrium misconduct. Second, to study the effect of enforcement transparency, we compare the equilibrium misconduct under the transparent enforcement setting to that in the opaque enforcement benchmark. We find that both the equilibrium misconduct levels and the effect of enforcement transparency depend on the principal's level of patience.

4.2.1 Impatient Principal

When $\delta_P < \frac{1}{2}$, the next proposition shows that if the penalty L is sufficiently large, neither agent engages in misconduct in period 1.[13]

[10]Although x_2^i also may be set to $\phi_{1|2} = \frac{x_1^i}{1-2\delta_P}$ in (11), this term does not depend on x_1^{-i}, so the information externality is not reflected here.

[11]Since the two agents are symmetric, the enforcement externalities can either be interpreted as peer agent $-i$'s enforcement risk being affected by agent i's misconduct, or agent i's enforcement risk being affected by peer agent $-i$'s misconduct.

[12]Agent i's enforcement risk in (11) is $F(\phi_{1|2}) = F(\frac{x_1^i}{1-2\delta_P})$, which does not depend on x_1^{-i}. So there is no enforcement externality in this case.

[13]To ensure tractability, the next proposition requires $L > \hat{L}(\delta_P)$. We show numerically in Section 5.1 that the qualitative results remain unchanged if we lift this restriction. In the more general case, the period-1 equilibrium misconduct levels are not necessarily zero, but they continue to be low.

Proposition 4 (Equilibrium Misconduct: $\delta_P < \frac{1}{2}$)
For every $\delta_P < \frac{1}{2}$, there exists $\hat{L}(\delta_P) < \frac{1+\delta-\delta_P}{h(0)}$ such that when $L > \hat{L}(\delta_P)$, there exists a unique equilibrium in pure strategies in which neither agent engages in misconduct in either period, i.e., $x_t^i = 0$, for $i = A, B$ and $t = 1, 2$.

Although the agents choose a positive misconduct level $x^* > 0$ in the opaque-enforcement benchmark (Proposition 1), Proposition 4 shows that they refrain from any misconduct in the presence of peer learning for a range of penalty L. Thus, the principal benefits when enforcement is transparent as the transparency reduces misconduct. The next corollary summarizes the comparison.

Corollary 3 (Effect of Enforcement Transparency: $\delta_P < \frac{1}{2}$)
When $\delta_P < \frac{1}{2}$, there exists $\hat{L}(\delta_P) < \frac{1+\delta-\delta_P}{h(0)}$ such that when $L > \hat{L}(\delta_P)$, enforcement transparency strictly decreases the agents' misconduct in both periods. Further, introducing enforcement transparency strictly reduces the principal's total discounted cost for all types $\theta > 0$.[14]

To see the intuition underlying Corollary 3, when $\delta_P < \frac{1}{2}$, there is a *positive* information externality between the two agents, inducing an agent to free ride on its peer's experimentation and to reduce its own period-1 misconduct. In addition, there is a *positive* enforcement externality, which also drives down the agents' misconduct. To see this, suppose both agents start with the opaque benchmark equilibrium misconduct $x_1^A = x_1^B = x^* > 0$. Given agent B chooses to engage in misconduct x^*, agent A can shield behind agent B's misconduct. Hence, agent A will choose to deviate to a low level of misconduct that falls outside the proximity cone, which significantly lowers its chance of being disciplined. Similarly, agent B also has an incentive to deviate. Thus, in equilibrium, $x_1^A = x_1^B = 0$.

Corollary 3 implies that enforcement transparency allows the principal to "make an example" of a nefarious agent, thereby discouraging agent misconduct and reducing the principal's total cost.

[14] Numerically we show in Section 5.1 that Corollary 3 still holds if we lift the restriction of $L > \hat{L}(\delta_P)$.

4.2.2 Patient Principal

When $\delta_P \geqslant \frac{1}{2}$, the principal's period-1 strategy is to either discipline both agents or discipline neither, depending on whether $x_1^A + x_1^B$ is sufficiently high (Proposition 3). The next proposition characterizes the agents' equilibrium misconduct.

Proposition 5 (Equilibrium Misconduct: $\delta_P \geqslant \frac{1}{2}$)
For $\delta_P \geqslant \frac{1}{2}$, there exists a unique equilibrium in pure strategies in which the agents choose the same misconduct level $x_1^A = x_1^B = x^{**}$ in period 1. The level $x^{**} > 0$ uniquely solves

$$\left(L + (1 + \delta - \delta_P)\phi_{0|2}(x^{**}, x^{**})\right) h\left(\phi_{0|2}(x^{**}, x^{**})\right) = 2 + \delta - 2\delta_P. \tag{13}$$

In addition, $x_2^i = 0$ if $(e_1^A, e_1^B) = (1, 1)$, and $x_2^i = \phi_{0|2}(x^{**}, x^{**})$ if $(e_1^A, e_1^B) = (0, 0)$.

In contrast to when the principal is impatient and enforcement transparency is beneficial, resulting in neither agent engaging in misconduct, Proposition 5 shows that when the principal is patient, enforcement transparency is no longer beneficial as it actually heightens the agents' misconduct and the principal's total costs. This finding is characterized in the next corollary.

Corollary 4 (Effect of Enforcement Transparency: $\delta_P \geqslant \frac{1}{2}$)
Suppose $\delta_P \geqslant \frac{1}{2}$. Enforcement transparency yields strictly higher agent misconduct in both periods. Further, introducing enforcement transparency results in weakly higher levels of the principal's total discounted cost C for all types $\theta > 0$, and the increase is strict for all types $\theta > \phi(x^*)$.

The information externality is always positive, which reduces misconduct, regardless of whether the principal is patient or not. In contrast, although the enforcement externality is positive when the principal is impatient, it is negative when the principal is patient. Indeed, when the principal is patient, the negative enforcement externality, which encourages misconduct, overwhelms the positive information externality, which discourages it. Hence, enforcement transparency leads to more misconduct.

To see why the effect of the enforcement externality dominates, suppose both agents start with the equilibrium misconduct level in the opaque enforcement benchmark, $x_1^A = x_1^B = x^* > 0$. In the benchmark, an agent has no incentive to deviate from the equilibrium level x^*, so the total marginal benefits equal the marginal cost. Accordingly, when an agent marginally increases its misconduct to $x^* + \mathrm{d}x$, the agent's expected payoff in (5) in the benchmark becomes

$$\Bigg(\underbrace{x^* + \mathrm{d}x}_{\substack{\text{private}\\\text{benefit}}} + \delta \cdot \underbrace{\phi\left(x^* + \mathrm{d}x\right)}_{\text{learning}}\Bigg)\Bigg(1 - \underbrace{F\left(\phi\left(x^* + \mathrm{d}x\right)\right)}_{\substack{\text{enforcement}\\\text{risk}}}\Bigg) + (-L)\,\underbrace{F\left(\phi\left(x^* + \mathrm{d}x\right)\right)}_{\substack{\text{enforcement}\\\text{risk}}}. \tag{14}$$

As the marginal benefits equal the marginal cost, the marginally higher private benefit, $\mathrm{d}x$, and the marginally higher benefit from learning (because an agent learns that it can choose higher period-2 misconduct without being punished), $\phi\left(x^* + \mathrm{d}x\right) - \phi\left(x^*\right)$, are offset by the marginally heightened enforcement risk in period 1, $F\left(\phi\left(x^* + \mathrm{d}x\right)\right) - F\left(\phi\left(x^*\right)\right)$.

