

Everlasting Fraud

Law Working Paper N° 617/2021

Revised January 2026

Vivian W. Fang

Indiana University and ECGI

Wenyu Wang

Kelley School of Business, Indiana University

Nan Li

University of Minnesota

Gaoqing Zhang

Carnegie Mellon University

© Vivian W. Fang, Wenyu Wang, Nan Li and Gaoqing Zhang 2026. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
<https://doi.org/10.67929/ecgi-law-617-2021>

ECGI Working Paper Series in Law

Everlasting Fraud

Working Paper N° 617/2021
January 2026

Vivian W. Fang
Wenyu Wang
Nan Li
Gaoqing Zhang

Abstract

This paper studies the dynamic interaction between corporate fraud and regulatory detection. We develop a theoretical framework in which managerial manipulation endogenously reduces the informativeness of earnings reports, while a regulator chooses detection effort to restore it. As manipulation goes undetected, information uncertainty accumulates and investors rely more heavily on the new earnings report; detection resolves fraud-induced uncertainty but resets the cycle. In multi-firm settings, the regulator reallocates attention based on relative fraud-induced uncertainty, yielding mean-reverting manipulation and, depending on parameters, either synchronized or sequential fraud patterns. We then structurally estimate the model to identify key parameters governing fraud accumulation, detection intensity, and uncertainty recovery. Finally, we present reduced-form evidence consistent with the model: earnings response coefficients drop sharply upon fraud revelation and gradually recover thereafter.

Keywords: Earnings Manipulation, Corporate Fraud, Regulation

JEL Classification: G32, G38, M40, M48

Vivian W. Fang

Indiana University and ECGI
[ORCID 0000-0001-7868-1393](https://orcid.org/0000-0001-7868-1393)

Wenyu Wang

Kelley School of Business, Indiana University

Nan Li

University of Minnesota

Gaoqing Zhang

Carnegie Mellon University

Everlasting Fraud *

Vivian W. Fang

Nan Li

Indiana University, ECGI

University of Minnesota

Wenyu Wang

Gaoqing Zhang

Indiana University

Carnegie Mellon University

January 31, 2026

Abstract

This paper studies the dynamic interaction between corporate fraud and regulatory detection. We develop a theoretical framework in which managerial manipulation endogenously reduces the informativeness of earnings reports, while a regulator chooses detection effort to restore it. As manipulation goes undetected, information uncertainty accumulates and investors rely more heavily on the new earnings report; detection resolves fraud-induced uncertainty but resets the cycle. In multi-firm settings, the regulator reallocates attention based on relative fraud-induced uncertainty, yielding mean-reverting manipulation and, depending on parameters, either synchronized or sequential fraud patterns. We then structurally estimate the model to identify key parameters governing fraud accumulation, detection intensity, and uncertainty recovery. Finally, we present reduced-form evidence consistent with the model: earnings response coefficients drop sharply upon fraud revelation and gradually recover thereafter.

JEL Classification: G32, G38, M40, M48

Keywords: Earnings Manipulation, Corporate Fraud, Regulation

*We are grateful to helpful comments from Christian Leuz (Editor), an anonymous referee, Associate Editor, Jeremy Bertomeu, Philip Bond, Sylvain Chassang, Rebecca Hann, Xu Jiang, Jon Karpoff, Lorien Stice-Lawrence, Ivan Marinovic, Valeri Nikolaev, Hao Xue, Liyan Yang, Anastasia Zakolyukina, Frank Zhou, seminar and conference participants at the Accounting and Economics Society webinar, China International Conference in Finance, Chinese University of Hong Kong, Columbia Burton Accounting Conference, Duke University, FARS meeting, HKUST Accounting Research Symposium, MFA meeting, Rice University, SEC Conference on Financial Market Regulation, University of Maryland, University of Miami, University of Science and Technology of China, and University of Texas, Austin. Please send correspondence to vfang@iu.edu.

The central regulatory concern about corporate fraud is informational: earnings manipulation not only biases reported performance but also erodes the informativeness of earnings reports, distorting how investors learn about firm fundamentals and price the firm. Regulatory intervention is therefore often framed as an effort to restore the informativeness of financial reporting. This paper studies the dynamic interaction between managerial manipulation and regulatory detection. Our main contribution is to characterize an accumulation-and-reset cycle for each firm: undetected manipulation builds information uncertainty and increases the market’s reliance on earnings news in pricing the firm, while detection temporarily resolves that uncertainty and resets the cycle. Across firms, regulatory reallocation of detection resources shapes how these accumulation-and-reset cycles overlap across firms, generating cross-sectional patterns in fraud and detection.

We begin with a single-firm model with a representative firm, a regulator, and competitive investors. In period t , the manager privately observes the firm’s economic earnings and issues a potentially biased earnings report. Investors update beliefs about fundamentals and price the firm based on the report. The regulator inspects the report using a detection technology that uncovers fraud with an endogenous probability and, if successful, reveals the firm’s true earnings to the market. To keep the model parsimonious and the economic mechanism transparent, we assume both parties are myopic: managerial manipulation affects the stock price and regulatory detection shapes the information environment in the current period, but neither party internalizes how current actions affect future payoffs.

The key state variable is cumulative fraud-induced information uncertainty carried into period t , denoted Φ_{t-1} . Undetected manipulation lowers the information content of past reports and increases Φ_{t-1} over time. If fraud is detected, true earnings are revealed and Φ_{t-1} collapses, restoring reporting informativeness. By summarizing the information consequences of past manipulation, Φ_{t-1} links the incentives of the manager and the regulator.

The manager chooses the manipulation amount in period t to boost the firm’s market value. A higher Φ_{t-1} makes the period- t report more value relevant because less informative

past reports lead investors to place greater weight on the current report. At the same time, a higher Φ_{t-1} raises the manager’s expected detection risk, as he anticipates that the regulator will intensify scrutiny when fraud-induced uncertainty is high. Indeed, the regulator optimally increases detection intensity with Φ_{t-1} because detecting fraud yields larger information benefits when uncertainty is higher, even though stronger enforcement requires more regulatory resources. In equilibrium, manipulation and detection are jointly determined by Φ_{t-1} : manipulation increases uncertainty, higher uncertainty induces stronger detection, and detection risk feeds back into manipulation incentives.

The single-firm model delivers the basic intuition of the accumulation-and-reset manipulation cycle. We then extend the model to a two-firm setting to study the cross-sectional dynamics of fraud and detection. The regulator allocates detection effort based on firms’ *relative* cumulative fraud-induced uncertainty, reflecting differences in their current levels of undetected manipulation. We label the firm with higher (lower) uncertainty as the H-firm (L-firm). Because detection resources are limited, the regulator concentrates scrutiny on the H-firm, leaving the L-firm with little or no detection. Concentrated scrutiny at the H-firm disciplines its manipulation, while strengthening manipulation incentives at the L-firm and allowing uncertainty to build there. As relative uncertainty shifts, regulatory attention rotates between firms. The two-firm model delivers an additional insight beyond the single-firm setting: firms’ manipulation choices become interdependent because scrutiny directed at one firm necessarily reduces detection at the other. These spillovers are contemporaneous, as one firm’s manipulation choice alters the other’s detection risk within the same period through the regulator’s reallocation of attention.¹

The framework outlined above generates testable predictions about the evolution of fraud-induced uncertainty, the information effects of detection, and the role of regulatory reallocation across firms. We then take the model to the data to assess its empirical relevance.

¹This allocation problem is consistent with the fact that most regulators face limited enforcement resources, so they have to decide which firms to prioritize in detecting fraud (see, e.g., Holzman et al. (2024)). However, persistent fraud in our model does not require high enforcement costs: when expected fraud is low, the regulator may optimally exert little effort because the information gains from detection are limited.

Because fraud-induced uncertainty is not directly observable, we use the earnings response coefficient (ERC), a standard measure of earnings-report value relevance (e.g., Fang et al. (2017)), as a proxy. We test the model predictions in two complementary ways.

We first structurally estimate a three-firm, fully forward-looking setting of the model to identify parameters governing fraud accumulation, detection intensity, and uncertainty recovery. Relative to the myopic benchmarks, this setting adds empirical realism in two ways. First, it delivers quantitative effects closer to those in the data, while the myopic benchmarks may understate the magnitudes. As such, it generates a possible non-monotonic manipulation–uncertainty relationship that cannot arise under myopia. When both the manager and the regulator internalize how their current actions affect future payoffs, equilibrium manipulation peaks at an intermediate level of fraud-induced uncertainty. Intuitively, the manager’s incentive to manipulate is muted when uncertainty is low because the gains are small, and it is also muted when uncertainty is high because expected detection costs are high.

Second, this non-monotonicity motivates a setting that allows for asymmetric, endogenously rotating regulatory attention across high-, medium-, and low-uncertainty firms, which requires at least three firms; in contrast, the two-firm setting implies only symmetric back-and-forth rotation with a single peer. The three-firm setting is the smallest environment that generates such rotations, and the resulting overlapping accumulation-and-reset cycles likely better match observed fraud dynamics. Adding more firms sharply increases computational cost with limited gains for identification. In a way, one can interpret the three-firm setting as a stylized simplification that characterizes the broader economy as a triad of close peers: when the regulator examines a focal firm, it reallocates detection effort within that peer group rather than across the full corporate universe, consistent with enforcement scrutiny often concentrating among similar firms.²

As before, we label firms with high, medium, and low cumulative fraud-induced uncer-

²As anecdotal evidence, following the WorldCom scandal in 2002, regulators increased scrutiny of other telecommunications firms with similar revenue-recognition practices, and after the 2008 crisis they investigated financial institutions with comparable underwriting and securitization practices.

tainty as the H-, M-, and L-firm. The three-firm setting generates interesting detection paths and fraud patterns: scrutiny concentrates on the H-firm, disciplining manipulation there while allowing the M- and L-firms to manipulate and accumulate uncertainty; detection then collapses uncertainty at the H-firm and shifts attention to the next-highest-uncertainty firm, allowing the previously disciplined firm to reenter the manipulation cycle. These dynamics produce predictable, overlapping accumulation-and-reset manipulation cycles even without aggregate shocks, with fraud appearing sequential when detection outpaces accumulation and more synchronized when accumulation outpaces detection. This mechanism is therefore distinct from aggregate regulatory “waves” that may arise from political forces or shifts in macroeconomic conditions (e.g., Bertomeu and Magee (2011); Hail et al. (2018)).

The estimation targets moments related to detection frequency, cumulative fraud at detection, and the speed at which information uncertainty rebuilds after detection. Results show that the model-implied moments align well with the empirical counterparts in terms of the frequency of detection (with an average rate of 0.7% per quarter for firms in the lowest tercile of uncertainty and a higher rate of 0.9% per quarter for firms in higher terciles), cumulative fraud levels upon detection (amounting to 11% of total assets), and the duration of uncertainty rebound post-detection (averaging 6 quarters).

Additionally, parameter estimates suggest that managers face substantial expected penalties upon detection, amounting to nearly five times the illegal gains from manipulation. Importantly, this estimate reflects the manager’s total expected cost, including legal, monetary, reputational, and career losses. In contrast, regulators face relatively low detection costs and focus enforcement on firms with high information uncertainty, reinforcing the model’s intuition that persistent fraud is not purely driven by regulatory budget constraints. Intriguingly, despite the estimation revealing high manipulation costs for managers and low detection costs for regulators—factors that would typically discourage fraud—the model indicates persistent fraud, as we observe in the data.³ Lastly, our estimates indicate that synchronized fraud is

³The finding that managers face high manipulation costs is consistent with prior structural work including Zakolyukina (2018), Bird et al. (2019), Bertomeu et al. (2021), Terry (2023), and Cheynel et al. (2024).

the more empirically relevant manifestation of firms’ overlapping accumulation-and-reset cycles. Following detection, reduced regulatory scrutiny allows firms to resume manipulation, leading fraud levels to converge across firms within approximately six quarters.

Finally, we present reduced-form evidence that the average ERC drops sharply upon fraud revelation and gradually recovers thereafter. This pattern is consistent with detection temporarily resolving fraud-induced uncertainty, after which uncertainty rebuilds as manipulation reaccumulates. In contrast, matched firms without detection exhibit no comparable decline-and-rebound pattern in the ERC.

To our best knowledge, this paper is the first to develop a dynamic theory of earnings manipulation with endogenous detection. We construct and estimate a model in which managers rationally choose manipulation by trading off valuation benefits against detection risk, while a regulator optimally allocates detection resources to maximize the informativeness of financial reporting. A key insight is that persistent fraud can arise endogenously because the regulator’s optimal detection policy is state-dependent and feeds back into managers’ manipulation incentives over time and across firms. By endogenizing both manipulation and detection, our study stands in contrast to most earnings manipulation theories, which either endogenize manipulation but treat detection as exogenous or implicit or endogenize detection in static or reputational settings without endogenous manipulation.

Indeed, theories that endogenize detection are rare. The closest related paper is Chen et al. (2025), which develops a structural model of manipulation and detection motivated by regulatory policy. Their model differs from ours in two key respects. First, their model is static, whereas ours is dynamic, allowing fraud, information uncertainty, and enforcement incentives to co-evolve over time. Second, the marginal benefit of manipulation is exogenous in their model and recovered from the data, whereas in ours it is endogenous and state-dependent, evolving with cumulative fraud through its effect on information uncertainty and investors’ reliance on current reports.

Relatedly, Cheynel et al. (2024) develop a dynamic model of manipulation, but the mech-

anism and role of detection differ from ours. In their model, manipulation escalates through a learning channel: repeated nondetection leads managers to infer weaker regulatory monitoring and therefore to increase fraud. Moreover, detection effort itself is not an endogenous choice in their model; instead, what evolves endogenously are managers’ beliefs about monitoring effectiveness, which are updated based on past nondetections.

Our paper also contributes to the literature that structurally estimates earnings manipulation. Existing studies typically abstract from an explicit regulatory optimization problem. For example, Zakolyukina (2018) studies a finite-horizon model of manipulation of book value, motivated by the puzzle that observed restatements are rare relative to suspected manipulation. She estimates the frequency and magnitude of undetected manipulation and shows how optimal manipulation varies with the manager’s horizon, generating distinctive tenure patterns. Beyer et al. (2019) build a dynamic manipulation model in which managers face manipulation costs and internalize both short- and long-term incentives when reporting. Their model features persistent information asymmetry driven by exogenous reporting noise; unlike in our model, strategic manipulation does not reduce the informativeness of reports. The goal of their structural estimation is to disentangle fundamental economic uncertainty from investor uncertainty arising from reporting distortions using the time-series properties of earnings. Bertomeu et al. (2021) develop a signaling model with a myopic manager and no regulator. In estimating the model, they first recover the market’s valuation sensitivity to reported earnings (similar to ERC) nonparametrically from the data and then use this object to infer manipulation costs and the level of manipulation. In all three papers, detection is either absent or captured in reduced form through exogenous detection probabilities or manipulation costs.