In contrast, when enforcement is transparent, the marginal benefits exceed the marginal cost, causing an agent to engage in more misconduct than when enforcement is opaque. When enforcement is transparent, an agent's payoff in (10) after the deviation to $x + \mathrm{d}x$ simplifies to

$$\Bigg(\underbrace{x^* + \mathrm{d}x}_{\substack{\text{private}\\\text{benefit}}} + \delta \cdot \underbrace{\phi\left(x^* + \frac{\mathrm{d}x}{2}\right)}_{\text{learning}}\Bigg)\Bigg(1 - \underbrace{F\left(\phi\left(x^* + \frac{\mathrm{d}x}{2}\right)\right)}_{\text{enforcement risk}}\Bigg) + (-L)\,\underbrace{F\left(\phi\left(x^* + \frac{\mathrm{d}x}{2}\right)\right)}_{\text{enforcement risk}}. \tag{15}$$

after applying the definition of ϕ in (3) and the relation that $\phi(\frac{x_A + x_B}{2}) = \phi_{0|2}(x_A, x_B)$. Comparing (14) to (15), we observe that enforcement transparency does not affect the agent's marginal private benefit, $\mathrm{d}x$, whereas it alters the marginal benefit from learning and the marginal enforcement risk. The information externality reduces the marginal benefit from learning in period 1 from the benchmark level $\phi\left(x^* + \mathrm{d}x\right) - \phi\left(x^*\right)$ to $\phi\left(x^* + \mathrm{d}x/2\right) - \phi\left(x^*\right)$, because the agent is less willing to experiment on its own when it can free ride on its peer's experimentation. This effect reduces the total marginal benefit and pushes misconduct down. On the other hand, the enforcement externality reduces the agent's marginal enforcement risk from the benchmark level $F\left(\phi\left(x^* + \mathrm{d}x\right)\right) - F\left(\phi\left(x^*\right)\right)$ to $F\left(\phi\left(x^* + \mathrm{d}x/2\right)\right) - F\left(\phi\left(x^*\right)\right)$

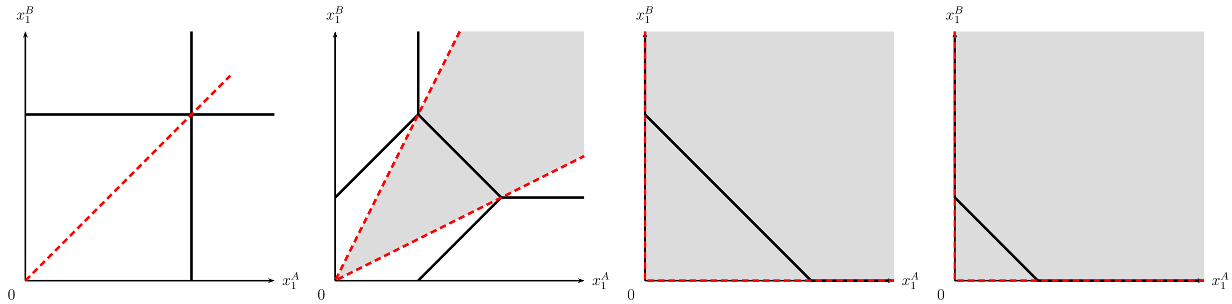
as part of the heightened enforcement risk is shifted to its peer. This effect reduces the marginal cost and pushes misconduct up.

As the marginal private benefit is unaffected by enforcement transparency, the *total* marginal benefit is reduced by a relatively lower proportion than the marginal cost. Consequently, an agent chooses a higher level of misconduct when enforcement is transparent than in the benchmark level. In short, the negative enforcement externality is stronger than the positive information externality, yielding greater misconduct.

Effect of Enforcement Transparency and Principal's Decision Horizon. Comparing Corollary 3, when the principal is impatient, with Corollary 4, when the principal is patient, shows that whether a principal should embrace enforcement transparency depends on its *decision horizon* (i.e., discount factor δ_P).

When the principal is impatient ($\delta_P < \frac{1}{2}$), the peer learning resulting from enforcement transparency is beneficial, because it enables the principal to make an example of a nefarious agent to deter misconduct. This result echos that of the traditional collective experimentation literature, which recognizes that peer learning reduces or slows down equilibrium experimentation due to information spillovers (Bolton and Harris, 1999; Keller et al., 2005). The positive enforcement externality from high peer misconduct, which is absent in those papers but at play in our setting when $\delta_P < \frac{1}{2}$, adds to the effect of the information spillover by inducing an agent to reduce its own misconduct and shield behind its peer's high misconduct.

When the principal is patient ($\delta_P \geqslant \frac{1}{2}$), however, the positive enforcement externality loses power while the negative enforcement externality kicks in, reversing the effect of the information spillover and causing enforcement transparency (and the resulting peer learning) to harm the principal. This is because a more patient principal is increasingly concerned about building a reputation for tough enforcement. Consequently, the principal is more reluctant to punish only one agent and is more inclined to punish both agents. This reluctance is reflected by the increasing area of the proximity cone (where the principal disciplines both agents or neither agent) as δ_P gets larger. Indeed, Figure 6 shows that as δ_P increases, the proximity cone, which is depicted by shaded area, increases from being non-existent when



(a) $\delta_P \to 0$ (b) $\delta_P = \frac{1}{4}$ (c) $\delta_P = \frac{1}{2}$ (d) $\delta_P = \frac{3}{4}$

Figure 6: Principal's period-1 strategy as δ_P varies. The shaded area reflects the proximity cone.

$\delta_P \to 0$ to the entire (x_1^A, x_1^B) quadrant when $\delta_P \geqslant \frac{1}{2}$.

As the positive enforcement externality, which discourages misconduct, arises only when the principal might only punish one of the agents (see Figure 5), such an externality becomes less relevant when the principal is patient and less willing to punish only one agent. In contrast, the negative enforcement externality, which encourages misconduct, arises exactly when the principal punishes both agents or neither.

Misconduct, Principal's Total Cost, and Principal's Decision Horizon. Having studied the effect of enforcement transparency on misconduct when the principal's patience level δ_P is fixed, we next investigate how misconduct will change when δ_P varies. We find that the result depends crucially on whether enforcement is transparent. When enforcement is opaque, Corollary 1 implies that a principal with a longer decision horizon (a higher δ_P) is more capable of deterring misconduct. When enforcement is transparent, however, misconduct is non-monotonic in the principal's patience δ_P, as established in the next corollary. Consequently, agents may engage in more misconduct when the principal is more patient.

Corollary 5 (Effect of δ_P: Transparent Enforcement)
Consider any L satisfying Assumption 1. For all $\delta_P < \frac{1}{2}$ such that $L > \hat{L}(\delta_P)$, the agents' misconduct remains zero in both periods. At $\delta_P = \frac{1}{2}$, the agents' misconduct jumps up in both periods. For $\delta_P > \frac{1}{2}$, the agents' misconduct strictly decreases in δ_P in both periods.

To see the intuition underlying the non-monotonicity, observe that a longer decision horizon, or equivalently, a stronger incentive to build a reputation for tough enforcement, has two *countervailing* effects on the agents' misconduct when enforcement is transparent. First, the principal applies a more stringent enforcement criterion in period 1 to signal being a tough type. This effect, which is the only effect present in the opaque enforcement benchmark, deters agent misconduct. Second, the principal is more inclined to discipline both agents simultaneously and less willing to punish only one agent. Therefore, an agent is less willing to reduce its misconduct to shield behind its peer, yielding higher misconduct.

5 Extensions

We extend our analysis by considering a wider parameter range, and we show, with numerical examples, that the main qualitative results are unchanged.

5.1 Impatient Principal: Asymmetric Equilibrium

In Proposition 4, where the principal is impatient, we suppose the penalty $L > \hat{L}(\delta_P)$, and we characterize the unique equilibrium. It is symmetric with $x_1^A = x_1^B = 0$. When the penalty $L < \hat{L}(\delta_P)$, we may lose both symmetry and uniqueness; however, the result in Corollary 3 that transparent enforcement deters misconduct and reduces the principal's total costs still holds.

To show that enforcement transparency reduces misconduct, let $\delta = 1$, $\delta_P = 0.1$, and the prior belief about the principal's type follows the exponential distribution with a c.d.f. of $F(\theta) = 1 - e^{-\theta}$. Assumption 1 requires $L \in [1, 1.9)$. The unique period-1 equilibrium is $x_1^A = x_1^B = 0$ when $L > \hat{L}(\delta_P) = 1.8$. For a smaller L, $L = 1.6$, there exist two asymmetric equilibria, where one agent chooses zero misconduct and the other engages in positive misconduct, $(x_1^i, x_1^{-i}) = (0.088, 0)$, $i = A, B$. For an even smaller L, $L = 1$, there remain two asymmetric equilibria, where both agents engage in positive misconduct: $(x_1^i, x_1^{-i}) = (0.372, 0.024)$. Observe that when L is sufficiently low, symmetry is broken as one of the two agents becomes the "pioneer" in learning while the other shields behind it. That said, the comparison with the opaque enforcement benchmark still holds. For $L = 1.6$, the benchmark yields period-1 misconduct of $x_1^A = x_1^B = 0.142$, higher than both misconduct levels when enforcement is transparent. For $L = 1$, the period-1 misconduct is $x_1^A = x_1^B = 0.426$

in the benchmark, compared to $(0.372, 0.024)$ when enforcement is transparent. Therefore, an impatient principal can leverage enforcement transparency to deter misconduct and lower total discounted costs, as in the main model.

In addition, the non-monotonicity property in Corollary 5 also holds, implying that misconduct can be higher when the principal is more patient. Fixing $L = 1$, the expected misconduct in period 1 starts with $\frac{x_1^A + x_1^B}{2} = 0.377$ at $\delta_P = 0$, monotonically declines to zero as $\delta_P \to \frac{1}{2}$, jumps up to 0.333 and then declines again for $\delta_P > \frac{1}{2}$.

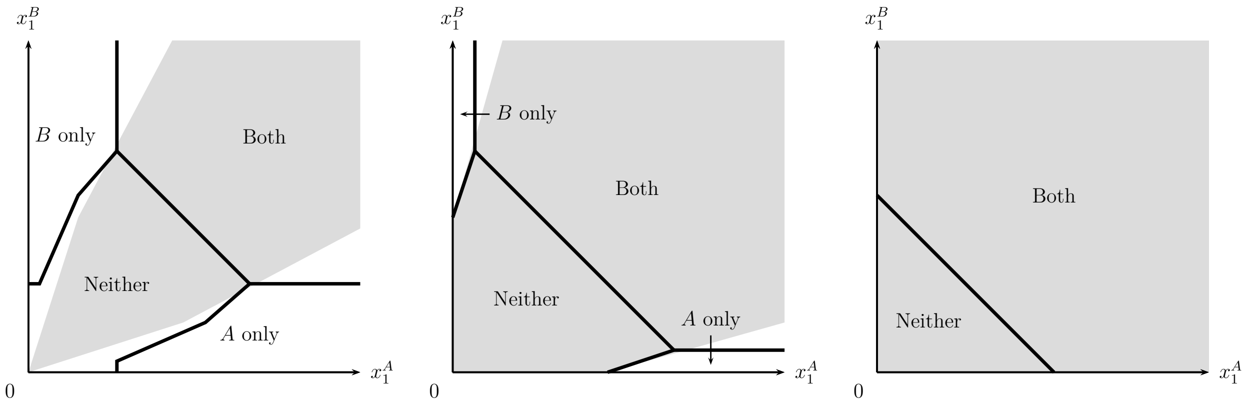
5.2 Interior Period-2 Misconduct

The main model assumes the penalty L cannot be too small (see Assumption 1(ii)), which causes the period-2 misconduct to be the highest level that avoids being disciplined, that is, the lower bound of the support of the updated beliefs about the principal's type. In this extension we relax the restriction by allowing for $Lh(0) < 1$. Intuitively, when the penalty L is not intimidating, agents may accept some enforcement risk in period 2 by choosing misconduct in the interior of the support of the updated beliefs. The next lemma specifies the interior solution.

Lemma 3 *In period 2, if the posterior belief of agent i is the prior with truncated support $[\theta_L, \theta_H]$, then $x_2^i = x_2^*(\theta_L, \theta_H) \in [\theta_L, \theta_H)$ is the unique solution to $(L + x_2^i)h(x_2^i) = \frac{F(\theta_H) - F(x_2^i)}{1 - F(x_2^i)}$ if $(L + \theta_L)h(\theta_L) < \frac{F(\theta_H) - F(\theta_L)}{1 - F(\theta_L)}$ and $x_2^i = \theta_L$ otherwise.*

Lemma 3 implies that a weak penalty L can result in period-2 misconduct levels being higher than the lower bound of the truncated belief support. With weak deterrence, the principal's reputation boost from enforcement is less valuable, and intuitively, it is less willing to discipline agents in period 1. Analogous to the main model, we can derive the principal's enforcement strategy in period 1. Lemma 2 still holds because its proof does not rely on the strategies of the agents in period 2. As the analytical form of the principal's strategy is convoluted, we offer the following numerical example to show the resemblance to the main model.

Let $\delta = 1$, $L = 1$, and suppose $F(\theta) = 1 - e^{-\theta/2}$ is an exponential distribution of types. Since $Lh(0) = 1/2 < 1$, Assumption 1 is violated. Figure 7 shows the enforcement zones



(a) $\delta_P = 1/3$ (b) $\delta_P = 1/2$ (c) $\delta_P = 2/3$

Figure 7: Principal's enforcement zones in period 1, when $Lh(0) = \frac{1}{2} < 1$. Grey areas show proximity sets. Parameters: $L = 1$, $F(\theta) = 1 - e^{-\theta/2}$, $\theta = 2$

in period 1 for a principal of type $\theta = 2$. When the principal is impatient ($\delta_P = 1/3$), the four enforcement zones resemble those in the main model, although the boundaries of the proximity cone are no longer linear. When $\delta_P = 1/2$, the principal's enforcement strategy is more tolerant than in the main model as the gain from a reputation boost is less valuable. When the principal is even more patient ($\delta_P = 2/3$), the enforcement zones become isomorphic to Figure 4(a) in the main model. The resemblance generates similar incentives for the agents in period 1. For $\delta_P = 1/3$, pure strategy equilibrium does not exist, but a mixed strategy equilibrium exists in which $x_1^A = 0.422$ with probability 0.764 and $x_1^A = 0.552$ with probability 0.236, while $x_1^B = 0.023$. This strategy profile gives an expected period-1 misconduct of 0.227, lower than 1.028 in the opaque enforcement benchmark, which is consistent with the result in Corollary 3. For $\delta_P = 2/3$, the unique pure strategy equilibrium features $x_1^A = x_1^B = 0.887$, higher than 0.550 in the benchmark, which is consistent with the result in Corollary 4. Finally, the comparative statics for δ_P is similar to Corollary 5: the expected period-1 misconduct is 0.227 at $\delta_P = 1/3$, declines to 0 as δ_P converges to 1/2, before jumping up to 1.268 at $\delta_P = 1/2$, and then declines again to 0.887 at $\delta_P = 2/3$.

6 Implications for Specific Settings

In this section, we identify institutional settings to which the model and its results may be applicable. We also discuss the policy implications of our model and our paper's relation to the literature in those settings.

Capital Budgeting Decisions. Consider a corporate setting in which a CEO/Board cares about the total firm value and the division managers have empire-building aspirations. When the CEO/Board's decision horizon is lengthy (e.g., incentive package has a long-term duration, or the CEO/Board is likely to be reelected in the near future), then the divisional manager's empire-building behavior is encouraged and the CEO/Board is worse-off when its interventions in the divisions are transparent. In contrast, if the CEO/Board's horizon is short, then enforcement transparency is beneficial. In addition, a CEO/Board's longer term decision horizon reduces agent empire-building when the CEO/Board's interventions are opaque, while it may heighten empire-building when enforcement is transparent. Our paper is related to the literature on capital budgeting (Harris and Raviv (1998); Bernardo, Cai, and Luo (2001); Malenko (2019)).[15]

Interventions by Common Owners. Consider a common institutional owner and self-interested managers of its portfolio firms who may engage in behavior that destroys firm value. Our results imply that whether interventions in the portfolio firms should be disclosed depends on whether the institutional common owner adopts a short-term or a long-term investment strategy. In addition, having a shorter investment horizon may not necessarily be harmful to the common owner because it may reduce misconduct when interventions in management of portfolio firms are transparent. In this sense, our work is related to the literature on institutional common ownership (Edmans, Levit, and Reilly (2019); Gilje, Gormley, and

[15]Harris and Raviv (1998) study the capital allocation process within firms. Bernardo, Cai, and Luo (2001) study optimal capital allocation and managerial compensation mechanisms for decentralized firms when division managers have an incentive to misrepresent project quality and to minimize privately costly but value-enhancing effort. Malenko (2019) studies a dynamic capital allocation process in an organization in which a division manager has empire-building preferences. These papers focus on the private information possessed by the agents (divisions), while we investigate the reputation management of the principal (headquarters).

Levit (2020)),[16] short-termism of institutional investors (Dasgupta and Piacentino (2015); Song (2019); Burkart and Dasgupta (2021)),[17] and shareholder activism spillovers and threat of shareholder activism (Lee and Park (2009); Gantchev, Gredil, and Jotikasthira (2019); Lakkis (2021)).

Regulations of Societally-Damaging Misconduct. Consider a regulatory environment with two regulated firms (agents) and a regulator (principal). A firm may misbehave to obtain a higher profit at the expense of the social welfare. The regulator, who internalizes the social damage from the misconduct, may intervene and correct each firm's misconduct at a cost.[18] A regulator with a longer decision horizon is interpreted as the regulator having a longer term or a higher probability of remaining in office. The following regulatory environments fit our setting, as they feature (a) a regulator's forbearance of less egregious misbehavior, (b) uncertainty of the regulatory standards, and (c) misconduct largely observable to the regulator and peer firms.

- Environmental, Social, and Governance (ESG) reporting[19]
- Regulations of over-collection of user data[20]

[16]Edmans, Levit, and Reilly (2019) show that governance is strengthened with common ownership. Gilje, Gormley, and Levit (2020) show that drivers of common ownership can diminish managerial motives to internalize externalities. We take a different perspective by investigating the common owner's strategic interventions in its portfolio firms and disclosure of the interventions.