More tangentially related, Bird et al. (2019) and Terry (2023) study the bunching of reported earnings just above analyst targets. Specifically, Bird et al. (2019) develop a single-period model of benchmark meeting in which managers trade off the valuation benefits of narrowly meeting expectations against manipulation costs, and use a regression discon-

tinuity design together with the earnings-surprise distribution to recover the implied extent of manipulation; their setting is static and abstracts from regulatory response. Terry (2023) embeds target-driven earnings manipulation into a quantitative general-equilibrium endogenous-growth model to study how short-term pressure disciplines managers while incentivizing R&D cuts; the dynamics in his model operate through investment and growth rather than through a strategic interaction between manipulation and detection.⁴

1 Model Framework

1.1 Single Firm

1.1.1 Model Setup

We consider a baseline setting in which a representative firm generates economic earnings s_t in each period $t \in \{1, 2, \dots, \infty\}$. We assume that s_t follows an AR(1) process such that

$$s_t = \rho s_{t-1} + \varepsilon_t, \quad (1)$$

where the correlation coefficient $\rho \in (0, 1)$ and the random variable $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. In each period, the firm manager privately learns the realization of the firm's economic earnings s_t and issues a report r_t . Investors use the report to update V_t , which is assumed to be set by a competitive market and equals the firm's total discounted future earnings in expectation:

$$V_t = \sum_{k=t}^{\infty} \delta^{k-t} E^I[s_k | \mathcal{F}_t] = \frac{E^I[s_t | \mathcal{F}_t]}{1 - \delta\rho}, \quad (2)$$

⁴Some papers are more remotely related in that they study disclosure or monitoring in settings that differ fundamentally from our focus on endogenous fraud and detection. To highlight two, Bourveau et al. (2025) study strategic disclosure under bounded rationality, focusing on how firms' depths of reasoning shape disclosure behavior rather than manipulation. Corona and Randhawa (2018) analyze a dynamic inspection and reputation framework in which monitoring responds to past outcomes, but misconduct is exogenous and the strategic choice concerns disclosure rather than the intensity of fraud.

where $E^I[\cdot|\mathcal{F}_t]$ denotes the investors' expectation, $\mathcal{F}_t \equiv \{r_t, r_{t-1}, \dots, r_1\}$ denotes the set of the firm's reports up to time t , and $\delta \in (0, 1)$ denotes the investors' discount factor.

We model the manager's earnings manipulation decision following Gao and Zhang (2019). In each period t , after observing the true economic earnings s_t , the manager chooses manipulation $m_t \geq 0$ that generates the following report r_t :

$$r_t = s_t + m_t(1 - q) + \sqrt{m_t q (1 - q)} \eta_t, \quad (3)$$

where $q \in (0, 1]$ is the probability that the firm's internal control system successfully prevents the manager's manipulation affecting the report, and $\eta_t \sim N(0, 1)$ is a standard normal random variable independent of all other variables in the model.⁵ The choice of manipulation m_t is observable only to the manager. Equation (3) suggests that manipulation has a dual effect on the report: m_t increases the mean of the report but decreases its precision.⁶

In each period t , the regulator chooses a level of detection intensity that detects the firm's fraudulent activity with a probability of $d_t \in (0, 1)$ and consumes regulatory resources of $\frac{\kappa}{2} (d_t)^2$. If the regulator successfully detects fraud, she would require the manager to restate the report to equal the economic earnings, i.e., $r_t = s_t$, and impose a penalty C_t on the manager that is proportional to the fraudulent amount:

$$C_t = c(r_t - s_t), \quad (4)$$

⁵Note that the noise term η_t implies managers do not have complete control over reporting outcomes. Gao and Zhang (2019) motivate this term by noting that internal controls can block managers' attempts to issue good earnings reports by overstating revenues or understating expenses. Similarly, under ASC 606, managers may attempt to shift revenue timing by redefining performance obligations, but the impact on reported revenue is not fully predictable because it also depends on when and how those obligations are ultimately satisfied (Dye and Zhang, 2025).

⁶Gao and Zhang (2019) show that the process generating r_t in equation (3) can be micro-founded as follows. In each period t , upon observing economic earnings s_t , the manager chooses manipulation m_t that adds m_t errors $\{\xi_l\}_{l=1}^{m_t}$ to s_t . Each error equals 0 with probability $q \in (0, 1]$ and 1 with probability $1 - q$. The report is then given by $r_t = s_t + \sum_{l=1}^{m_t} \xi_l$. Using the central limit theorem, we can approximate the distribution of $\sum_{l=1}^{m_t} \xi_l$ as $N(m_t(1 - q), m_t q (1 - q))$. We adopt this functional form for analytical tractability, but the key economic force is more general as long as managerial manipulation increases reporting noise.

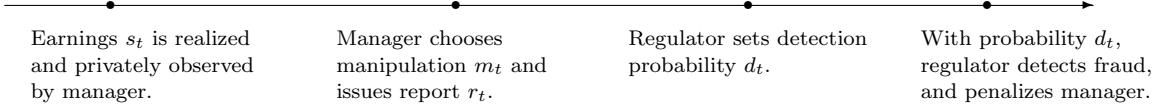


Figure 1: Timeline of the period- t game

where the coefficient $c > 0$, which is assumed to be not too excessive so that the equilibrium manipulation $m_t^* > 0$.⁷

The manager may also incur a personal cost of manipulating the report (e.g., reputational concerns or psychic costs), which we model as proportional to the amount of manipulation, i.e., $\mu(r_t - s_t)$. Dropping this personal cost does not affect the main equilibrium mechanism. We include it to facilitate structural estimation because such private costs are an intuitive component of managerial utility and help explain reporting behavior beyond formal regulatory penalties. A key difference between the regulatory cost and the personal cost is that C_t is imposed only if the manager is caught by the regulator, so its deterrent effect depends on the endogenous probability of fraud detection, whereas μ is incurred regardless of detection and is therefore taken as exogenous. Figure 1 summarizes the timing of events in each period.

The manager's payoff We assume that the manager derives utility from his compensation (or career prospects) that is proportional to firm valuation. To ease notation, we scale up the manager's utility so that it simply equals the valuation set by investors. In each period t , the manager maximizes the total present value of his future expected payoffs by choosing manipulation m_t :

$$U_t = \max_{m_t} E^M \left[\sum_{k=0}^{\infty} \beta_F^k u_{t+k} \middle| s_t, \mathcal{F}_{t-1} \right], \quad (5)$$

where $\beta_F \in [0, 1)$ is the manager's discount factor. $E^M [\cdot | s_t, \mathcal{F}_{t-1}]$ denotes the manager's expected utility in period t based on his information set, which includes s_t , his privately

⁷While the expected value of the penalty C_t is always positive, its realized value can be negative given that the noise in the report r_t is normally distributed. Nonetheless, we note that this feature is unlikely to affect the main economic forces of our model, since the manager only cares about the expected penalty at the time of making his manipulation choice m_t in each period t .

observed true earnings of the firm for the period, and $\mathcal{F}_{t-1}$, the firm's publicly released earnings reports in the past up until period $t - 1$. The manager's period- t payoff is:

$$u_t = d_t^* \left(\frac{s_t}{1 - \delta\rho} - C_t \right) + (1 - d_t^*) \frac{E^I[s_t|\mathcal{F}_t]}{1 - \delta\rho} - \mu(r_t - s_t), \quad (6)$$

where d_t^* is the regulator's period- t detection probability anticipated by the manager. The first two terms of equation (6) represent the manager's utility under two different scenarios, respectively. In the first scenario that occurs with a probability of d_t^* , the manager's fraudulent behavior is detected. As a result, the firm's true earnings s_t are revealed to investors, who would then update their valuation of the firm to $\frac{s_t}{1 - \delta\rho}$ based on s_t . The manager suffers a penalty of C_t proportional to m_t as shown in equation (4). In the second scenario that occurs with the remaining probability of $1 - d_t^*$, the manager's fraudulent behavior goes undetected. Thus, the firm's true earnings remain unknown to investors, who would then have to set its valuation of the firm to $\frac{E^I[s_t|\mathcal{F}_t]}{1 - \delta\rho}$ based on the firm's public reports $\mathcal{F}_t$. The manager incurs no penalty. Equation (6) also makes it clear that, except for the personal manipulation cost, $\mu(r_t - s_t)$, the manager's manipulation choice m_t affects firm valuation only when it is undetected.

The regulator's payoff In each period t , the regulator derives utility from the informativeness of the set of reports $\mathcal{F}_t \equiv \{r_t, r_{t-1}, \dots\}$ about the firm value V_t . This reflects the regulator's objective in ensuring the integrity of financial reporting. Since earnings follow an AR(1) process, s_t in period t is a sufficient statistic to estimate all of the firm's future earnings and the firm value. Thus, the informativeness of $\mathcal{F}_t$ can be measured by the informativeness of s_t , or the conditional variance of s_t :

$$\Phi_t \equiv \text{var}(s_t|\mathcal{F}_t). \quad (7)$$

We specify the regulator's period- t utility as

$$v_t = -(1 - d_t) \Phi_t - \frac{\kappa}{2} d_t^2. \quad (8)$$

where $\frac{\kappa}{2} d_t^2$ represents the cost of detection.⁸ Equation (8) sums up the regulator's expected utility in period t over two scenarios. If detection succeeds with probability d_t , then the true earnings s_t is revealed and the conditional variance Φ_t drops to zero. Alternatively, if detection fails with probability $1 - d_t$, then the conditional variance remains at $\Phi_t > 0$.

We assume that the regulator sets the detection probability d_t to maximize her lifetime utility W_t :

$$W_t = \max_{d_t} E^I \left[\sum_{k=0}^{\infty} \beta_R^k v_{t+k} \middle| \mathcal{F}_t \right], \quad (9)$$

where $\beta_R \in [0, 1)$ is the regulator's discount factor, $E^I [\cdot | \mathcal{F}_t]$ indicates that the regulator has the same information set as the investors.

1.1.2 Discussion of assumptions

In this section, we discuss and motivate several assumptions in our model. First, following Gao and Zhang (2019), we assume that manipulation increases uncertainty. Unlike an alternative assumption that manipulation generates no uncertainty or valuation distortion, our assumption helps rationalize regulators' concern about manipulation and allows us to endogenize the regulator's incentives to detect fraud. In our framework, neither periodic nor cumulative manipulation is fully unraveled, so both reduce the informativeness of financial

⁸The first component in v_t can be interpreted as the expected dis-utility stemming from the informational uncertainty upon the failure of fraud detection. Such dis-utility may be micro-founded in a model in which an external investor of the firm makes an investment decision based on the information about the firm's economic earnings s_t subject to a quadratic cost. Under that micro-foundation, the dis-utility of informational uncertainty equals the *ex ante* difference between the payoff that the investor earns when facing the informational uncertainty and that under perfect information (i.e., the investor's payoff loss from the informational uncertainty). Standard results from a single-person decision model with quadratic payoffs and Gaussian distributions imply the *ex ante* payoff loss from the informational uncertainty is proportional to the conditional variance $\Phi_t \equiv \text{var}(s_t | \mathcal{F}_t)$. Our specification of the regulator's payoff function therefore abstracts from other regulatory considerations, including minimizing the cumulative amount of fraud. In our model, only the portion of cumulative fraud that is not "unraveled" by investors matters, insofar as it affects information uncertainty Φ_t .

reports and increase uncertainty. The regulator therefore invests in detection, which yields two informational benefits: (1) successful detection resolves fraud-induced uncertainty, and (2) the threat of penalties reduces the manager’s *ex ante* incentive to manipulate, thereby improving report informativeness and lowering uncertainty. In addition, the assumption that manipulation increases uncertainty is crucial for generating the accumulation-and-reset cycle that lies at the heart of our empirical analysis. Absent this assumption, the reaccumulation of manipulation following fraud detection and the subsequent decline in manipulation induced by tighter regulation would have no effect on information uncertainty.

Second, for tractability, we abstract from the empirical regularity that manipulation reverses over time, e.g., the reversal of discretionary accruals (e.g., Beyer, Guttman, and Marinovic (2019)). That said, our model generates an economic force to explain how this reversal pattern may arise endogenously: because the regulator’s equilibrium detection effort increases with cumulative manipulation, firms that have manipulated more in the past have stronger incentives to reduce manipulation going forward, implying a reversal pattern in equilibrium. Also for tractability, we assume that the probability of successful detection depends only on the regulator’s detection choice and does not vary directly with the magnitude of fraud. Nevertheless, our analysis implies a positive relation in equilibrium: as fraud accumulates and increases uncertainty, the regulator optimally increases detection efforts.

Third, we assume that the regulator focuses on detecting and penalizing the manager’s current manipulation rather than past or cumulative manipulation. Two points are in order. First, it yields the greatest informational gain for the regulator to focus on fraud in the current-period report r_t and to uncover true earnings s_t , because s_t is a sufficient statistic for estimating all future earnings (see equation (2)). Conditional on the revelation of s_t , detecting fraud in past reports $\{r_{t-1}, r_{t-2}, \dots\}$ would incur additional costs but generate no incremental informational benefits. Second, if the regulator could instead uncover and penalize past or cumulative fraud, the manager would have weaker incentives to commit fraud. However, in terms of economic forces, this modification is similar to imposing a larger

penalty on current-period fraud, i.e., a higher c .

1.1.3 Analysis

In this section, we analyze the manager’s optimal manipulation choice m_t^* and the regulator’s equilibrium detection choice d_t^* . For brevity, we present only the equations needed to convey the model’s key intuitions, leaving detailed derivations to Appendix A.1. The dynamic interaction between the manager and the regulator gives rise to a complex dynamic programming problem that does not admit a closed-form solution in general. To keep the analysis tractable, we focus on a baseline myopic version of the model in this section, in which both the manager and the regulator are myopic (i.e., $\beta_F = \beta_R = 0$) and choose m_t and d_t to maximize their contemporaneous payoffs u_t and v_t in each period. This simplified setting allows us to characterize the equilibrium analytically and to highlight the core economic mechanisms of the model. We also set $\mu = 0$ in our baseline analysis, because while the personal cost can be empirically important (e.g., Karpoff et al. (2008)), it does not drive the key economic forces in the model. We present the full forward-looking dynamic model, including a personal cost, in Appendix A.2. There, we solve the model numerically and show that the key insights from the myopic benchmark continue to hold.

Baseline model with myopic manager and regulator

We begin by solving the manager’s optimal manipulation choice m_t^* , anticipating the regulator’s optimal detection choice d_t^* . The first step is to analyze how firm valuation varies with m_t in each period:

$$E^I[s_t|\mathcal{F}_t] = (1 - w_t) \cdot \rho E^I[s_{t-1}|\mathcal{F}_{t-1}] + w_t \cdot [r_t - \hat{m}_t(1 - q)]. \quad (10)$$

As shown, firm valuation is set as the weighted average of the investors’ prior of s_t (the first term) and the incremental information that investors gain from seeing the report r_t

(the second term). The prior builds on the AR(1) process of s_t and equals $\rho E^I [s_{t-1} | \mathcal{F}_{t-1}]$. $\hat{m}_t$ in equation (10) is the investors' conjectured manipulation by the manager, and the manager factors in the investor's conjecture in his maximization problem. In equilibrium, this conjecture equals the manager's optimal manipulation choice, i.e., $\hat{m}_t = m_t^*$. To extract information from the new report, investors rationally subtract the expected manipulation $m_t^*(1 - q)$, leading to a refined signal $r_t - m_t^*(1 - q)$. The weight

$$w_t = \frac{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2}{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2 + m_t^* q (1 - q)}, \quad (11)$$

captures the value relevance of the earnings report, with Φ_{t-1} being the inverse precision of the prior, as defined in equation (7). When Φ_{t-1} is larger, the prior is less precise and so the investors have to place a greater weight on the current report to infer firm fundamentals.

Equation (10) suggests that undetected manipulation m_t has a contemporaneous effect as well as an intertemporal effect on the investors' expected firm value. To see the contemporaneous effect, note that m_t inflates the current earnings report r_t , which in turn boosts firm valuation in the current period $E^I [s_t | \mathcal{F}_t]$; this effect works through the second term of the equation. To see the intertemporal effect, note that $E^I [s_t | \mathcal{F}_t]$ serves as the prior for the investors to conjecture future earnings s_{t+1} , so as m_t inflates $E^I [s_t | \mathcal{F}_t]$, it also boosts firm valuation in the next period $E^I [s_{t+1} | \mathcal{F}_{t+1}]$; this effect works through the first term of the equation. In fact, such bias propagates to all future s_{t+k} for $k > 0$ through the recursive form of equation (10).⁹ Specifically, assuming that the firm remains undetected, we can summarize the contemporaneous effect ($k = 0$) and the intertemporal effect ($k > 0$) of m_t below as

$$\frac{\partial E^M [E^I [s_{t+k} | \mathcal{F}_{t+k}] | s_t, \mathcal{F}_t]}{\partial m_t} = \begin{cases} w_t (1 - q) & \text{if } k = 0 \\ \rho^k \prod_{l=1}^k (1 - w_{t+l}) w_t (1 - q) & \text{if } k > 0 \end{cases}. \quad (12)$$

⁹This can be easily seen by shifting equation (10) forward by k periods from t to $t + k$.

When the discount factor $\beta_F = 0$, the manager cares only about the contemporaneous effect of his manipulation choice (i.e., the part with $k = 0$ in equation (6)). Taking the derivative of u_t in equation (6) with respect to m_t and then substituting in equation (12) derived above, we obtain the first-order condition (F.O.C.) as¹⁰

$$\underbrace{c(1-q)d_t^*}_{\text{MC of } m_t} = \underbrace{(1-d_t^*)\frac{w_t(1-q)}{1-\delta\rho}}_{\text{MB of } m_t \text{ from the contemporaneous effect}}. \quad (13)$$

The MC of manipulation, expressed on the left-hand side (LHS) of the F.O.C., increases with the regulator's optimal choice of detection probability d_t^* , which is correctly conjectured by the manager. The MB of manipulation, expressed on the right-hand side (RHS) of the F.O.C., arises from the contemporaneous effect of m_t in boosting the firm valuation in the current period $E^I[s_t|\mathcal{F}_t]$, taking into account the likelihood of being detected in the current period $(1-d_t^*)$. Substituting the expression of w_t in equation (11) into equation (13) gives the equilibrium manipulation m_t^* as a function of the fraud-induced uncertainty Φ_{t-1} and the equilibrium detection strength d_t^* , as stated in the following lemma.

Lemma 1 *In each period t , given the regulator's equilibrium detection choice d_t^* conjectured by the manager, the manager chooses the optimal manipulation*

$$m_t^*(d_t^*, \Phi_{t-1}) = \frac{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2}{q(1-q)} \left(\frac{1-d_t^*}{cd_t^*} \frac{1}{1-\delta\rho} - 1 \right), \quad (14)$$

where Φ_{t-1} is the fraud-induced uncertainty about s_{t-1} , as defined in equation (7). m_t^* is decreasing in d_t^* and increasing in the last-period uncertainty Φ_{t-1} .

Equation (14) illustrates some equilibrium properties of manipulation m_t^* . First, holding

¹⁰Note that we focus on a fully revealing equilibrium in which the market perfectly anticipates the manager's equilibrium manipulation choice m_t^* . The F.O.C. implies that, given that the market rationally conjectures the equilibrium manipulation level m_t^* , the MB of manipulation equals the MC of manipulation such that the manager (weakly) prefers the equilibrium manipulation choice m_t^* . Hence m_t^* constitutes a rational-expectation equilibrium in which the manager's equilibrium choice is consistent with the market conjecture.

d_t^* constant, m_t^* is increasing in the level of fraud-induced uncertainty Φ_{t-1} . Intuitively, all else equal, the manager has greater incentives to commit fraud as the market faces a higher uncertainty about the firm and relies on the manager's report to a larger extent (i.e., higher w_t). Second, m_t^* is decreasing in the detection strength d_t^* as a higher d_t^* increases the MC of manipulation and deters manipulation in equilibrium.

Given the manager's optimal manipulation choice, Φ_t evolves endogenously over time, and standard Bayesian updating yields its law of motion, as shown in Lemma 2 below:¹¹

Lemma 2 *In each period t , if the regulator detects fraud, the fraud-induced uncertainty about the firm's earnings s_t drops to zero, i.e., $\Phi_t \equiv 0$. If the regulator fails to detect fraud, the period- t uncertainty Φ_t is a function that is decreasing in the period- t equilibrium detection intensity d_t^* and increasing in the last-period Φ_{t-1} :*

$$\Phi_t(d_t^*, \Phi_{t-1}) = \frac{m_t^* q (1 - q) (\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2)}{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2 + m_t^* q (1 - q)} \quad (15)$$

$$= (\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2) \left[1 - (1 - \delta\rho) \frac{cd_t^*}{1 - d_t^*} \right]. \quad (16)$$

The law of motion sheds light on two informational benefits of fraud detection. First, if detection succeeds in period t , it reveals the firm's true economic earnings and thus eliminates the cumulative fraud-induced uncertainty Φ_t . Furthermore, from the law of motion (15), the future fraud-induced uncertainty is increasing in the prior uncertainty; in that vein, the period- t information gain from the fraud detection success propagates into all future periods, reducing the future levels of fraud-induced uncertainty.

Second, even if the regulator fails to detect fraud, the manager's *ex ante* anticipation of the regulatory threat still produces a disciplinary effect in curbing the growth of the fraud-induced uncertainty. To see this, we substitute the equilibrium manipulation choice

¹¹Also, observe that iterating (15) over time suggests that Φ_t essentially depends on the manager's undetected equilibrium manipulation accumulated in the past, i.e., $\{m_t^*, m_{t-1}^*, \dots\}$. We thus hereafter refer to the state variable Φ_t as either the information uncertainty about the firm fundamentals in period t or the cumulative fraud-induced uncertainty interchangeably.

m_t^* in equation (14) into the law of motion (15) and obtain equation (16) that expresses the period- t fraud-induced uncertainty as a decreasing function of the equilibrium detection intensity d_t^* at any given level of the past uncertainty. Intuitively, anticipating the regulatory detection, the manager refrains from further manipulation, improving the informativeness of the period- t report and lowering the uncertainty about the firm's economic earnings. Of course, like its effect in the case of fraud detection success, the disciplinary effect of regulatory detection upon detection failure in period t also propagates to all future periods through the law of motion on Φ_t .

After characterizing the equilibrium manipulation choice m_t^* and fraud-induced uncertainty Φ_t as a function of the equilibrium level of regulation intensity d_t^* rationally anticipated by the manager and investors, we now analyze the regulator's choice of the detection intensity d_t to close the model. As implied by our discussion of the law of motion on Φ_t , the regulator's choice of detection likelihood in period t carries two effects. First, a higher d_t increases the regulator's period- t utility by boosting the chance of detection success (upon which Φ_t is decreased to zero); this is the contemporaneous effect of detection. Second, clearing Φ_t also reduces the expected level of $\Phi_{t+\ell}$ for all $\ell > 0$ because Φ_t affects all future $\Phi_{t+\ell}$ through the law of motion (16); this is the intertemporal effect of detection. When the discount factor $\beta_R = 0$, however, the regulator cares only about the contemporaneous effect. Taking the derivative of v_t in equation (8) with respect to d_t , and using the law of motion (16) gives the optimal detection choice d_t^* :

$$\underbrace{\kappa d_t^*}_{\text{MC of } d_t} = \Phi_t = \underbrace{(\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2) \left[1 - (1 - \delta\rho) \frac{cd_t^*}{1 - d_t^*} \right]}_{\text{MB of } d_t \text{ from the contemporaneous effect}}. \quad (17)$$

As in the F.O.C. for the manager's problem, we express the MC of detection on the LHS and the MB of detection on the RHS. The MC is proportional to the amount of detection intensity set by the regulator, d_t . The MB is increasing in the cumulative level of fraud Φ_t as it comes from clearing uncertainty about s_t in the current period (the contemporaneous

effect). In particular, applying the implicit function theorem to equation (17) shows that the regulator's choice of optimal detection strength d_t^* is increasing in the prior fraud-induced uncertainty Φ_{t-1} . That is, the regulator matches the strength of fraud detection with the severity of fraud in equilibrium. This result is intuitive because the manager's manipulation in the past adds noise to the firm's reports and decreases informativeness. Since the regulator's objective is to clear fraud and restore informativeness of the firm's reports, her gains are higher from detecting reports with more extensive fraud. We formally state this property of the equilibrium detection intensity in the following lemma.

Lemma 3 *In each period t , the regulator chooses the optimal detection probability $d_t^* \in (0, 1)$ that is the unique solution to equation (17). d_t^* is increasing in the last-period fraud-induced uncertainty Φ_{t-1} .*

Figure 2 illustrates an example of the optimal detection policy, d , chosen by the regulator as a function of the last-period cumulative fraud-induced uncertainty, Φ . In creating this figure, we set β_F and β_R to be zero so that the manager and the regulator are both myopic. Other parameters are set close to the values we estimated from data in Section 2.2. Although d appears almost linear in Φ over the range shown in the figure, its true functional form is nonlinear. The near-linearity arises because Φ is evaluated over a narrow interval that is chosen to match the empirically relevant range of d in the data. Over a broader range of Φ , the detection policy is increasing and concave in Φ .

Finally, to complete characterizing the equilibrium manipulation and detection choices, we use the expression of d_t^* in equation (17) to express the equilibrium manipulation m_t^* in equation (14) as a function of the fraud-induced uncertainty Φ_{t-1} only. Interestingly, there are two countervailing effects of Φ_{t-1} on m_t^* . To see this, recall from Lemma 1 that fixing the regulator's detection choice, m_t^* is increasing in Φ_{t-1} . However, Lemma 1 also suggests that as Φ_{t-1} increases, the regulator would invest more heavily in the detection technology, which deters manipulation and reduces m_t^* . Nonetheless, in the myopic baseline model, we find that the disciplinary effect of detection always dominates such that m_t^* is

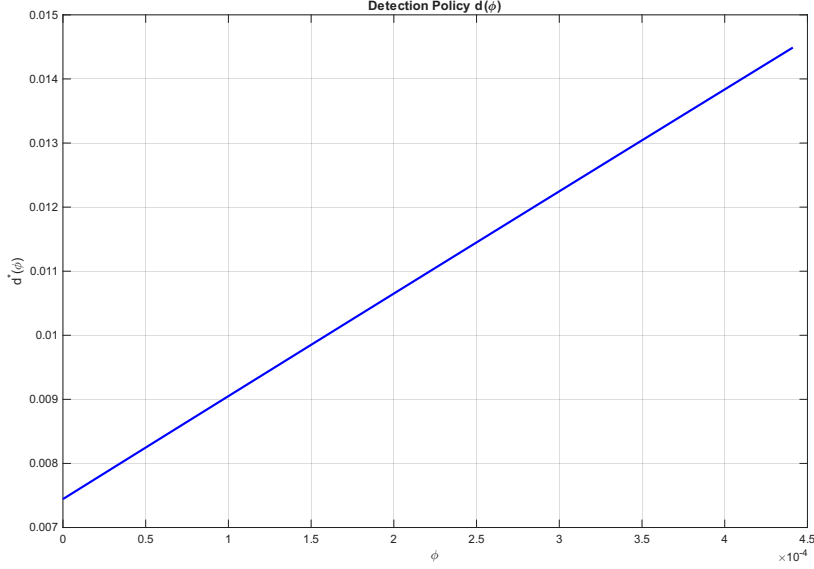


Figure 2: Equilibrium detection probability $d^*(\Phi)$. This figure plots the regulator's optimal choice of detection intensity, d^* , as a function of the last-period information uncertainty induced by cumulative fraud, Φ , in the myopic model. The parameters used in generating this figure are: $\beta_F = \beta_R = 0$, $\delta = 0.9$, $\rho = 0.7$, $\sigma_\varepsilon = 0.015$, $q = 0.5$, $\kappa = 0.03$, and $c = 2.7$.

strictly decreasing in Φ_{t-1} . This simple monotone relationship between m_t^* and Φ_{t-1} does not necessarily extend to the full model with the forward-looking manager and regulator; in that setting, the two countervailing effects of Φ_{t-1} go together, leading to possible non-monotonic relations between m_t^* and Φ_{t-1} over time.

For future reference, we summarize in the following proposition the equilibrium manipulation choice and detection strength $\{m_t^*(\Phi_{t-1}), d_t^*(\Phi_{t-1})\}$ under the myopic model.