[17]Dasgupta and Piacentino (2015) show that the competition for investor capital creates short-term incentives among money manager blockholders, weakening their incentives to exit and dampening their governance roles. Song (2019) finds that while blockholder short-term incentives are harmful in a single blockholder structure, they may enhance blockholder governance when there is a mixture of blockholders with both long-term and short-term incentives. Burkart and Dasgupta (2021) show how competition for flow among activist funds can induce excessive leverage in target firms. Our work proposes a potential channel through which a blockholder's short-term incentive deters firm misbehavior, and this channel relies on whether the interventions in firms are publicized.

[18]The cost may include staff time, agency resources, and litigation costs; it may also include any potential negative externality that an enforcement action imposes on the society, such as market instability and discouraged productive activities.

[19]In early 2021, the U.S. Securities and Exchange Commission (SEC) proposed rules for mandatory reporting relating to areas such as climate change (https://www.sec.gov/news/press-release/2022-46), human capital management, board diversity and cybersecurity risk governance. Also in 2021, the European Commission (EC) adopted a proposal for a corporate sustainability reporting directive (CSRD) to include specific European sustainability reporting standards (ESRS) (https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en).

[20]In data privacy regulations, the over-collection of user data by applications is publicly observable (i.e., available in the user agreement), pervasive, and often undisciplined (e.g., Fan et al., 2020). Meanwhile, companies face substantial uncertainty in acceptable levels of compliance (Drake, 2017).

- Regulations for payment for order flow (PFOF)[21]
- SEC 2003 mutual fund voting rule that requires mutual funds to vote in the best interest of fund shareholders.[22]

Our model has policy implications for the design of a regulatory environment. It characterizes two aspects of the environment: (i) whether enforcement actions should be publicly disclosed, and (ii) whether optimal regulation involves enforcement externalities among the regulated entities. Our results indicate that enforcement transparency is beneficial (harmful) when the regulator has a shorter (longer) decision horizon. In addition, the regulator's optimal enforcement strategy may induce either a positive enforcement externality or a negative enforcement externality.[23] Lastly, longer administrative terms for regulators may be harmful because they may induce higher misconduct.

To the extent that our setting applies to regulatory environments, our paper is related to the literature on regulatory reputation management (Boot and Thakor (1993); Morrison and White (2013); Shapiro and Skeie (2015); Huang (2017)),[24] regulatory enforcement transparency (see Goldstein and Sapra (2013) and Goldstein and Leitner (2020) for surveys of

[21] PFOF is a system by which wholesale trading giants (e.g., Citadel) pay retail brokerages (e.g., Robinhood) to execute their clients' orders. Information on PFOF is publicly available because market centers are required to disclose information online according to SEC Rule 605 and Rule 606. PFOF may result in inferior execution prices for individual brokerage clients (see https://www.bloomberg.com/news/articles/2022-06-08/sec-chief-takes-aim-at-payment-for-order-flow-in-sweeping-plans). Aggressive practice of PFOF may subject a broker to SEC enforcement. For example, in 2020, the SEC charged Robinhood Financial with misleading customers about revenue sources and failing to satisfy its duty of best execution (See https://www.sec.gov/news/press-release/2020-321).

[22] See https://www.sec.gov/news/press/2003-12.htm.

[23] In practice, there exists a positive enforcement externality among offenders in that the probability of an individual being punished for violating the law decreases with the number of individuals violating the law (Ehrlich, 1973). There also exists a negative enforcement externality in the form of regulatory spillover, in the sense that one agent's aggressive offense increases another agent's chance of being regulated (e.g., Donelson, Flam, and Yust, 2022). Our theory provides a justification for these externalities.

[24] Boot and Thakor (1993) and Morrison and White (2013) consider settings in which the regulator's reputation is about its ability to monitor banks ex ante, instead of its preference to punish banks ex post. Shapiro and Skeie (2015) consider a setting where a regulator, who privately knows its cost of injecting capital, decides whether to inject capital into two short-lived banks in a sequential fashion. Unlike our model, their papers involve neither strategic experimentation nor a mutually dependent enforcement strategy. Huang (2017) considers a dynamic regime change setting in a global game framework, where the policy maker, who faces a continuum of speculators that may coordinate to attack the current regime, builds a reputation for strong preference to maintain the status quo. In contrast to our model, there is no strategic experimentation in the study because the speculators are atomistic and hence behave myopically.

the literature that focuses on banking regulations), and crime (Bond and Hagerty (2010)).[25]

7 Conclusion

We examine a two-period model in which two agents make misconduct decisions in the presence of a strategic principal whose tolerance for misconduct is unknown to the agents. We depart from the existing literature on collective experimentation by allowing the object about which the players are learning to be a *strategic* principal. We show that optimal enforcement creates endogenous enforcement externalities. As a result, enforcement transparency, which allows the principal to "make an example" of the nefarious agent by publicly punishing it, may actually increase equilibrium misconduct. We also show that a principal with a longer decision horizon may induce higher misconduct. Our results have implications for enforcement transparency and a principal's decision horizon. Our findings apply to various institutional settings, including relations between headquarters and division managers, common owners and portfolio firm managers, and regulators and firms.

[25]Bond and Hagerty (2010) study the design of enforcement mechanisms, in which the fixed total enforcement budget for the agents generates a positive enforcement externality, because when other individuals commit more crime, enforcement resources are stretched. Our enforcement externalities differ because (i) the positive externality still arises even if the enforcement cost on one agent does not depend on whether the other agent is disciplined, and (ii) the enforcement externalities can also be negative.

Appendix: Proofs

Proof of Lemma 1

Proof. (i) Since period 2 is the terminal period, the principal minimizes its continuation cost $\sum_{i=A,B}[x_2^i + e_2^i(\theta - x_2^i)]$. Therefore, $e_2^i = 1$ ($e_2^i = 0$) if $x_2^i > \theta$ ($x_2^i < \theta$). When $x_2^i = \theta$, Assumption 2 requires that $e_2^i = 0$.

(ii) In period 2, an agent i maximizes its continuation payoff $\mathbb{E}\left[x_2^i(1-e_2^i) - Le_2^i|q^i\right] = x_2^i(1-\mathbb{E}[e_2^i|q^i]) - L\mathbb{E}[e_2^i|q^i]$. According to the principal's strategy, $\mathbb{E}[e_2^i|q^i] = \Pr(x_2^i > \theta|q^i)$.

If the posterior belief is the prior truncated to $[\theta_L, \theta_H)$, then agent i, who chooses misconduct $x_2^i \in [\theta_L, \theta_H)$, receives a continuation payoff:

$$x_2^i \frac{F(\theta_H) - F(x_2^i)}{F(\theta_H) - F(\theta_L)} - L\frac{F(x_2^i) - F(\theta_L)}{F(\theta_H) - F(\theta_L)}.$$

Differentiating w.r.t. x_2^i: $\frac{f(x_2^i)}{F(\theta_H)-F(\theta_L)}\left(\frac{F(\theta_H)-F(x_2^i)}{f(x_2^i)} - (L + x_2^i)\right) \leqslant \frac{f(x_2^i)}{F(\theta_H)-F(\theta_L)}\left(\frac{1}{h(0)} - L\right) < 0$, where the first inequality is due to Assumption 1 (i) and the second inequality is due to Assumption 1 (ii). This means the continuation payoff is strictly decreasing in $x_2^i \in [\theta_L, \theta_H)$. If $x_2^i < \theta_L$, then the payoff is x_2^i. If $x_2^i \geqslant \theta_H$, the payoff is $-L$. Therefore, the globally optimal choice is $x_2^i = \theta_L$. ■

Proof of Proposition 1

Proof. Given x_1^i, the principal's cost (1) from agent i, as a continuous function of θ, has a derivative of at least 1 almost everywhere. In contrast, the cost (2) has a derivative of at most $\delta_P < 1$ almost everywhere. Moreover, as $\theta \to \infty$, (1) is greater than (2). Therefore, the principal's strategy in period 1 must feature a cutoff $\theta^\dagger \geqslant 0$ such that $e_1^i = 1$ if and only if $\theta < \theta^\dagger$. The indifference is broken in favor of not disciplining because of Assumption 2. Then, according to Bayes' rule, agent i's posterior belief is the prior distribution truncated to $[\theta^\dagger, \infty)$ if $e_1^i = 0$, and truncated to $[0, \theta^\dagger)$ if $e_1^i = 1$. By Lemma 1, agent i chooses $x_2^i = \theta^\dagger$ and $x_2^i = 0$, respectively. Plugging into (1) and (2) and observe that (1) and (2) are equal at the cutoff type $\theta = \theta^\dagger$, we have $\theta^\dagger + \delta_P \cdot 0 = x_1^i + \delta_P \cdot \theta^\dagger$. Solving the equation, we have $\theta^\dagger = \phi(x_1^i) = \frac{x_1^i}{1-\delta_P}$.

In period 1, agent i's payoff is $(x_1^i + \delta\phi(x_1^i))\,(1 - F\,(\phi(x_1^i))) + (-L + \delta \cdot 0)F\,(\phi(x_1^i))$, its derivative having the same sign as $(1 + \delta - \delta_P) - (L + (1 + \delta - \delta_P)\phi(x_1^i))\,h\,(\phi(x_1^i))$. It is strictly decreasing in x_1^i by Assumption 1. It diverges to $-\infty$ as $x_1^i \to \infty$ and evaluates to $1 + \delta - \delta_P - Lh(0) > 0$ when $x_1^i = 0$. By Intermediate value theorem, there exists a unique $x^* > 0$ such that $(L + (1 + \delta - \delta_P)\phi(x_1^i))\,h\,(\phi(x_1^i)) = 1 + \delta - \delta_P$, and agent i's payoff is maximized at $x_1^i = x^*$. In addition, x^* is global optimal because the first derivative is positive when $x < x^*$ and negative when $x > x^*$. ■

Proof of Corollary 1

Proof. Equation (6) can be rewritten as:

$$G(\delta_P, \phi(x^*)) \equiv (L + (1 + \delta - \delta_P)\phi(x^*))\,h\,(\phi(x^*)) - (1 + \delta - \delta_P) = 0.$$

Notice that $\frac{\partial G}{\partial \phi(x^*)} > 0$, and $\frac{\partial G}{\partial \delta_P} = 1 - \phi(x^*)h(\phi(x^*)) = \frac{Lh(\phi(x^*))}{1+\delta-\delta_P} > 0$ with the second equality being guaranteed by $G(\delta_P, \phi(x^*)) = 0$. By the Implicit function theorem, $\frac{\mathrm{d}\phi(x^*)}{\mathrm{d}\delta_P} = -\frac{\partial G}{\partial \delta_P} / \frac{\partial G}{\partial \phi(x^*)} < 0$. Since $x^* = \phi(x^*)(1 - \delta_P)$, we have x^* strictly decreases in δ_P. ■

Proof of Lemma 2

Proof. (i) Suppose for some principal type θ, we have $x_1^i > x_1^j$, $e_1^i = 0$ and $e_1^j = 1$. Let $\theta_L \leqslant \theta$ be the infimum of such types. The principal's total cost is at least $\theta + x_1^i + \delta_P(2\theta_L)$. By deviating to $e_1^i = 1$ and $e_1^j = 0$, the total cost is at most $\theta + x_1^j + \delta_P(2\theta)$. As $\theta \to \theta_L$, this deviation strictly lowers cost, a contradiction.

(ii) Depending on the principal's enforcement in period 1, agent i chooses potentially random misconduct level x_2^i in period 2 according to its posterior. Assume $x_1^A \geqslant x_1^B$ w.l.o.g. By disciplining both agents, one agent, and neither agent in period 1, the principal respectively incurs the following total costs as continuous functions of θ:

$$\begin{aligned} C_2(\theta) &\equiv 2\theta + \delta_P \sum_i \mathbb{E}\left[\min\left\{\theta, x_2^i(x_1^A, x_1^B, e_1^A = 1, e_1^B = 1)\right\} \middle| \theta\right], \\ C_1(\theta) &\equiv \theta + x_1^B + \delta_P \sum_i \mathbb{E}\left[\min\left\{\theta, x_2^i(x_1^A, x_1^B, e_1^A = 1, e_1^B = 0)\right\} \middle| \theta\right], \\ C_0(\theta) &\equiv x_1^A + x_1^B + \delta_P \sum_i \mathbb{E}\left[\min\left\{\theta, x_2^i(x_1^A, x_1^B, e_1^A = 0, e_1^B = 0)\right\} \middle| \theta\right]. \end{aligned}$$

Define $\theta^* \equiv \inf\{\theta \geqslant 0 : C_0(\tilde{\theta}) \leqslant \min\{C_1(\tilde{\theta}), C_2(\tilde{\theta})\},\ \forall\ \tilde{\theta} > \theta\}$ and $\theta^{**} \equiv \sup\{\theta \geqslant 0 : C_2(\tilde{\theta}) < \min\{C_0(\tilde{\theta}), C_1(\tilde{\theta})\},\ \forall\ 0 \leqslant \tilde{\theta} < \theta\}$. By definition, $\theta^{**} \geqslant 0$. Since $C_2(\theta) > C_1(\theta) > C_0(\theta)$ as $\theta \to \infty$, we know $\theta^* < \infty$ and $\Pr\{\theta : (e_1^A, e_1^B) = (0,0)\} > 0$.

We prove a single-crossing property such that if $C_1(\theta_0) - C_0(\theta_0) \geqslant 0$ for some θ_0, then it remains true for all $\theta > \theta_0$. Suppose towards contradiction that $C_1(\theta_0) - C_0(\theta_0) \geqslant 0$ and $C_1(\theta_1) - C_0(\theta_1) < 0$ for some $\theta_1 > \theta_0$. Since $\lim_{\theta\to\infty} C_1(\theta) - C_0(\theta) = \infty$, by continuity there exist $\theta' \in [\theta_0, \theta_1)$ and $\theta'' > \theta_1$ such that $C_1(\theta') - C_0(\theta') = C_1(\theta'') - C_0(\theta'') = 0$ and $C_1(\theta) - C_0(\theta) < 0$ for all $\theta \in (\theta', \theta'')$. This means a type $\theta \in (\theta', \theta'')$ never chooses $(e_1^A, e_1^B) = (0,0)$, but since $(e_1^A, e_1^B) = (0,0)$ is on equilibrium path, the agents assign zero conditional probability to $\theta \in (\theta', \theta'')$ upon seeing $(e_1^A, e_1^B) = (0,0)$. As a result, upon seeing $(e_1^A, e_1^B) = (0,0)$, x_2^i is never in the interval (θ', θ''), and thus for all $\theta \in (\theta', \theta'')$:

$$\begin{aligned} & C_1'(\theta) - C_0'(\theta) \\ = & 1 + \sum_i \delta_P \left(\Pr(x_2^i(x_1^A, x_1^B, e_1^A = 1, e_1^B = 0) \geqslant \theta \middle| \theta) - \Pr(x_2^i(x_1^A, x_1^B, e_1^A = 0, e_1^B = 0) \geqslant \theta \middle| \theta)\right) \\ = & 1 + \sum_i \delta_P \left(\Pr(x_2^i(x_1^A, x_1^B, e_1^A = 1, e_1^B = 0) \geqslant \theta \middle| \theta) - \Pr(x_2^i(x_1^A, x_1^B, e_1^A = 0, e_1^B = 0) \geqslant \theta'')\right), \end{aligned}$$

which decreases in θ wherever defined. This implies $C_1(\theta) - C_0(\theta)$ is concave on $\theta \in (\theta', \theta'')$, a contradiction to the definition of θ' and θ''. Similarly, one can show that the function $C_2(\theta) - C_0(\theta)$ have the same single-crossing property.

There are two possibilities regarding $C_1(\theta)$. First, consider the case where $\{\theta : (e_1^A, e_1^B) = (1,0)\} \neq \emptyset$. The above argument works for the function $C_2(\theta) - C_1(\theta)$ to show single-crossing property. Then by definition of θ^* and θ^{**} and the three single-crossing properties, we have the cutoff strategy described in the lemma. Also, since $\{\theta : (e_1^A, e_1^B) = (1,0)\} = [\theta^{**}, \theta^*) \neq \emptyset$, we must have $\theta^{**} < \theta^*$.