Proposition 1 *Consider the myopic model with the manager's and the regulator's discount factors $\beta_F = \beta_R = 0$. For a given level of accumulated past fraud Φ_{t-1} , the regulator's equilibrium detection choice $d_t^*(\Phi_{t-1})$ and the manager's equilibrium manipulation choice $m_t^*(\Phi_{t-1})$ are given by the following set of equations, with the two agents rationally anticipating each*

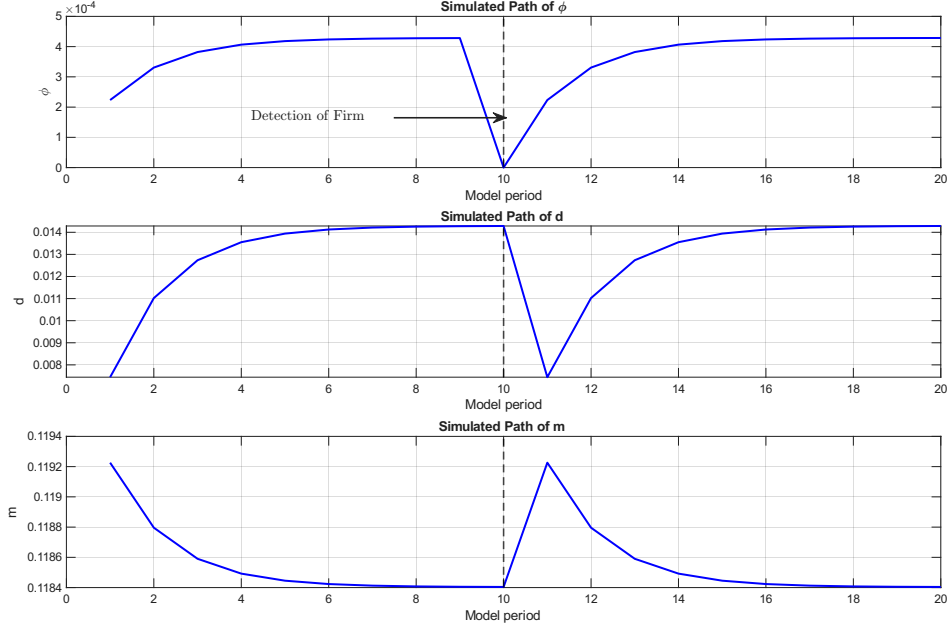


Figure 3: Simulated dynamics in the one-firm myopic model. The figure plots the simulated paths of fraud-induced uncertainty Φ (top panel), the regulator's optimal detection intensity d (middle panel), and the manager's optimal manipulation choice m (bottom panel). The simulation is generated from the one-firm myopic version of the model. A detection event is imposed exogenously at $t = 10$, which resets Φ to zero. Parameter values are set to $\beta_F = \beta_R = 0$, $\delta = 0.9$, $\rho = 0.7$, $\sigma_\varepsilon = 0.015$, $q = 0.5$, $\kappa = 0.03$, and $c = 2.7$.

other's optimal policy function:

$$d_t^*(\Phi_{t-1}) = \frac{1}{\kappa} (\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_t^*(\Phi_{t-1})}{1 - d_t^*(\Phi_{t-1})} \right), \quad (18)$$

$$m_t^*(\Phi_{t-1}) = \frac{\kappa}{cq(1-q)} \frac{1 - d_t^*(\Phi_{t-1})}{1 - \delta\rho}. \quad (19)$$

The equilibrium detection $d_t^*(\Phi_{t-1})$ is increasing in Φ_{t-1} , whereas the equilibrium manipulation $m_t^*(\Phi_{t-1})$ is decreasing in Φ_{t-1} .

Figure 3 illustrates simulated paths of the dynamic interaction among the last-period fraud-induced uncertainty Φ , the regulator's detection intensity d , and the manager's manipulation choice m in the one-firm myopic model. The firm starts with a modest level of Φ , so detection intensity is relatively low. Anticipating a low probability of being detected, the manager chooses a high level of manipulation, which leads to a rapid accumulation of

fraud-induced uncertainty. As Φ increases, the regulator optimally responds by increasing detection intensity, and the manager, facing greater scrutiny, rationally reduces manipulation. This gradual adjustment reflects the endogenous feedback between the state variable Φ and the regulator's and manager's choices d and m .

The simulation also features a realized detection event at period $t = 10$, at which point Φ is reset to zero. This detection generates sharp adjustments by both the regulator and the manager in the subsequent periods: detection intensity drops immediately following the reset, while manipulation rises as the manager exploits the temporarily relaxed regulatory environment. Over the subsequent periods, manipulation again drives the accumulation of Φ , prompting the regulator to increase detection intensity, and the system converges back toward its long-run equilibrium. Overall, the figure highlights the cyclical nature of the model, in which periods of low detection and high manipulation sow the seeds for tighter enforcement, while detection events reset the state and initiate a new cycle of strategic interaction between the regulator and the firm.

While the myopic model is useful for isolating the core economic forces, it abstracts from intertemporal considerations that may affect both managerial manipulation and regulatory enforcement. In particular, it ignores how current actions influence future payoffs through continuation values. To address these limitations, we develop in Appendix A.2 a full forward-looking version of the model in which both the manager and the regulator maximize the present value of expected payoffs and internalize the effects of current choices on future outcomes. We solve the forward-looking model numerically and show that the key qualitative insights from the myopic benchmark continue to hold.

1.2 Multi-firm Model

1.2.1 Model Setup

In this section, we extend the single-firm framework to a multi-firm setting in which the regulator monitors multiple firms and allocates limited detection resources across them. This allocation problem is corroborated by recent empirical studies of enforcement at the U.S. Securities and Exchange Commission, the primary securities regulator in the United States (see, e.g., Holzman et al. (2024)), which show that the agency operates under constrained enforcement resources and must therefore prioritize which firms to investigate.

While the single-firm model highlights the regulator’s intertemporal allocation of detection effort, the multi-firm setting adds a cross-sectional dimension: the regulator must allocate detection not only over time but also across firms. This extension allows us to study how regulatory priorities vary across firms and how firms respond heterogeneously to anticipated detection intensity, generating richer dynamics of fraud and enforcement.

The model setup closely follows the single-firm case, with two key modifications. First, the economy consists of N firms whose economic earnings are independent of one another. This assumption rules out mechanical correlations in fraud arising from correlated fundamentals and allows us to abstract from information spillovers across firms, which are not the focus of this study. Second, the regulator must allocate limited detection resources across the N firms. In each period, the regulator conducts an independent inspection of each firm’s report. Let $d_{it} \in [0, 1]$ denote the probability that the inspection uncovers fraud in firm i ’s report at time t , where $i \in \{1, 2, \dots, N\}$. The choice of d_{it} reflects the regulator’s allocation of detection technology and enforcement intensity to firm i .

We assume that the regulator faces a convex cost of detection in each period, given by

$$\frac{\kappa}{2} \left(\sum_{i=1}^N d_{it} \right)^2. \quad (20)$$

This cost structure reflects limited regulatory resources or capacity. Dedicating more detection effort to one firm raises the marginal cost of detecting fraud at other firms, creating a trade-off that disciplines the regulator’s cross-sectional allocation of enforcement effort.

1.2.2 Analysis

Similar to our analysis of the single-firm model, we proceed in two steps. To keep the analysis transparent and to isolate the key economic mechanisms, we first study a baseline myopic version of the multi-firm model. In this setting, both managers and the regulator are myopic and choose their actions to maximize contemporaneous payoffs, which allows us to characterize the equilibrium allocation of detection across firms analytically. This benchmark highlights how limited regulatory capacity shapes the regulator’s cross-sectional enforcement priorities and how firms respond to heterogeneous detection threats.

We then consider the full forward-looking multi-firm model, in which both managers and the regulator internalize the intertemporal consequences of their actions through continuation values. Because the resulting dynamic programming problem does not admit a closed-form solution, we solve the model numerically and relegate the details to Appendix A.4. We show that the key qualitative insights from the myopic benchmark remain intact.

Baseline model with myopic regulator and two firms

Even though the multi-firm setting has no restrictions on the maximum number of firms in the model, a 2-firm setting is sufficient to deliver the key implications of a general multi-firm model. As a result, our analyses in this section focus on the case with two firms (i.e., $N = 2$) to retain more analytical tractability.

In the multi-firm setting, the manipulation decision of each manager and the detection decision of the regulator can be similarly characterized as in the single-firm model, with the critical difference that now their decisions are interdependent via the regulator’s allocation of detection resource across firms. Both m_{it}^* and d_{it}^* can be written recursively as functions

of the cumulative levels of past fraud at all firms, $\{\Phi_{1t-1}, \Phi_{2t-1}\}$. Hence, we can treat $\{\Phi_{1t-1}, \Phi_{2t-1}\}$ as the set of state variables for period t and characterize the equilibrium in the proposition below. For ease of notation, we omit the time subscript and denote variables of the next period with a prime.

Proposition 2 *Consider the myopic model with two firms $i \in \{1, 2\}$ and the manager's and the regulator's discount factors $\beta_F = \beta_R = 0$. Without loss of generality, assume that the levels of accumulated past fraud at the two firms $\Phi_1 \geq \Phi_2$. The equilibrium manipulation decision and detection strength at the two firms $\{m_1^*, m_2^*, d_1^*, d_2^*\}$ are given as follows:*

1. *If Φ_1 is much larger than Φ_2 (i.e., $\Phi_1 > H(\Phi_2)$), $d_2^* = 0$ and d_1^* solves*

$$(\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) = \kappa d_1^*, \quad (21)$$

and is increasing in Φ_1 , where $H(\Phi_2)$ is an increasing function defined in the proof; the equilibrium manipulation decision $m_2^* = \infty$, and, conditional on d_1^* , the equilibrium manipulation decision m_1^* is the same as that in equation (14) of the single-firm case;

2. *If Φ_1 and Φ_2 are of similar sizes (i.e., $\Phi_1 \leq H(\Phi_2)$), the pair of $\{d_1^*, d_2^*\}$ solves*

$$(\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) = \kappa (d_1^* + d_2^*), \quad (22)$$

$$(\rho^2 \Phi_2 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_2^*}{1 - d_2^*} \right) = \kappa (d_1^* + d_2^*). \quad (23)$$

d_1^* is increasing in Φ_1 and decreasing in Φ_2 whereas d_2^* is increasing in Φ_2 and decreasing in Φ_1 ; conditional on $\{d_1^*, d_2^*\}$, the equilibrium manipulation decisions $\{m_1^*, m_2^*\}$ are the same as in equation (14) of the single-firm case.

Appendix A.3 provides a detailed proof to this proposition. This proposition points to a new insight relative to those from the single-firm model. Recall that in the single-firm model, the regulator matches the strength of fraud detection at a given firm with the

severity of fraud at that firm. With multiple firms, the regulator faces the additional task of optimally allocating the regulatory resources across firms. Consequently, the detection intensity imposed by the regulator on a given firm depends on not only the firm's own information uncertainty from cumulative fraud but also how it compares to information uncertainty about the other firms in the economy. Stated differently, when there are multiple firms, the regulator determines the strength of fraud detection at each firm based on the *relative* severity of fraud. When the fraud-induced uncertainty at one firm is much higher than that at the other firms, the regulator will devote most of the regulatory resource to detect fraud at that firm and almost none to the other firms. Furthermore, even when the regulator directs regulatory resources to both firms, she will direct fewer regulatory resources to a firm when she anticipates the fraud at the other firms to rise. This, in turn, boosts the incentive of the underlying firm to manipulate. In this light, an increase of fraud at a firm generates spillovers to the other firms through the interdependence of detection and manipulation choices across firms.

Figure 4 illustrates the regulator's optimal detection policy in the two-firm setting. Using the same parameter values as in the one-firm model, we solve for the equilibrium detection intensities d_1 and d_2 as functions of both firms' fraud-induced uncertainty, Φ_1 and Φ_2 . Unlike the single-firm case, detection decisions are now interdependent: the optimal detection intensity on a given firm depends not only on its own level of cumulative fraud, but also on the peer firm's level of information uncertainty. The figure presents the model solution in two heatmaps. The left panel plots d_1 as a function of (Φ_1, Φ_2) , while the right panel plots d_2 as a function of the same state variables. In each panel, the horizontal axis represents the peer firm's uncertainty and the vertical axis represents the firm's own uncertainty, and the color scale indicates the magnitude of detection intensity. This representation highlights how regulatory resources are endogenously reallocated across firms as their relative levels of cumulative fraud change and makes transparent the strategic substitutability in detection: when one firm's Φ stands out, the regulator concentrates enforcement effort on that firm and

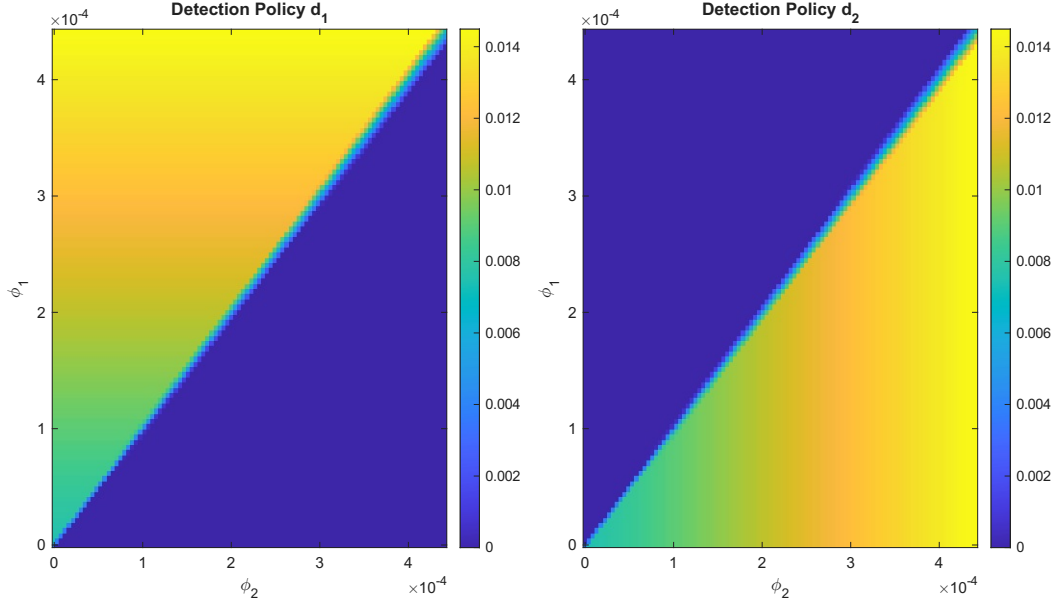


Figure 4: Equilibrium detection probabilities in the two-firm myopic model. This figure plots the regulator's optimal detection intensities d_1^* (left panel) and d_2^* (right panel) as functions of the information uncertainty induced by cumulative fraud, (Φ_1, Φ_2) . In each panel, the vertical axis represents the firm's own level of fraud-induced uncertainty, while the horizontal axis represents the peer firm's uncertainty. The depth of color indicates the magnitude of detection intensity. Unlike the one-firm case, detection decisions are interdependent: the optimal detection intensity imposed on a firm depends not only on its own Φ_i , but also on the peer firm's Φ_j . Parameter values are set to $\beta_F = \beta_R = 0$, $\delta = 0.9$, $\rho = 0.7$, $\sigma_\varepsilon = 0.016$, $q = 0.44$, $\kappa = 0.03$, and $c = 2.7$.