Second, consider the case where $\{\theta : (e_1^A, e_1^B) = (1,0)\} = \emptyset$. Then by definition of θ^* and θ^{**} and the single-crossing property of $C_2(\theta) - C_0(\theta)$, we have the cutoff strategy described in the lemma. Also, since $\{\theta : (e_1^A, e_1^B) = (1,0)\} = \emptyset$, we must have $\theta^{**} = \theta^*$. ■

Proof of Proposition 2

Proof. W.l.o.g., suppose $x_1^A \geqslant x_1^B$. (i) Let $(x_1^A, x_1^B) \in P(\delta_P)$. Suppose towards contradiction that $\theta^{**} < \theta^*$. According to Lemma 2, type-θ^{**} principal incurs cost $\theta^{**} + x_1^B + 2\delta_P\theta^{**}$ when punishing only agent A, cost $2\theta^{**} + 2\delta_P \cdot 0$ when punishing both, and cost $x_1^A + x_1^B + 2\delta_P\theta^{**}$ when punishing neither. By definition, we must have $\theta^{**} + x_1^B + 2\delta_P\theta^{**} = 2\theta^{**} + 2\delta_P \cdot 0 < x_1^A + x_1^B + 2\delta_P\theta^{**}$, which yields $x_1^A > \theta^{**} = \frac{x_1^B}{1-2\delta_P}$, contradicting the fact that $(x_1^A, x_1^B) \in P(\delta_P)$. Therefore, $\theta^{**} = \theta^*$. Then the indifference of type θ^* requires $2\theta^* + 2\delta_P \cdot 0 = x_1^A + x_1^B + 2\delta_P\theta^*$, i.e., $\theta^{**} = \theta^* = \frac{x_1^A + x_1^B}{2(1-\delta_P)} = \phi_{0|2}(x_1^A, x_1^B)$.

(ii) Let $(x_1^A, x_1^B) \notin P(\delta_P)$. Suppose towards contradiction that $\theta^{**} = \theta^*$. According to Lemma 2, type-θ^* principal incurs cost $2\theta^* + 2\delta_P \cdot 0$ when punishing both, and cost $x_1^A + x_1^B + 2\delta_P\theta^*$ when punishing neither. By definition, we must have $2\theta^* + 2\delta_P \cdot 0 = x_1^A + x_1^B + 2\delta_P\theta^*$, which yields $\theta^* = \phi_{0|2}(x_1^A, x_1^B)$. Instead, it incurs cost at most $\theta^* + x_1^B + 2\delta_P\theta^*$ when punishing only agent A, regardless of the agents' beliefs. Since $(\theta^* + x_1^B + 2\delta_P\theta^*) - (2\theta^* + 2\delta_P \cdot 0) = \frac{x_1^B - (1-2\delta_P)x_1^A}{2(1-\delta_P)} < 0$, we have a profitable deviation, contradicting the assumption $\theta^{**} = \theta^*$. Therefore, $\theta^{**} < \theta^*$. For type θ^{**} to be indifferent between punishing both and punishing agent A, we require $2\theta^{**} + 2\delta_P \cdot 0 = \theta^{**} + x_1^B + 2\delta_P\theta^{**}$, i.e., $\theta^{**} = \frac{x_1^B}{1-2\delta_P} = \phi_{1|2}(x_1^A, x_1^B)$. For type θ^* to be indifferent between punishing agent A and punishing neither, we need $\theta^* + x_1^B + 2\delta_P\theta^{**} = x_1^A + x_1^B + 2\delta_P\theta^*$, i.e., $\theta^* = \frac{x_1^A(1-2\delta_P) - x_1^B(2\delta_P)}{(1-2\delta_P)^2} = \phi_{0|1}(x_1^A, x_1^B)$. ■

Proof of Corollary 2

Proof. According to Proposition 2, a type-θ principal punishes both agents if and only if either $(x_1^A, x_1^B) \in P(\delta_P)$ and $\theta < \phi_{0|2}(x_1^A, x_1^B)$, or $(x_1^A, x_1^B) \notin P(\delta_P)$ and $\theta < \phi_{1|2}(x_1^A, x_1^B)$. Taking the union of the two cases, we have the condition $\theta < \phi_{0|2}(x_1^A, x_1^B)$ and $\theta < \phi_{1|2}(x_1^A, x_1^B)$.

A type-θ principal punishes one agent if and only if $(x_1^A, x_1^B) \notin P(\delta_P)$ and $\phi_{1|2}(x_1^A, x_1^B) \leqslant$

$\theta < \phi_{0|1}(x_1^A, x_1^B)$. Since the latter condition already implies $(x_1^A, x_1^B) \notin P(\delta_P)$, we only require $\phi_{1|2}(x_1^A, x_1^B) \leqslant \theta < \phi_{0|1}(x_1^A, x_1^B)$.

A type-θ principal punishes neither agent if and only if either $(x_1^A, x_1^B) \in P(\delta_P)$ and $\theta \geqslant \phi_{0|2}(x_1^A, x_1^B)$, or $(x_1^A, x_1^B) \notin P(\delta_P)$ and $\theta \geqslant \phi_{0|1}(x_1^A, x_1^B)$. Taking the union of the two cases, we have the condition $\theta \geqslant \phi_{0|2}(x_1^A, x_1^B)$ and $\theta \geqslant \phi_{0|1}(x_1^A, x_1^B)$. ■

Proof of Proposition 3

Proof. W.l.o.g., suppose $x_1^A \geqslant x_1^B$. Suppose towards contradiction that $\theta^{**} < \theta^*$. A type-θ^{**} principal's cost is $2\theta^{**}$ when punishing both agents, $\theta^{**} + x_1^B + 2\delta_P\theta^{**}$ when punishing only agent A, and $x_1^A + x_1^B + 2\delta_P\theta^{**}$ when punishing neither. Then the indifference condition implied by the definition of θ^{**} requires $2\theta^{**} = \theta^{**} + x_1^B + 2\delta_P\theta^{**}$, or equivalently,

$$x_1^B + (2\delta_P - 1)\theta^{**} = 0. \tag{16}$$

If $x_1^B > 0$, (16) is impossible, and hence we conclude $\theta^{**} = \theta^* = \phi_{0|2}(x_1^A, x_1^B)$.

If $x_1^B = 0$ and $\delta_P > \frac{1}{2}$, then (16) requires $\theta^{**} = 0$. The definition of θ^* requires $x_1^A + (2\delta_P - 1)\theta^* = 0$, which contradicts the assumption that $\theta^* > \theta^{**} \geqslant 0$. Therefore, again we conclude $\theta^{**} = \theta^* = \phi_{0|2}(x_1^A, x_1^B)$.

If $x_1^B = 0$ and $\delta_P = \frac{1}{2}$, then the definition of θ^* requires $x_1^A = \theta^{**}$. Fixing any type $\theta \in (\theta^{**}, \theta^*)$, we know $e_1^A(x_1^A, 0) = 1$ and $e_1^B(x_1^A, 0) = 0$. However, since $\phi_{0|2}(x_1^A, 0) = x_1^A = \theta^{**} < \theta$, for sufficiently small $\varepsilon > 0$ we have $\phi_{0|2}(x_1^A, \varepsilon) < \theta$. Then, $e_1^A(x_1^A, \varepsilon) = e_1^B(x_1^A, \varepsilon) = 0$, violating Assumption 2. We hence conclude $\theta^{**} = \theta^* = \phi_{0|2}(x_1^A, x_1^B)$. ■

Proof of Proposition 4

Proof. First, we show that when L is close enough to $\frac{1+\delta-\delta_P}{h(0)}$, a pair (x_1^A, x_1^B) in the interior of the proximity cone is not part of an equilibrium. Suppose towards contradiction that it is, then FOC of (10) w.r.t. x_1^A and x_1^B yields $x_1^A = x_1^B = x^{**}$, where x^{**} satisfies (13). Then, agent i's payoff simplifies to $U\left(\frac{x^{**}}{1-\delta_P}, L\right)$, where $U(\theta, L) \equiv (1+\delta-\delta_P)\theta(1-F(\theta)) - LF(\theta)$. Since $\frac{\partial U(\theta,(1+\delta-\delta_P)/h(0))}{\partial\theta} = -(1-F(\theta))(1+\delta-\delta_P)\left(\frac{h(\theta)}{h(0)} - 1 + \theta h(\theta)\right) < 0$ for all $\theta > 0$, by

Mean value theorem, $U\left(\frac{x^{**}}{1-\delta_P}, \frac{1+\delta-\delta_P}{h(0)}\right) < U\left(0, \frac{1+\delta-\delta_P}{h(0)}\right) = 0$. Because (13) defines x^{**} as a continuous function of L, we know $U\left(\frac{x^{**}}{1-\delta_P}, L\right) < 0$ for $L \in \left(\tilde{L}(\delta_P), \frac{1+\delta-\delta_P}{h(0)}\right)$ for some $\tilde{L}(\delta_P) < \frac{1+\delta-\delta_P}{h(0)}$. However, by choosing $x_1^i = 0$, agent i can secure a non-negative payoff from (11), a profitable deviation.

Next, we show that equilibrium cannot feature $x_1^A \neq x_1^B$. Suppose towards contradiction that $x_1^A > x_1^B$ in equilibrium. Since (x_1^A, x_1^B) is not in the interior of the proximity cone, it must be $x_1^A \geqslant \frac{x_1^B}{1-2\delta_P}$. In this regime, (12)'s derivative w.r.t. x_1^A is

$$\frac{1-F\left(\phi_{0|1}\right)}{1-2\delta_P}\left[(1+\delta-2\delta_P) - \left(L + x_1^A + \delta(\phi_{0|1} - \phi_{1|2})\right) h\left(\phi_{0|1}\right)\right]. \tag{17}$$

Given the definition of $\phi_{0|1}$ and $\phi_{1|2}$, the expression in the brackets is decreasing in x_1^A.When evaluated at $x_1^A = \frac{x_1^B}{1-2\delta_P}$, the expression becomes $(1+\delta-2\delta_P) - \left(L + \frac{x_1^B}{1-2\delta_P}\right) h\left(\frac{x_1^B}{1-2\delta_P}\right) \leqslant 1+\delta-2\delta_P - Lh(0) < 0$ when $L > \frac{1+\delta-2\delta_P}{h(0)}$. That is, if $L > \frac{1+\delta-2\delta_P}{h(0)}$, optimization of agent A requires $x_1^A = \frac{x_1^B}{1-2\delta_P}$. To be consistent with an equilibrium, agent B must be willing to choose $x_1^B = x_1^A(1-2\delta_P)$. However, the derivative of (11) w.r.t. x_1^B, evaluated at $x_1^B = x_1^A(1-2\delta_P)$, is $-\frac{1-F(x_1^A)}{(1-2\delta_P)^2}((1+\delta-2\delta_P)(2\delta_P) + (1-2\delta_P)((L+(1+\delta-2\delta_P)x_1^A)h(x_1^A) - 1)) < 0$, so that agent B will deviate to some lower x_1^B.

Finally, the only remaining possibility is $x_1^A = x_1^B = 0$, and we verify it as an equilibrium. Given $x_1^B = 0$, agent A's payoff is described by (12), with derivative described by (17). Since $x_1^B = 0$, we have $\phi_{0|1} = \frac{x_1^A}{1-2\delta_P}$ and $\phi_{1|2} = 0$. When $L = \frac{1+\delta-2\delta_P}{h(0)}$, (17) evaluated at $x_1^A = 0$ is zero. Therefore, for all larger L, $x_1^A = 0$ is the best response for agent A. The same applies to x_1^B.

For all above arguments to hold, we require $L > \hat{L}(\delta_P) \equiv \max\left\{\tilde{L}(\delta_P), \frac{1+\delta-2\delta_P}{h(0)}\right\}$. Since $x_1^i = 0$ in period 1, no agent is disciplined, and both choose zero misconduct in period 2. ■

Proof of Corollary 3

Proof. When $\delta < \frac{1}{2}$ and $L \in \left(\hat{L}(\delta_P), \frac{1+\delta-\delta_P}{h(0)}\right)$, with transparent enforcement the agents choose misconduct level 0 in period 1, and again 0 in period 2, according to Proposition 4. With opaque enforcement, the agents choose misconduct level $x^* > 0$ in period 1 that satisfies (6), and chooses either $\phi(x^*)$ or 0 in period 2 depending on whether $\theta \geqslant \phi(x^*)$ or not. Therefore, in either period, the expected misconduct is higher when enforcement is opaque.

With transparent enforcement, a type-θ principal's expected total cost is 0. With opaque enforcement, a type-θ principal's expected total discounted cost simplifies to $\min\{2\theta, 2\phi(x^*)\}$, which is larger than zero for all $\theta > 0$. ■

Proof of Proposition 5

Proof. For $\delta_P \geqslant \frac{1}{2}$, any (x_1^A, x_1^B) is in the proximity cone. The FOC of (10) w.r.t. x_1^i requires

$$(2 + \delta - 2\delta_P) - \left(L + x_1^i + \delta\phi_{0|2}\right) h\left(\phi_{0|2}\right) \leqslant 0, \tag{18}$$

and equality holds if $x_1^i > 0$. The left-hand side of (18) is decreasing in x_1^i, and it diverges to $-\infty$ as $x_1^i \to \infty$. Therefore, given any $x_1^{-i} \geqslant 0$, there exists a unique best response $x_1^i(x_1^{-i}) \geqslant 0$ of agent i. If the best response $x_1^i(x_1^{-i}) > 0$, the Implicit function theorem implies that

$$\frac{\mathrm{d}x_1^i}{\mathrm{d}x_1^{-i}} = -1 + \frac{2(1-\delta_P)h\left(\phi_{0|2}\right)^2}{(2+\delta-2\delta_P)\left(h\left(\phi_{0|2}\right)^2 + h'\left(\phi_{0|2}\right)\right)} \in (-1, 0).$$

If the best response $x_1^i(x_1^{-i}) = 0$ for some x_1^{-i}, then due to the fact that the left-hand side of (18) decreases in x_1^{-i}, we have $x_1^i(x_1^{-i}) = 0$ for all larger x_1^{-i}. Consequently, the best response functions intersect at most once, ensuring uniqueness of the equilibrium. Guessing $x_1^A, x_1^B > 0$, the FOC's hold with equality. Solving the system, we have $x_1^A = x_1^B \equiv x^{**}$, where x^{**} satisfies (13). The left-hand side of (13) strictly increases in x^{**}. At $x^{**} = 0$, the left-hand side becomes $Lh(0)$, which is small than the right-hand side $2 + \delta - 2\delta_P$ according to Assumption 1. As $x^{**} \to \infty$, the left-hand side diverges to $\infty > 2 + \delta - 2\delta_P$. Therefore, (13) has a unique solution $x^{**} > 0$ by the Intermediate value theorem, and our guess of the equilibrium is verified. The misconduct in period 2 follows from Lemma 1. ■

Proof of Corollary 4

Proof. When $\delta_P \geqslant \frac{1}{2}$, with transparent enforcement the agents choose misconduct level x^{**} in period 1 that satisfies (13), according to Proposition 5. In period 2, misconduct is either $\phi_{0|2}(x^{**}, x^{**})$ or 0, depending on whether $\theta \geqslant \phi_{0|2}(x^{**}, x^{**})$ or not. The expected misconduct in period 2 is $\left(1 - F\left(\phi_{0|2}(x^{**}, x^{**})\right)\right) \phi_{0|2}(x^{**}, x^{**})$.

With opaque enforcement, the agents choose misconduct level $x^* > 0$ in period 1 that satisfies (6), and chooses either $\phi(x^*)$ or 0 in period 2 depending on whether $\theta \geqslant \phi(x^*)$ or not. The expected misconduct in period 2 is $(1 - F(\phi(x^*)))\,\phi(x^*)$.

Since $2 + \delta - 2\delta_P > 1 + \delta - \delta_P$ and the function $(L + (1 + \delta - \delta_P)\theta)\,h(\theta)$ is increasing in x, we know $\phi_{0|2}(x^{**}, x^{**}) > \phi(x^*)$, i.e., $x^{**} > x^*$, so that misconduct is higher in period 1 when enforcement is transparent. For period 2, notice that the function $(1 - F(\theta))\theta$ has derivative $(1 - F(\theta))(1 - \theta h(\theta))$. Since

$$\begin{aligned} 2 + \delta - 2\delta_P &= \left(L + (1 + \delta - \delta_P)\phi_{0|2}(x^{**}, x^{**})\right) h\left(\phi_{0|2}(x^{**}, x^{**})\right) \\ &\geqslant Lh(0) + (1 + \delta - \delta_P)\phi_{0|2}(x^{**}, x^{**}) h\left(\phi_{0|2}(x^{**}, x^{**})\right) \\ &> 1 - \delta_P + (1 + \delta - \delta_P)\phi_{0|2}(x^{**}, x^{**}) h\left(\phi_{0|2}(x^{**}, x^{**})\right), \end{aligned}$$

we know that $\phi_{0|2}(x^{**}, x^{**}) h\left(\phi_{0|2}(x^{**}, x^{**})\right) < 1$, implying that the derivative $(1 - F(\theta))(1 - \theta h(\theta))$ is positive for $\theta = \phi_{0|2}(x^{**}, x^{**})$. According to Assumption 1, $(1 - F(\theta))(1 - \theta h(\theta))$ is also positive all $\theta \leqslant \phi_{0|2}(x^{**}, x^{**})$. Therefore, $(1 - F(\theta))\,\theta$ is lower if evaluated at $\theta = \phi(x^*)$ than if evaluated at a higher level $\theta = \phi_{0|2}(x^{**}, x^{**})$, and misconduct in period 2 is higher when enforcement is transparent.