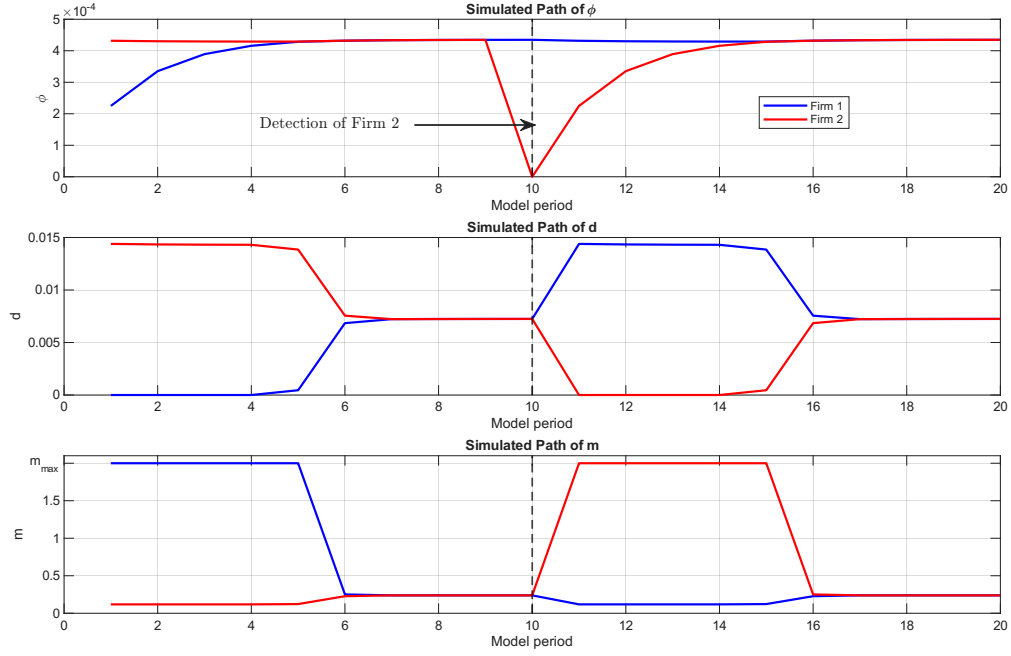


Figure 5: Simulated dynamics in the two-firm myopic model. The figure plots the simulated paths of fraud-induced uncertainty Φ_i (top panel), the regulator's optimal detection intensities d_i (middle panel), and the managers' optimal manipulation choices m_i (bottom panel) for both firms. The simulation is generated from the two-firm myopic version of the model using the equilibrium policy functions for detection and manipulation. A detection event is imposed exogenously on firm 2 at a period $t = 10$, which resets its Φ_i to zero. Parameter values are set to $\beta_F = \beta_R = 0$, $\delta = 0.9$, $\rho = 0.7$, $\sigma_\varepsilon = 0.016$, $q = 0.44$, $\kappa = 0.03$, and $c = 2.7$.

relaxes scrutiny on the other.

Figure 5 uses the model to illustrate the joint dynamics of fraud, detection, and manipulation for two firms. Starting from different initial levels of fraud-induced uncertainty, we simulate the paths of Φ_{it} , d_{it} , and m_{it} forward using the equilibrium policy functions derived from the model. The top panel plots the evolution of fraud-induced uncertainty for both firms, the middle panel shows the corresponding detection intensities chosen by the regulator, and the bottom panel reports the managers' manipulation choices. As in the one-firm case, Φ_{it} evolves endogenously according to the law of motion implied by manipulation and detection, while d_{it} and m_{it} are chosen optimally given the current state of both firms.

In this simulation, firm 1 starts out with a relatively low Φ , compared to firm 2, and in

the first few periods, the regulator concentrates detection on firm 2 and leaves firm 1 under the radar. As a result, firm 1 manipulates aggressively, leading to a quick convergence of its Φ to that of firm 2. After six periods, the two firms' fraud-induced uncertainty levels converge, and thus the regulator allocates her detection intensity equally between the two firms.

To highlight the role of realized detection, the simulation imposes a detection event for firm 2 at period $t = 10$, at which point that firm's fraud-induced uncertainty is reset to zero. This event triggers a sharp reallocation of regulatory attention across firms and induces immediate adjustments in managerial behavior. In particular, right after detection, firm 2's uncertainty is cleaned up and thus it temporarily faces lower detection intensity. The regulator shifts her attention to firm 1, who remains highly suspicious. Interestingly, this renders two consequences. First, firm 2, with a low level of Φ , now becomes more aggressive in manipulation; Second, firm 1, with a high level of Φ becomes more conservative. Over time, the interaction between manipulation, detection, and the accumulation of fraud-induced uncertainty causes the two firms' states to converge again, generating cyclical patterns in enforcement and fraud. Together, these dynamics illustrate how cross-firm strategic interactions and endogenous reallocation of regulatory resources shape the joint evolution of fraud and detection in the two-firm economy.

While the myopic two-firm model provides sharp analytical insights into the regulator's cross-sectional allocation of detection and firms' strategic responses, it abstracts from intertemporal incentives. In Appendix A.4, we present and solve the full forward-looking model, in which both firms and the regulator internalize the continuation value of their actions. We solve the model numerically and show that the qualitative insights from the myopic benchmark continue to hold as the number of firms increases and as intertemporal incentives are taken into account.

2 Empirical Results

The theoretical framework outlined above yields testable predictions about the evolution of fraud-induced uncertainty, the information effects of detection, and the role of regulatory reallocation across firms. In this section, we assess the model’s empirical relevance by examining whether the observed dynamics between earnings manipulation and regulatory detection align with these predictions. To do so, we first structurally estimate a three-firm, forward-looking setting of the model, and then conduct reduced-form analysis that requires no assumptions of the structural model.

2.1 Sample and Variable Measurement

The model speaks to the joint dynamics of earnings manipulation and detection. Because undetected earnings manipulation is unobservable, we capture detection events and detected manipulation using accounting restatement data from Audit Analytics. The initial sample covers 18,340 restatements announced by 10,404 unique firms between 1995Q1 and 2019Q3. We merge these data with firm financial data from Compustat Quarterly, analyst forecasts from I/B/E/S, and stock returns from CRSP Daily files. If a firm has multiple restatements in our sample period, we focus on the restatement with the largest cumulative impact on earnings based on Audit Analytics.

The model’s predictions hinge on the state variable Φ_{t-1} , which captures cumulative fraud-induced information uncertainty and governs how much weight investors place on new earnings information. Because Φ_{t-1} is unobservable, we test the model’s implications through observable price responses. Specifically, we use the ERC as a proxy for this mechanism. Using the ERC has two advantages. First, it is a widely accepted measure of earnings-report value relevance in the literature (e.g., Fischer and Verrecchia (2000); Fang et al. (2017)), making the results easy to interpret. Second, in our model, a higher Φ_{t-1} implies an information environment in which the market places greater weight on earnings news at

the announcement, suggesting a positive relation between Φ_{t-1} and the ERC.¹²

As is standard in the literature, we define the ERC as the coefficient on earnings surprise from a regression of earnings announcement returns on earnings surprise. Specifically, for quarter t , we measure earnings surprise as reported quarterly earnings per share (EPS) minus the I/B/E/S median consensus EPS forecast immediately prior to the announcement, scaled by the stock price two days before the announcement. We measure earnings announcement returns as the three-day market-adjusted abnormal return centered on the earnings announcement date.

2.2 Estimation

2.2.1 Three-Firm, Forward Looking Model

While the myopic benchmarks in Section 1 excel at illustrating the economic forces as cleanly as possible, they miss the intertemporal effects of manipulation and detection, which significantly understates the dynamic interaction between them. To provide a more reliable quantitative assessment of the model, we estimate a three-firm, forward-looking model in this section. Appendix A.4 summarizes the model's setup and solution. This full-fledged model shares the same intuition as the myopic models, but offers two empirical advantages.

First, relative to the myopic settings, the forward-looking model is more likely to deliver quantitative effects closer to those in the data, whereas the myopic benchmarks may understate the magnitudes. As such, the forward-looking setting implies a possible non-monotonic relationship between fraud-induced uncertainty and equilibrium manipulation. This is because, when uncertainty is low, the manager's incentive to manipulate is limited because the gains are small; when uncertainty is high, it is limited because expected detection costs are high. As a result, equilibrium manipulation is greatest at intermediate levels of

¹²Mathematically, we can write $ERC_t = \frac{\rho^2 \Phi_{t-1} + \sigma_\epsilon^2}{\rho^2 \Phi_{t-1} + \sigma_\epsilon^2 + m_t^* q(1-q)} \cdot \frac{1}{1-\delta\rho}$. Taking the derivative with respect to Φ_{t-1} gives $\frac{\partial ERC_t}{\partial \Phi_{t-1}} = \frac{\rho^2 m_t^* q(1-q)}{(1-\delta\rho)(\rho^2 \Phi_{t-1} + \sigma_\epsilon^2 + m_t^* q(1-q))^2} > 0$ for $m_t^* > 0$ and $0 < q < 1$, so ERC_t increases in Φ_{t-1} .

uncertainty. This mechanism generates accumulation-and-reset cycles of manipulation and detection, even when regulation corrects severe information distortions in the short run.

Second, a three-firm model is the smallest environment that delivers a realistic, asymmetric rotation of regulatory attention across high-, medium-, and low-uncertainty firms (H-, M-, and L-firms, respectively). This richer cross section allows enforcement to shift endogenously over time, generating overlapping accumulation-and-reset cycles that more closely match the fraud patterns observed in the data. By contrast, the two-firm model implies mechanically symmetric rotation and therefore less diverse dynamics. An N -firm extension is computationally infeasible for structural estimation, and adding firms beyond three offers little additional insight or identification gain.

2.2.2 Simulation and Identification

We estimate the model using the Simulated Method of Moments (SMM), which requires us to find the unique set of parameters that make the model fit the data as closely as possible. We first explain how we simulate the model to generate a panel of firms who manipulate their earnings report and a regulator who allocates resources to detect fraud conducted by these firms.

Since our model features a three-firm setting, we start our simulation by creating a panel of three-firm groups. Each group contains three firms with different initial levels of $\Phi_{i,0}$ that implies ERC of different levels, labeled as high (H), medium (M), and low (L), and groups are independent from each other. As we simulate the model forward, each period t , we set the regulator’s detection intensity on each firm, $d_{i,t}$, and the firms’ manipulation, $m_{i,t}$, based on the model’s prediction, which are both functions of the firms’ $\Phi_{i,t-1}$. Successful detection, $\mathbb{I}_{i,t}$, occurs randomly with a probability of the detection intensity the regulator imposes on each firm, $d_{i,t}$. If a detection event is triggered at period t for firm i , so that $\mathbb{I}_{i,t} = 1$, we reset $\Phi_{i,t}$ to zero; otherwise, $\mathbb{I}_{i,t} = 0$, and $\Phi_{i,t}$ evolves according to the law of motion implied by the model. Each period, we update each firm’s ERC label based on its model-implied ERC,

with H , M , and L indicating the firm belonging to the top, mid, and bottom tercile of ERC. This simulation process is repeated for 500 periods, with the first 50 periods disregarded as burn-in.

While we are able to track Φ , m , and d for each firm perfectly in the model, they are unobservable in the data. We leverage the positive relation between Φ and ERC implied by the model and use ERC observed in the data to classify firms in our sample, just as we do in the simulation. Specifically, we divide firms into terciles each quarter based on their ERC and assign an indicator taking values of L , M , or H for the low-, medium-, and high-ERC tercile, respectively.¹³ These terciles in the data can be mapped to firms with model-implied ERC in the corresponding tercile.

To fit the model to the data, we first calibrate a few parameters. They are not central to our analysis and therefore are left out of the estimation. We first set the market’s discount rate δ to be 0.974 per quarter, corresponding to 0.9 per annum. We set the regulator’s and manager’s discount rate, β_R and β_F to be the same as δ , so that we do not consider managers’ or regulator’s short-termism or impatience.¹⁴

We then calibrate the persistence and volatility of the true earnings process, ρ and σ_ϵ , which are taken as exogenous to the model. In our model, all firms manipulate but only a small fraction of them are eventually detected. As a result, the reported earnings by undetected firms are manipulated and distorted, which do not reflect the true earnings process. We therefore use the restated earnings by firms that were detected in our sample. The assumption is that the restated earnings provide useful information regarding the true earnings process. We calibrate ρ to be 0.7 and σ_ϵ to be 0.016.¹⁵

¹³In the model, ERC can be computed using the formula in footnote 12 for each period. Empirically, this part of the analysis requires firm-quarter-level ERC estimates. Accordingly, we estimate a firm’s ERC in quarter t by regressing earnings announcement return on earnings surprise using a rolling window of the 12 quarters prior to t

¹⁴Empirically, Armstrong et al. (2015) find that managers retain approximately 92–94% of granted equity annually, while Edmans et al. (2017) document an even lower annual sale rates of 1–2%. Moreover, newly granted equity each year further offsets sales of vested shares. Overall, managers tend to hold the firm’s equity in the long-run, and thus managerial impatience does not appear to be a major concern in the sample.

¹⁵We calibrate these parameters using industry- and seasonally-adjusted restated earnings. Seasonal adjustment is performed using the X11-Autoregressive Integrated Moving Average (X11-ARIMA) procedure.

We are left with four structural parameters to be estimated: the likelihood of successful manipulation, $1 - q$; the regulator’s detection cost, κ ; and the manager’s manipulation costs, which include the regulatory penalty imposed upon detection, c , and the manager’s private cost, μ , which captures his reputation risks, career concerns, psychic costs, among many others that incur even without detection. These parameters are chosen to align the model’s key moments with their empirical counterparts, and we next discuss the key moments used in the estimation, how they are constructed both in the model and in the data, and how they help identify these parameters.

The first set of moments are the detection probabilities for firms in different ERC terciles. In the model, we record each detection event as a dummy $\mathbb{I}_{i,t}$ in the simulation, which equals one if firm i is detected in period t and zero otherwise. The probability of fraud detection for ERC tercile g is therefore defined as $E[\mathbb{I}_{i,t} | i \in g]$. In the data, we also compute the average detection likelihood for firms in each ERC tercile following the same approach.

The second set of moments include the magnitude of cumulative fraud upon detection and the change in ERC around detection event. In the data, for each detection event, we obtain the cumulative restatement amount associated with it. We measure cumulative fraud by aggregating the restatement’s total dollar impact on the income statement (as reported by Audit Analytics) and scaling it by total assets at the beginning of the earliest affected quarter disclosed in the announcement. Typically, the restatement periods cover 8 quarters instead of the whole period since last detection if ever. To ensure the same moment construction in the simulation, let t^* be the detection time for a detection event, then we define the cumulative fraud upon detection in the model as $\sum_{t=t^*-7}^{t^*} m_t(1 - q)$, which measures the total realized manipulation during the same 8-quarter period before the detection.