With transparent enforcement, a type-θ principal's expected total discounted cost simplifies to $\min\{2\theta, 2\phi_{0|2}(x^{**}, x^{**})\}$. With opaque enforcement, a type-θ principal's expected total discounted cost is $\min\{2\theta, 2\phi(x^*)\}$. Since $\phi_{0|2}(x^{**}, x^{**}) > \phi(x^*)$, the former cost is weakly higher than the latter, and strictly higher if $\theta > \phi(x^*)$. ■

Proof of Corollary 5

Proof. When $\delta_P < \frac{1}{2}$ and $L > \hat{L}(\delta_P)$, the conclusion follows from Proposition 4 and the proof of Corollary 3. When $\delta_P = \frac{1}{2}$, the conclusion follows from Proposition 5 and the proof of Corollary 4.

When $\delta_P > \frac{1}{2}$, (13) can be rewritten as:

$$G(\delta_P, \phi_{0|2}(x^{**}, x^{**})) \equiv (L + (1 + \delta - \delta_P)\phi_{0|2}(x^{**}, x^{**}))h(\phi_{0|2}(x^{**}, x^{**})) - (2 + \delta - 2\delta_P) = 0.$$

Notice that G strictly increases in $\phi_{0|2}(x^{**}, x^{**})$, and $\frac{\partial G}{\partial \delta_P} = 2 - \phi_{0|2}(x^{**}, x^{**})h(\phi_{0|2}(x^{**}, x^{**})) = \frac{Lh(\phi_{0|2}(x^{**}, x^{**})) + \delta}{1 + \delta - \delta_P} > 0$. By the Implicit function theorem, $\phi_{0|2}(x^{**}, x^{**})$ strictly decreases in δ_P. Therefore, x^{**}, which is equal to $\phi_{0|2}(x^{**}, x^{**})(1 - \delta_P)$, also strictly decreases in δ_P.

The expected misconduct in period 2 is $\left(1 - F\left(\phi_{0|2}(x^{**}, x^{**})\right)\right)\phi_{0|2}(x^{**}, x^{**})$ according to the proof of Corollary 4. From $G(\delta_P, \phi_{0|2}(x^{**}, x^{**})) = 0$, we have:

$$\phi_{0|2}(x^{**}, x^{**})h(\phi_{0|2}(x^{**}, x^{**})) = \frac{2 + \delta - 2\delta_P - Lh(\phi_{0|2}(x^{**}, x^{**}))}{1 + \delta - \delta_P} < \frac{1 + \delta - 2\delta_P}{1 + \delta - \delta_P} < 1.$$

Since the function $(1 - F(\theta))\theta$ has derivative $(1 - F(\theta))(1 - \theta h(\theta))$, which is positive whenever $\theta \leqslant \phi_{0|2}(x^{**}, x^{**})$. Therefore, a higher δ_P, leading to a lower $\phi_{0|2}(x^{**}, x^{**})$, also lowers the expected misconduct in period 2. ■

Proof of Lemma 3

Proof. If the posterior belief is the prior truncated to $[\theta_L, \theta_H)$, then agent i, who chooses misconduct $x_2^i \in [\theta_L, \theta_H)$, receives a continuation payoff:

$$x_2^i \frac{F(\theta_H) - F(x_2^i)}{F(\theta_H) - F(\theta_L)} - L\frac{F(x_2^i) - F(\theta_L)}{F(\theta_H) - F(\theta_L)}.$$

Differentiating w.r.t. x_2^i, we have $\frac{1 - F(x_2^i)}{F(\theta_H) - F(\theta_L)}\left(\frac{F(\theta_H) - F(x_2^i)}{1 - F(x_2^i)} - (L + x_2^i)h(x_2^i)\right)$. KKT condition requires $\frac{F(\theta_H) - F(x_2^i)}{1 - F(x_2^i)} - (L + x_2^i)h(x_2^i) \leqslant 0$, with equality if $x_2^i = \theta_L$. Since the left-hand side is decreasing in x_2^i, this condition is also sufficient. ■

References

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991). Optimal learning by experimentation. *Review of Economic Studies 58*(4), 621–654.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, 719–752.

Bergemann, D. and J. Välimäki (2000). Experimentation in markets. *Review of Economic Studies 67*(2), 213–234.

Bergemann, D. and J. Välimäki (2006). Bandit problems. Cowles Foundation discussion paper.

Bernardo, A. E., H. Cai, and J. Luo (2001). Capital budgeting and compensation with asymmetric information and moral hazard. *Journal of Financial Economics 61*(3), 311–344.

Bolton, P. and C. Harris (1999). Strategic experimentation. *Econometrica 67*(2), 349–374.

Bonatti, A. and J. Hörner (2011). Collaborating. *American Economic Review 101*, 632–663.

Bond, P. and K. Hagerty (2010). Preventing crime waves. *American Economic Journal: Microeconomics 2*(3), 138–59.

Boot, A. W. and A. V. Thakor (1993). Self-interested bank regulation. *American Economic Review 83*(2), 206–212.

Burkart, M. and A. Dasgupta (2021). Competition for flow and short-termism in activism. *Review of Corporate Finance Studies 10*(1), 44–81.

Chen, Y. (2020). A revision game of experimentation on a common threshold. *Journal of Economic Theory 186*, 104997.

Dasgupta, A. and G. Piacentino (2015). The wall street walk when blockholders compete for flows. *Journal of Finance 70*(6), 2853–2896.

Donelson, D. C., R. W. Flam, and C. G. Yust (2022). Spillover effects in disclosure-related securities litigation. *The Accounting Review 97*(5), 275–299.

Drake, G. (2017). Navigating the Atlantic: Understanding EU data privacy compliance amidst a sea of uncertainty. *Southern California Law Review 91*, 163.

Edmans, A., D. Levit, and D. Reilly (2019). Governance under common ownership. *Review of Financial Studies 32*(7), 2673–2719.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy 81*(3), 521–565.

Fan, M., L. Yu, S. Chen, H. Zhou, X. Luo, S. Li, Y. Liu, J. Liu, and T. Liu (2020). An empirical evaluation of GDPR compliance violations in Android mHealth apps. In *2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)*, pp. 253–264. IEEE.

Gantchev, N., O. R. Gredil, and C. Jotikasthira (2019). Governance under the gun: Spillover effects of hedge fund activism. *Review of Finance 23*(6), 1031–1068.

Gilje, E. P., T. A. Gormley, and D. Levit (2020). Who's paying attention? measuring common ownership and its impact on managerial incentives. *Journal of Financial Economics 137*(1), 152–178.

Goldstein, I. and Y. Leitner (2020). Stress tests disclosure: Theory, practice, and new perspectives. In J. D. Farmer, A. M. Kleinnijenhuis, T. Schuermann, and T. Wetzer (Eds.), *Handbook of Financial Stress Testing*. Cambridge University Press.

Goldstein, I. and H. Sapra (2013). Should banks' stress test results be disclosed? An analysis of the costs and benefits. *Foundations and Trends in Finance 8*(1), 1–54.

Harris, M. and A. Raviv (1998). Capital budgeting and delegation. *Journal of Financial Economics 50*(3), 259–289.

Hörner, J. and A. Skrzypacz (2016). Learning, experimentation and information design. In *Advances in Economics and Econometrics: Eleventh World Congress*, Volume 1, pp. 63–98.

Huang, C. (2017). Defending against speculative attacks: The policy maker's reputation. *Journal of Economic Theory 171*, 1–34.

Keller, G., S. Rady, and M. Cripps (2005). Strategic experimentation with exponential bandits. *Econometrica 73*(1), 39–68.

Lakkis, E. (2021). Threat of hedge fund activism and risky investment. *Available at SSRN 3868341*.

Lee, D. W. and K. S. Park (2009). Does institutional activism increase shareholder wealth? evidence from spillovers on non-target companies. *Journal of Corporate Finance 15*(4), 488–504.

Malenko, A. (2019). Optimal dynamic capital budgeting. *Review of Economic Studies 86*(4), 1747–1778.

Manso, G. (2011). Motivating innovation. *Journal of Finance 66*(5), 1823–1860.

Morrison, A. D. and L. White (2013). Reputational contagion and optimal regulatory forbearance. *Journal of Financial Economics 110*(3), 642–658.

Nanda, R. and M. Rhodes-Kropf (2017). Financing risk and innovation. *Management Science 63*(4), 901–918.

Rothschild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory 9*(2), 185–202.

Shapiro, J. and D. Skeie (2015). Information management in banking crises. *Review of Financial Studies 28*(8), 2322–2363.

Song, F. (2019). Blockholder short-term incentives, structures, and governance. European Corporate Governance Institute (ECGI)-Finance Working Paper.