The change in ERC following detection is computed as $\frac{ERC_{t^*+1}}{ERC_{t^*-1}} - 1$ in the model in percentage term. Since our simulation tracks the evolution of ERC over time for each firm, it is straightforward to calculate its changes around the detection event. In the data, ERC also declines sharply post detection, and we measure the change in ERC caused by

the detection as the difference between the ERC right before the detection quarter and the trough of the ERC within 4 quarters post detection, again expressed in percentage term.

Finally, we measure the duration of ERC recovery as the time it takes for ERC to rise from the trough post detection to its peak within the subsequent 8 periods. We implement the same approach of measuring this moment in both the data and the simulation.

How do these moments help us pin down the model parameters? In SMM, all moments are used jointly to identify the parameters, and thus there is no one-to-one mapping between moments and parameters. However, some moments can be particularly informative of certain parameters or carry clear economic implications. To make the estimation process as transparent as possible, we offer some insight into the relation between the moments and parameters.

In the model, the detection cost κ has a first-order effect on the overall detection rate in the economy. A higher κ makes regulatory monitoring and detection more costly, thereby reducing the likelihood of detection across all firms. We therefore use the average detection rates across the low-, medium-, and high-ERC groups to pin down κ .

Both the regulatory penalty c and the manager's personal cost μ discourage fraudulent behavior, but they differ in how they influence firms' incentives, allowing us to identify them separately. Since c affects only detected firms, its deterrence effect is stronger for those under close regulatory scrutiny. Consequently, as c increases, the gap in detection rates between the low- and high-ERC groups widens. We exploit this moment to distinguish c from μ , which influences manipulation decisions more uniformly across firms.

Conditional on c and κ , the parameter μ governs both the cumulative fraud observed upon detection and the change in ERC after detection. As μ rises, managers face higher private costs of manipulation, leading them to commit less total fraud given the same detection intensity. Lower cumulative fraud reduces the noise in past earnings reports and increases the precision of investors' valuation beliefs, which in turn dampens the change in ERC following detection. Hence, we expect higher μ to be associated with smaller cumulative

fraud and smaller post-detection changes in ERC.

Finally, these two moments—cumulative fraud and ERC change—also help identify the likelihood of manipulation success, $1 - q$. As $1 - q$ increases, each unit of manipulation is more likely to be successfully incorporated into reported earnings, raising cumulative fraud, *ceteris paribus*. Meanwhile, the uncertainty generated by each unit of manipulation is proportional to $\sqrt{q(1 - q)}$. When $1 - q = 0.5$, this term reaches its maximum, implying that uncertainty is highest at this midpoint and declines as $1 - q$ moves toward either 0 or 1. Therefore, cumulative fraud upon detection is monotonic in $1 - q$ while the change in ERC around detection is hump-shaped in $1 - q$. Given the distinct responses of the two moments to μ and $1 - q$, they provide sufficient information to identify these two parameters.

2.2.3 Estimation Results

In this section, we present the estimation results for the six targeted moments, the point estimates of model parameters, and the overall model fit. Panel A of Table 1 indicates a close alignment between the model and the data for the targeted moments. Starting with the probability of fraud detection, both the model and empirical estimates show that firms in the bottom tercile of ERC exhibit the lowest detection rate, approximately 0.75% per quarter, which corresponds to an annualized fraud detection rate of 3%. In contrast, firms in the mid- and top-tercile experience higher detection rates of around 0.8 - 0.9% per quarter or 3.2-3.6% annually, suggesting increased regulatory scrutiny for these firms. These detection rates are estimated with high precision, as evidenced by the small standard errors associated with them. In addition, the differences between the model-implied detection rates and empirical counterparts are statistically insignificant.

Turning to the magnitude of cumulative fraud upon detection, the model-implied magnitude of cumulative fraud upon detection closely resembles its empirical counterpart, representing approximately 11% of total assets; the difference between the two is statistically insignificant. This moment is again estimated with high precision.

Third, the model also does a good job in capturing the decline in ERC caused by detection events in the simulation. The model predicts an average drop of 8% in ERC due to the reset of Φ after detection, and it matches closely the decrease in ERC observed in the data, which is 9%.

Lastly, our model closely mirrors the data in estimating the duration of uncertainty recovery. Specifically, the empirical estimate suggests that it takes an average of 5.6 quarters for a firm’s ERC to rebound from its lowest level to its highest level post-detection, a duration that aligns closely with the model’s implied number. The standard error for this moment, however, is large, indicating that the variation in recovery time is large in the data.

Panel B of Table 1 presents the parameter estimates, which minimize the distance between the model-implied moments and their empirical counterparts reported in Panel A. The coefficient of detection cost, κ , is estimated to be 0.03. To evaluate its magnitude, we note that the detection cost, $\frac{1}{2}\kappa d^2$, is specified as the regulator’s disutility relative to the aggregate information uncertainty that it attempts to minimize. Our estimate of $\kappa = 0.03$ suggests that, from the regulator’s perspective, the cost of increasing the detection rate by one percentage-point is equivalent to seeing the aggregate information uncertainty to increase by 1%.¹⁶

We estimate the coefficient of penalty on detected fraud, c , to be 2.7. This implies that for each unit of fraud detected, the imposed regulatory penalty is about five times the illegal gain managers get from manipulation. Additionally, we estimate the manager’s personal manipulation cost, μ , to be 0.37, which corresponds to approximately 14% of the regulatory penalty cost.

Although penalty on detection is extremely costly to managers when it occurs, detection is relatively infrequent. By contrast, personal costs are incurred whenever managers manipulate, regardless of whether the firm is ultimately detected, making them more influential for average manipulation incentives. That said, penalty on detection remains crucial for

¹⁶Increasing the detection rate by one percentage-point from its sample mean incurs an incremental cost of $\frac{1}{2}\kappa[(d + 0.01)^2 - d^2] = 4 \cdot 10^{-6}$, which is about 1% of the steady-state uncertainty $\Phi = 4 \cdot 10^{-4}$.

the dynamic behavior of manipulation. In particular, it is the threat of enforcement that generates the accumulation-and-reset pattern observed in the data. For this reason, penalty on detection lies at the core of the mechanism, while personal costs mainly exert a steady, level effect on manipulation over time.

Based on the estimated parameters, our model indicates that eradicating fraud is unlikely due to a dynamic interplay between regulatory actions and firm behavior. After detecting fraud, the regulator optimally reduces the scrutiny on the offending firm. However, this reduced oversight enables the firm to resume fraudulent activities. In equilibrium, firms facing heightened regulatory scrutiny decrease their manipulation, while those hiding under the radar increase their fraudulent activities. This leads to a convergence in fraud levels and information uncertainty across firms. Quantitatively, our estimation suggests that this convergence occurs rather swiftly, typically within six quarters.

Finally, with the estimated model, we implement a counterfactual in which manipulation is completely shut down by setting the personal cost of manipulation to infinity, so that no earnings manipulation occurs in equilibrium. We compare several key features of interest between the baseline economy (with fraud) and this no-fraud counterfactual: the volatility and persistence of returns, and the volatility and persistence of ERC. The results are summarized in Panel C of Table 1.

The return process is substantially more volatile in the presence of fraud. This is because manipulation adds an endogenous source of variation to reported earnings, and at the same time increases the uncertainty of investors' valuation of the firm, which amplifies the market reaction to earnings announcements. Moreover, returns exhibit a more negative autocorrelation with fraud. To understand this effect, first note that in the absence of fraud, the firm's report perfectly reveals economic earnings, so returns depend solely on contemporaneous earnings. Because economic earnings follow an AR(1) process, returns exhibit mean reversion and thus negative autocorrelation. With fraud, however, the firm's report introduces an additional manipulation noise, η_t , into investors' valuations. Past manipulation noises

continue to affect valuations as long as the fraud remains undetected. Over time, their influence gradually decays, since valuations increasingly reflect more recent reports and hence more recent realizations of η_t . This gradual decay generates additional mean reversion in returns beyond that implied by the AR(1) dynamics of economic earnings, leading to return autocorrelation that is even more negative than in the no-fraud case.

The behavior of ERC also differs sharply across the two economies. In the no-fraud counterfactual, ERC is constant in steady state, implying zero volatility and a persistence of one. With fraud, ERC becomes state-dependent because it is a function of Φ_{t-1} , which evolves over time and is reset upon detection. As a result, ERC exhibits positive volatility and lower persistence in the baseline economy. This highlights that earnings management not only affects the level of reported earnings, but also fundamentally alters how informative earnings are for prices.

2.3 Reduced-Form Evidence

To complement the structural estimates, we conduct two reduced-form analyses. The first examines the model’s prediction of an accumulation-and-reset cycle within a firm—or, equivalently, a reset-and-reaccumulation cycle. When fraud is detected (proxied by the quarter in which the firm announces a restatement), the uncertainty that has accumulated under undetected manipulation is largely resolved. Accordingly, the model predicts that the market should place little weight on subsequent earnings news, implying a sharp decline in the ERC at detection followed by a gradual recovery as manipulation and information uncertainty reaccumulate.

To examine the dynamics of the ERC following a restatement, we regress three-day earnings announcement returns on earnings surprise, interactions between earnings surprise and a set of post-detection indicators, and control variables. We include observations from four quarters prior to the detection quarter through all subsequent quarters, yielding 106,414 firm-quarter observations from 2,731 firms. The interaction terms capture how ERCs evolve after

detection relative to the quarters immediately preceding detection. Post-detection indicators are labeled $PD1$ through $PD12$, where $PD(k)$ equals one for the k th post-detection quarter ($k = 1, \dots, 12$) and $PD12$ pools the twelfth and all subsequent post-detection quarters. As is standard in event-time analyses, we group all quarters beyond the twelfth post-detection quarter into a single indicator to capture the long-term average ERC relative to the pre-detection period. We control for firm size, market-to-book ratio, a fourth-quarter indicator, and a loss indicator, and include interactions between these controls and earnings surprise. We include firm and quarter fixed effects and cluster standard errors by firm and quarter.

Table 2 reports the results. Column (1) shows that the post-detection dynamics are consistent with the model’s prediction that fraud-induced uncertainty resets upon detection and then gradually reaccumulates. Specifically, the ERC falls below its pre-detection baseline for at least four quarters after the announcement and then rebounds. In economic magnitudes, the average pre-detection ERC is 2.93. Relative to this baseline, the ERC is lower from quarters $t + 1$ through $t + 6$ (with changes ranging from -0.19 to -0.33), with the largest declines occurring in $t + 3$ and $t + 4$. By $t + 7$, the change turns positive to 0.27, consistent with a recovery.

Prior studies also document a rebound in the ERC after restatements (Wilson (2008); Chen et al. (2014)), although the persistence of the post-restatement decline varies across samples. This finding is often interpreted as a temporary loss of reporting credibility that gradually fades over time. We propose a different explanation for this finding: in our framework, the rebound arises endogenously from the joint dynamics of manipulation and detection and therefore need not reflect investors’ short memory.¹⁷

The second analysis examines the model’s multi-firm prediction for how accumulation-and-reset cycles overlap across firms. The three-firm model predicts that regulatory attention

¹⁷Wilson (2008) shows that the ERC is significantly below its pre-restatement baseline for the first three post-restatement earnings announcements, but not thereafter. Chen et al. (2014) show that after material restatements, the ERC decline is both more persistent (up to 11 quarters) and more economically meaningful (as much as 56% below baseline), but is largely confined to the first post-restatement quarter for less material restatements.

is endogenously allocated toward firms with the greatest cumulative fraud-induced uncertainty (i.e., H-firms). Following fraud detection, regulatory scrutiny concentrates on these firms, diverting attention away from lower-uncertainty firms (M- and L-firms) and allowing uncertainty at those firms to continue accumulating.

To test this prediction, for each firm that announces a fraud-related restatement in quarter t (the focal firm), we match it with control firms that do not announce a restatement during the subsequent eight-quarter event window. We implement one-to-two matching with replacement by selecting control firms from the same two-digit SIC industry with total assets and market-to-book ratios closest to those of the focal firm in quarter t . The matching procedure helps ensure that focal and matched firms exhibit similar characteristics but differ in regulatory scrutiny induced by detection at the focal firm. We then compare the evolution of ERCs across the two groups over the same event window.

Column (2) of Table 2 presents the result. Consistent with the model’s prediction, we observe distinct ERC dynamics across the two groups. Focal firms experience a pronounced decline in the ERC following detection, consistent with a reset of accumulated uncertainty, whereas matched firms exhibit no comparable decline. Moreover, the baseline ERC of matched firms (1.89) is substantially lower than that of focal firms (2.93), consistent with the model’s prediction that firms that are ultimately detected exhibit higher ERCs prior to detection. These patterns suggest that, absent detection, fraud-induced uncertainty continues to accumulate at matched firms, increasing the market’s reliance on earnings news.

This result supports the model’s cross-firm prediction that regulatory attention directed toward H-firms allows manipulation and uncertainty to continue accumulating at other firms. Importantly, the absence of an ERC decline among matched firms helps rule out alternative explanations based on macro or industry shocks. In this sense, the analysis also serves as a placebo test: only firms that are caught experience an ERC reset, while otherwise similar firms continue along an accumulation path.

3 Conclusion

This paper develops a dynamic theory of corporate fraud with endogenous regulatory detection. In the single-firm setting, the model generates an accumulation-and-reset cycle: undetected manipulation builds fraud-induced uncertainty, increasing the market’s reliance on earnings news and strengthening incentives to manipulate. Detection temporarily resolves this uncertainty, resetting the information environment but initiating a new accumulation cycle. Because regulatory effort responds optimally to accumulated uncertainty, fraud persists even under effective enforcement.

Extending the framework to a multi-firm setting highlights a cross-sectional mechanism. The regulator optimally reallocates detection resources towards firms with the greatest fraud-induced uncertainty. While this concentration disciplines manipulation at targeted firms, it simultaneously relaxes scrutiny at non-targeted firms, allowing their uncertainty to continue accumulating. This mechanism produces overlapping accumulation-and-reset cycles across firms and a “whack-a-mole” equilibrium in which fraud is displaced rather than eliminated.

Both the structural estimation and reduced-form evidence support these predictions. The estimated model matches observed detection frequencies, cumulative fraud at detection, and the pace of post-detection recovery. In reduced-form analyses, the average ERC drops sharply following fraud detection and then gradually recovers. In contrast, matched firms without detection exhibit no comparable decline-and-rebound pattern in the ERC.

Overall, the results suggest that persistence in corporate fraud need not reflect weak enforcement. Instead, rational, targeted detection corrects severe information distortions when they become most acute, yet cannot prevent manipulation and information uncertainty from reaccumulating because regulatory attention shifts to and away from firms. In this sense, corporate fraud persists not because anti-fraud regulation fails; quite to the contrary, fraud is “everlasting” precisely because the regulatory approach is optimal.

References

- Armstrong, Chris, John E Core, and Wayne R Guay, 2015, Why do CEOs hold so much equity?, *Unpublished Manuscript* .
- Bertomeu, Jeremy, Edwige Cheynel, Edward Xuejun Li, and Ying Liang, 2021, How pervasive is earnings management? evidence from a structural model, *Management Science* 67, 5145–5162.
- Bertomeu, Jeremy, and Robert P Magee, 2011, From low-quality reporting to financial crises: Politics of disclosure regulation along the economic cycle, *Journal of Accounting and Economics* 52, 209–227.
- Beyer, Anne, Ilan Guttman, and Iván Marinovic, 2019, Earnings management and earnings quality: Theory and evidence, *The Accounting Review* 94, 77–101.
- Bird, Andrew, Stephen A Karolyi, and Thomas G Ruchti, 2019, Understanding the “numbers game”, *Journal of Accounting and Economics* 68, 101242.
- Bourveau, Thomas, Matthias Breuer, and Robert Stoumbos, 2025, Learning to disclose: Disclosure dynamics in the 1890s streetcar industry, *Review of Financial Studies* Forthcoming.
- Chen, Chuan, Yanrong Jia, Xiumin Martin, and Bernardo Silveira, 2025, Assessing the objective function of the sec against financial misconduct: A structural approach, *Journal of Accounting and Economics* 101794.
- Chen, Xia, Qiang Cheng, and Alvis K Lo, 2014, Is the decline in the information content of earnings following restatements short-lived?, *The Accounting Review* 89, 177–207.
- Cheynel, Edwige, Davide Cianciaruso, and Frank S Zhou, 2024, Fraud power laws, *Journal of Accounting Research* 62, 833–876.
- Corona, Carlos, and Ramandeep S Randhawa, 2018, The value of confession: Admitting mistakes to build reputation, *The Accounting Review* 93, 133–161.

- Dye, Ronald A, and Gaoqing Zhang, 2025, Whack-a-mole accounting regulation and multidimensional earnings management, *Unpublished Manuscript* .
- Edmans, Alex, Vivian W Fang, and Katharina A Lewellen, 2017, Equity vesting and investment, *Review of Financial Studies* 30, 2229–2271.
- Fang, Vivian W, Allen H Huang, and Wenyu Wang, 2017, Imperfect accounting and reporting bias, *Journal of Accounting Research* 55, 919–962.
- Fischer, Paul E, and Robert E Verrecchia, 2000, Reporting bias, *The Accounting Review* 75, 229–245.
- Gao, Pingyang, and Gaoqing Zhang, 2019, Accounting manipulation, peer pressure, and internal control, *The Accounting Review* 94, 127–151.
- Hail, Luzi, Ahmed Tahoun, and Clare Wang, 2018, Corporate scandals and regulation, *Journal of Accounting Research* 56, 617–671.
- Holzman, Eric R, Nathan T Marshall, and Brent A Schmidt, 2024, When are firms on the hot seat? an analysis of sec investigation preferences, *Journal of Accounting and Economics* 77, 101610.
- Karpoff, Jonathan M, D Scott Lee, and Gerald S Martin, 2008, The consequences to managers for financial misrepresentation, *Journal of Financial Economics* 88, 193–215.
- Terry, Stephen J, 2023, The macro impact of short-termism, *Econometrica* 91, 1881–1912.
- Wilson, Wendy M, 2008, An empirical analysis of the decline in the information content of earnings following restatements, *The Accounting Review* 83, 519–548.
- Zakolyukina, Anastasia A, 2018, How common are intentional GAAP violations? estimates from a dynamic model, *Journal of Accounting Research* 56, 5–44.

Table 1: Estimation Results

This table reports the overall model fit (Panel A), parameter estimates (Panel B), and a comparison between the baseline economy and the no-fraud counterfactual (Panel C). Panel A sorts firms into tercile groups based on their beginning-of-the-quarter ERC, and H (M, L) indicates firms with a high (medium, low) level of ERC. The targeted moments include the quarterly likelihood of detection, cumulative fraud upon detection, the change in ERC upon detection, and the duration of ERC recovery. Panel C reports the volatility and persistence of returns and ERC under the baseline and the no-fraud counterfactual economies.

Panel A. Model Fit

Moments				
		Model	Data	S.E.
Detection probability sorted by ERC (%)	L	0.75	0.73	0.04
	M	0.84	0.80	0.04
	H	0.86	0.83	0.04
Cumulative fraud upon detection		0.118	0.109	0.012
Change in ERC upon detection		-0.082	-0.091	0.041
Duration of ERC recovery (qtr)		6.00	5.60	4.11

Panel B. Parameter Estimates

Pre-set parameters			Value	
δ, β_R, β_F	discount rate for the market, regulator, and manager		0.974	
ρ	persistence of earnings		0.7	
σ_ϵ	conditional volatility of earnings		0.016	
Estimated parameters			Value	S.E.
κ	detection cost		0.03	0.006
c	penalty on detected fraud		2.7	0.33
μ	variable cost of manipulation		0.37	0.08
$1 - q$	probability of successful manipulation		0.56	0.07

Panel C. Counterfactual: Baseline vs. No-Fraud

	Std. Dev.		Autocorrelation	
	Baseline	No Fraud	Baseline	No Fraud
Return	0.4008	0.0511	-0.4736	-0.1489
ERC	0.0067	0.0000	0.3254	1.0000

Table 2: ERC Dynamics Following Restatements

This table reports the ordinary least squares (OLS) regression results examining the dynamics of ERC following a restatement. *CAR* is the three-day market-adjusted return centered on the earnings announcement date. *SUE* denotes earnings surprise, defined as the quarterly reported earnings per share (EPS) minus the I/B/E/S median consensus EPS forecast, scaled by the stock price two days prior to the announcement. *PD(k)* is an indicator equal to one for the k th quarter after the restatement, where $k = 1, \dots, 12$, and *PD12* pools the twelfth and all subsequent post-detection quarters. For matched control firms, *PD(k)* is defined analogously using the focal firm's event date (that is, relative to the matched restating firm). Standard errors are clustered by firm and quarter. T-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Sample Variable	(1) Detected Firms CAR	(2) Matched Control CAR
SUE	2.928*** (12.42)	1.898*** (3.26)
SUE×PD1	-0.236** (-2.03)	0.116 (0.53)
SUE×PD2	-0.186 (-1.64)	0.203 (0.91)
SUE×PD3	-0.328** (-2.59)	-0.039 (-0.17)
SUE×PD4	-0.315** (-2.27)	-0.151 (-0.72)
SUE×PD5	-0.198 (-1.65)	-0.247 (-1.16)
SUE×PD6	-0.193 (-1.34)	-0.041 (-0.16)
SUE×PD7	0.271* (1.89)	0.084 (0.35)
SUE×PD8	-0.164 (-1.23)	0.023 (0.14)
SUE×PD9	-0.068 (-0.48)	0.298 (1.17)
SUE×PD10	0.253** (2.06)	0.174 (0.74)
SUE×PD11	0.071 (0.42)	0.318 (1.62)
SUE×PD12	-0.108 (-0.55)	0.242 (1.17)
Observations	106,414	137,582
Adjusted R-squared	0.145	0.124
Control Variables	Yes	Yes
Firm Fixed Effects	Yes	Yes
Quarter Fixed Effects	Yes	Yes
Two-way Clustering	Yes	Yes

For Online Publication

Internet Appendix of “Everlasting Fraud”

Appendix A. Model Setup and Proofs

The section contains proofs that supplement the ones already presented in the main text.

A.1 Single-Firm Myopic Model

The single-firm myopic model setup is presented in the main text of the paper, and in this section, we provide proofs to the Lemmas and Propositions developed in the single-firm myopic model.

A.1.1 Proof of Lemma 2

Proof. of Lemma 2: We only consider the case in which the detection fails:

$$\begin{aligned}\Phi_t &= \text{var}(s_t | \mathcal{F}_{t-1}) - \text{var}(E^I[s_t | \mathcal{F}_t] | \mathcal{F}_{t-1}) \\ &= \text{var}(\rho s_{t-1} + \varepsilon_t | \mathcal{F}_{t-1}) - \text{var}(E^I[s_t | r_t, r_{t-1}, \dots] | \mathcal{F}_{t-1}) \\ &= \rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 - \text{var}\left(\frac{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2}{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 + m_t^* q (1 - q)} r_t | \mathcal{F}_{t-1}\right) \\ &= \rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 - \left[\frac{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2}{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 + m_t^* q (1 - q)}\right]^2 \text{var}(r_t | \mathcal{F}_{t-1}) \\ &= \frac{m_t^* q (1 - q) [\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2]}{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 + m_t^* q (1 - q)} \\ &= \frac{m_t^* q (1 - q) (\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2)}{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2 + m_t^* q (1 - q)}.\end{aligned}\tag{A.1}$$

The first equality uses the law of total variance. The third equality uses

$$\begin{aligned} E^I [s_t | \mathcal{F}_t] &= E^I [s_t | \mathcal{F}_{t-1}] + \frac{\text{cov}(r_t, s_t | \mathcal{F}_{t-1})}{\text{var}(r_t | \mathcal{F}_{t-1})} \{r_t - E[r_t | \mathcal{F}_{t-1}]\} \\ &= E^I [s_t | \mathcal{F}_{t-1}] + \frac{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2}{\rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 + m_t^* q (1 - q)} \{r_t - E[r_t | \mathcal{F}_{t-1}]\}, \end{aligned} \quad (\text{A.2})$$

where

$$\begin{aligned} \text{var}(r_t | \mathcal{F}_{t-1}) &= \text{var}(s_t | \mathcal{F}_{t-1}) + m_t^* q (1 - q) \\ &= \rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2 + m_t^* q (1 - q), \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} \text{cov}(r_t, s_t | \mathcal{F}_{t-1}) &= \text{var}(s_t | \mathcal{F}_{t-1}) \\ &= \rho^2 \text{var}(s_{t-1} | \mathcal{F}_{t-1}) + \sigma_\varepsilon^2. \end{aligned} \quad (\text{A.4})$$

The last step uses the definition of $\Phi_{t-1} \equiv \text{var}(s_{t-1} | \mathcal{F}_{t-1})$. ■

A.1.2 Proof of Lemma 3

Proof. of Lemma 3: We first prove that there exists a unique solution of $d_t^* \in (0, 1)$ to equation (17). Note first that at one extreme of $d_t^* = 0$, the left-hand side of equation (17) is 0 but the right-hand side is positive. At the other extreme of $d_t^* = 1$, the left-hand side of equation (17) is 1 but the right-hand side turns negative. By the intermediate value theorem, there must exist a solution $d_t^* \in (0, 1)$. In addition, the left-hand side of equation (17) is increasing in d_t^* whereas the right-hand side is decreasing. Such monotonicity, in turn, implies that the solution of d_t^* is unique in the interval $(0, 1)$. Lastly, we prove that d_t^* increases in Φ_{t-1} . Applying the implicit function theorem to equation (17) yields that:

$$\frac{\partial d_t^*}{\partial \Phi_{t-1}} = \frac{\rho^2 \left(1 - \frac{(1-\delta\rho)cd_t^*(\Phi_{t-1})}{1-d_t^*(\Phi_{t-1})} \right)}{\kappa} > 0. \quad (\text{A.5})$$

■

A.1.3 Proof of Proposition 1

Proof. of Proposition 1: We first reproduce the F.O.C. of m_t^* and d_t^* from Lemmas 1 and 3 below:

$$\begin{aligned} m_t^*(\Phi_{t-1}) &= \frac{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2}{q(1-q)} \left(\frac{1-d_t^*}{cd_t^*} \frac{1}{1-\delta\rho} - 1 \right), \\ d_t^*(\Phi_{t-1}) &= \frac{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2}{\kappa} \left(1 - (1-\delta\rho) \frac{cd_t^*}{1-d_t^*} \right). \end{aligned}$$

Dividing the two equations yields:

$$\frac{m_t^*(\Phi_{t-1})}{d_t^*(\Phi_{t-1})} = \frac{\kappa}{q(1-q)} \frac{\frac{1-d_t^*}{cd_t^*} \frac{1}{1-\delta\rho} - 1}{1 - (1-\delta\rho) \frac{cd_t^*}{1-d_t^*}}. \quad (\text{A.6})$$

Solving m_t^* as a function of d_t^* gives:

$$m_t^*(\Phi_{t-1}) = \frac{\kappa}{cq(1-q)} \frac{1-d_t^*(\Phi_{t-1})}{1-\delta\rho}. \quad (\text{A.7})$$

From Lemma 3, d_t^* is increasing in Φ_{t-1} , which implies that m_t^* is decreasing in Φ_{t-1} . ■

A.2 Single-Firm Full Dynamic Model

In this appendix, we present the full model with the forward-looking manager and regulator, and a personal manipulation cost (i.e., $\mu > 0$). For expositional purposes, we will focus on the additional economic forces brought up by the forward-looking features.

We begin by solving the manager's optimal choice of manipulation, m_t^* in each period t . Both the contemporaneous and the intertemporal effects of m_t on the investors' expectation of the firm's economic earnings are identical with those in the myopic model, as characterized in equation (12) of the main text. The forward-looking model, however, differs from the myopic model in that the manager now cares both the contemporaneous and the intertemporal effects of his manipulation choice. To illustrate, taking derivative of the manager's payoff

U_t in equation (5) with respect to m_t and then substituting in equation (12), we obtain the F.O.C. as

$$\underbrace{c(1-q)d_t^* + (1-q)\mu}_{\text{MC of } m_t} = \underbrace{(1-d_t^*)\frac{w_t(1-q)}{1-\delta\rho}}_{\text{MB of } m_t \text{ from the contemporaneous effect}} + \underbrace{\sum_{k=1}^{\infty} \beta_F^k \left[\prod_{\ell=0}^k (1-d_{t+\ell}^*) \right] \frac{\rho^k \left[\prod_{\ell=1}^k (1-w_{t+\ell}) \right] w_t(1-q)}{1-\delta\rho}}_{\text{MB of } m_t \text{ from the intertemporal effect}}. \quad (\text{A.8})$$

The MC of manipulation is similar to that in the baseline myopic model, except for the personal manipulation cost μ that we explicitly incorporate here. The MB of manipulation now includes both the contemporaneous effect and the intertemporal effect of m_t . While the MB from the contemporaneous effect is only affected by the likelihood of not being detected in the current period $(1-d_t^*)$, the MB from the intertemporal effect is affected by the likelihood of not being detected up to a future period of interest $\left[\prod_{\ell=0}^k (1-d_{t+\ell}^*) \right]$.

Augmenting the MB of manipulation with its intertemporal effects in the forward-looking model relative to the myopic model results in two new features of the equilibrium manipulation. First, the augmented MB of manipulation leads to a quantitatively higher level of manipulation in the forward-looking model. This quantitative change not only matters for our later calibration and structural estimation exercises, but may also cause qualitative changes in the equilibrium characteristics of manipulation, most notably, in overturning the monotone relationship between the past fraud-induced uncertainty Φ_{t-1} and current manipulation choice m_t^* under the myopic model (Proposition 1).

Second, the inclusion of the intertemporal effects leads to a positive association between the manager's manipulation choices over time. Specifically, substituting equation (11) for $w_{t+\ell}$ in equation (A.8) and solving for $m_{t+\ell}^*$, we find that the manager's current manipulation choice m_t^* is strictly increasing in his next-period manipulation choice m_{t+1}^* (and by induction, all the future manipulations $\{m_{t+2}^*, m_{t+3}^*, \dots\}$), stated in a recursive form in Lemma 4 below.

Lemma 4 *In each period t , given the regulator's equilibrium detection choice d_t^* conjectured by the manager, the manager chooses the optimal manipulation*

$$m_t^* = \frac{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2}{q(1-q)} \left[\frac{1 - d_t^*}{cd_t^* + \mu} \left(\frac{1}{1 - \delta\rho} + \beta_F \rho q (1 - q) \frac{(cd_{t+1}^* + \mu)m_{t+1}^*}{\rho^2 \Phi_t + \sigma_\varepsilon^2} \right) - 1 \right], \quad (\text{A.9})$$

where Φ_t is expressed recursively in equation (7).

Appendix A.2.1 in the next subsection provides all the proofs for the formal statements in Appendix A.2.

The intuition for the intertemporal correlation between the manager's manipulation choices is as follows. The manager's current manipulation choice m_t^* is disciplined by his anticipation of the sequence of future detection choices in equilibrium $\{d_{t+1}^*, d_{t+2}^*, \dots\}$ because these detection choices affect the MB of manipulation in period t as in equation (A.8). Analogously, the manager's next-period manipulation choice m_{t+1}^* also depends on the same sequence of future detection choices, $\{d_{t+1}^*, d_{t+2}^*, \dots\}$. In that vein, regulatory detections contribute to a positive association between manipulation choices over time.

Conditional on the sequence of the equilibrium manipulation choices $\{m_t^*, m_{t+1}^*, \dots\}$, the law of motion that governs the cumulative fraud-induced uncertainty Φ_t admits the same functional form as equation (15) in the myopic model. We thus skip that derivation and turn to the regulator's choice of optimal detection likelihood, d_t^* . Taking the derivative of the regulator's payoff function W_t in equation (9) yields the F.O.C. as

$$\underbrace{\kappa d_t}_{\text{MC of } d_t} = \underbrace{\Phi_t - 0}_{\text{MB of } d_t \text{ from the contemporaneous effect}} + \underbrace{\beta_R [W_{t+1}(0) - W_{t+1}(\Phi_t)]}_{\text{MB of } d_t \text{ from the intertemporal effect}}. \quad (\text{A.10})$$

$W_{t+1}(\Phi_t)$ denotes the regulator's objective function evaluated at an initial level of uncertainty Φ_t . Compared to the F.O.C. (17) in the myopic model, the MC of detection in the full model is the same whereas the MB of detection continues to be increasing in the fraud-

induced uncertainty Φ_t but is quantitatively larger, as it includes both the contemporaneous and intertemporal effects of detection in clearing uncertainty in the current period and decreasing prior uncertainty for all future periods (via the law of motion of Φ_t). Specifically, $W_{t+1}(0) - W_{t+1}(\Phi_t) > 0$ represents the capitalized value of resetting the initial uncertainty from Φ_t to zero for all future periods upon successful detection.

Lastly, we combine the F.O.C. for the manipulation and the detection choices to characterize the equilibrium under the full model. Because our model features an infinite horizon and both m_t^* and d_t^* can be written recursively as functions of Φ_{t-1} , we can treat Φ_{t-1} as the state variable for period t and characterize the equilibrium as a dynamic programming problem with the Bellman equations.

Proposition 3 *For a given level of accumulated past fraud Φ , the regulator's equilibrium detection choice $d^*(\Phi)$ and the manager's equilibrium manipulation choice $m^*(\Phi)$ are given by the following set of equations, with the two agents rationally anticipating each other's optimal policy function:*

$$d^*(\Phi) = \frac{\Phi' + \beta_R [W(0) - W(\Phi')]}{\kappa}, \quad (\text{A.11})$$

$$m^*(\Phi) = \frac{\rho^2 \Phi + \sigma_\varepsilon^2}{q(1-q)} \left[\frac{1 - d^*(\Phi)}{cd^*(\Phi) + \mu} \left(\frac{1}{1 - \delta\rho} + \beta_F \rho q (1 - q) \frac{(cd^*(\Phi') + \mu)m^*(\Phi')}{\rho^2 \Phi' + \sigma_\varepsilon^2} \right) - 1 \right], \quad (\text{A.12})$$

where we denote variables of the next period with a prime, and

$$\Phi'(\Phi) = \frac{m^*(\Phi) q (1 - q) (\rho^2 \Phi + \sigma_\varepsilon^2)}{\rho^2 \Phi + \sigma_\varepsilon^2 + m^*(\Phi) q (1 - q)}, \quad (\text{A.13})$$

$$W(\Phi) = -(1 - d^*)\Phi' - \frac{\kappa}{2} (d^*)^2 + \beta_R [d^* W(0) + (1 - d^*) W(\Phi')]. \quad (\text{A.14})$$

The dynamic programming problem in Proposition 3 does not permit a fully analytical solution so we solve it numerically. The main predictions and implications of the single-firm myopic model carry over to the full-blown model.

A.2.1 Proof of Lemma 4

Proof. of Lemma 4: Note that (A.8) can be simplified into

$$\begin{aligned}
& \frac{(1 - \delta\rho)(cd_t^* + \mu)}{1 - d_t^*} \left(1 + \frac{m_t^* q (1 - q)}{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2} \right) \\
&= 1 + \beta_F \rho (1 - d_{t+1}^*) \frac{m_{t+1}^* q (1 - q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1 - q)} \\
&+ \beta_F^2 \rho^2 (1 - d_{t+1}^*) (1 - d_{t+2}^*) \frac{m_{t+2}^* q (1 - q)}{\rho^2 \Phi_{t+1} + \sigma_\varepsilon^2 + m_{t+2}^* q (1 - q)} \frac{m_{t+1}^* q (1 - q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1 - q)} \\
&+ \dots
\end{aligned} \tag{A.15}$$

By induction, in period $t+1$, conditional on that the regulator fails to detect fraud in period t , the manager chooses m_{t+1}^* that satisfies:

$$\begin{aligned}
& \frac{(1 - \delta\rho)(cd_{t+1}^* + \mu)}{1 - d_{t+1}^*} \left(1 + \frac{m_{t+1}^* q (1 - q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2} \right) \\
&= 1 + \beta_F \rho (1 - d_{t+2}^*) \frac{m_{t+2}^* q (1 - q)}{\rho^2 \Phi_{t+1} + \sigma_\varepsilon^2 + m_{t+2}^* q (1 - q)} \\
&+ \beta_F^2 \rho^2 (1 - d_{t+2}^*) (1 - d_{t+3}^*) \frac{m_{t+3}^* q (1 - q)}{\rho^2 \Phi_{t+2} + \sigma_\varepsilon^2 + m_{t+3}^* q (1 - q)} \frac{m_{t+2}^* q (1 - q)}{\rho^2 \Phi_{t+1} + \sigma_\varepsilon^2 + m_{t+2}^* q (1 - q)} \\
&+ \dots
\end{aligned} \tag{A.16}$$

Multiplying both sides of equation (A.16) by $\beta_F \rho (1 - d_{t+1}^*) \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1-q)}$ yields:

$$\begin{aligned}
& \beta_F \rho (1 - \delta \rho) (cd_{t+1}^* + \mu) \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2} \\
&= \beta_F \rho (1 - d_{t+1}^*) \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1-q)} \\
&+ \beta_F^2 \rho^2 (1 - d_{t+2}^*) (1 - d_{t+1}^*) \frac{m_{t+2}^* q (1-q)}{\rho^2 \Phi_{t+1} + \sigma_\varepsilon^2 + m_{t+2}^* q (1-q)} \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1-q)} \\
&+ \beta_F^3 \rho^3 (1 - d_{t+3}^*) (1 - d_{t+2}^*) (1 - d_{t+1}^*) \frac{m_{t+3}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+3}^* q (1-q)} \\
&\times \frac{m_{t+2}^* q (1-q)}{\rho^2 \Phi_{t+1} + \sigma_\varepsilon^2 + m_{t+2}^* q (1-q)} \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2 + m_{t+1}^* q (1-q)} \\
&+ \dots
\end{aligned} \tag{A.17}$$

Substituting (A.17) into (A.15) yields:

$$\frac{(1 - \delta \rho) (cd_t^* + \mu)}{1 - d_t^*} \left(1 + \frac{m_t^* q (1-q)}{\rho^2 \Phi_{t-1} + \sigma_\varepsilon^2} \right) = 1 + \beta_F \rho (1 - \delta \rho) (cd_{t+1}^* + \mu) \frac{m_{t+1}^* q (1-q)}{\rho^2 \Phi_t + \sigma_\varepsilon^2}. \tag{A.18}$$

Solving (A.18) for m_t yields (A.9) in the lemma. ■

A.2.2 Proof of Proposition 3

Proof. of Proposition 3: The expressions for m^* , and Φ' have been derived earlier so we focus on deriving the expressions for the regulator's value function W and the F.O.C. of d^* .

Using the law of motion (15), we rewrite the regulator's payoff (9) recursively:

$$\begin{aligned}
W_t(\Phi_{t-1}) &= \max_{d_t} - (1 - d_t) \Phi_t(d_t^*, \Phi_{t-1}) - \frac{\kappa}{2} d_t^2 \\
&+ \beta_F E^I \left[\sum_{k=t+1}^{\infty} \beta_F^{k-(t+1)} \left(- (1 - d_k^*) \Phi_k - \frac{\kappa}{2} d_k^{*2} \right) \right] \\
&= \max_{d_t} - (1 - d_t) \Phi_t(d_t^*, \Phi_{t-1}) - \frac{\kappa}{2} d_t^2 \\
&+ \beta_F [d_t W_{t+1}(0) + (1 - d_t) W_{t+1}(\Phi_t(d_t^*, \Phi_{t-1}))],
\end{aligned} \tag{A.19}$$

where $\Phi_t(d_t^*, \Phi_{t-1})$ is given in equation (15). Note that the future cumulative level of fraud Φ_t depends on the equilibrium detection probability d_t^* and not on the actual detection probability d_t . This is because, the manager does not observe the regulator's detection choice at the time of choosing manipulation and his manipulation choice only depends on the equilibrium d_t^* . Dropping the time subscript yields the expression of W . Taking the derivative of W with respect to d yields the results. ■

A.3 Multi-Firm Myopic Model

The multi-firm myopic model setup is presented in the main text of the paper, and in this section, we provide proofs to the Lemmas and Propositions developed in the multi-firm myopic model.

A.3.1 Proof of Proposition 2

Proof. of Proposition 2: With zero discount factors and dropping the time subscript, the regulator's objective function becomes:

$$W\left(\{\Phi_i\}_{i \in \{1,2\}}\right) = -\sum_{i=1}^2 (1-d_i) \Phi'_i - \frac{\kappa}{2} \left(\sum_{i=1}^2 d_i\right)^2. \quad (\text{A.20})$$

In addition, following similar steps as in the single-firm case, we can derive the same law of motion for Φ_i (as in equation (16)) into:

$$\Phi'_i \equiv (\rho^2 \Phi_i + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_i^*}{1 - d_i^*}\right). \quad (\text{A.21})$$

Taking the first-order condition gives that

$$\frac{\partial W}{\partial d_i} = \Phi'_i - \kappa \left(\sum_{i=1}^2 d_i\right). \quad (\text{A.22})$$

Note that since $\Phi_1 \geq \Phi_2$, $\rho^2 \Phi_1 + \sigma_\varepsilon^2 \geq \rho^2 \Phi_2 + \sigma_\varepsilon^2$. Now consider two cases. First, suppose

that

$$\rho^2\Phi_2 + \sigma_\varepsilon^2 < (\rho^2\Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right), \quad (\text{A.23})$$

that is, Φ_1 is much larger than Φ_2 . We will restate condition (A.23) in terms of exogenous parameters after solving the equilibrium. We now conjecture the equilibrium is that $d_2^* = 0$ and $d_1^* > 0$, where d_1^* solves:

$$\Phi_1' = (\rho^2\Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) = \kappa d_1^*. \quad (\text{A.24})$$

To verify that this is indeed an equilibrium, note first that the solution to (A.24) is unique because the left-hand side is decreasing in d_1^* whereas the right-hand side is increasing in d_1^* . In addition, by the implicit function theorem, since the left-hand side is increasing in Φ_1 , d_1^* is increasing in Φ_1 . Next, using the first-order condition (A.24), we can rewrite the condition (A.23) as:

$$\kappa d_1^* = \Phi_1' = (\rho^2\Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) > \rho^2\Phi_2 + \sigma_\varepsilon^2. \quad (\text{A.25})$$

Since d_1^* is increasing in Φ_1 , the condition (A.23) holds if and only if Φ_1 is sufficiently large and/or Φ_2 is sufficiently small. In other words, we can rewrite the condition (A.23) as

$$\Phi_1 > H(\Phi_2), \quad (\text{A.26})$$

where $H(\cdot)$ is some given increasing function. Finally, we verify that $d_2^* = 0$. This is because,

at $d_2 = 0$, the first-order condition for d_2 is always negative, i.e.,

$$\begin{aligned}
\frac{\partial W}{\partial d_2} &= \Phi'_2 - \kappa d_1^* \\
&= (\rho^2 \Phi_2 + \sigma_\varepsilon^2) - \kappa d_1^* \\
&< (\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) - \kappa d_1^* \\
&= 0.
\end{aligned} \tag{A.27}$$

The third step uses (A.23). The last step uses (A.24).

Second, suppose that $\Phi_1 \leq H(\Phi_2)$, that is, Φ_1 and Φ_2 are of similar sizes. We now conjecture the equilibrium is that $d_1^* > 0$ and $d_2^* > 0$, where the pair of $\{d_1^*, d_2^*\}$ solves:

$$\Phi'_1 = (\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) = \kappa D^*, \tag{A.28}$$

$$\Phi'_2 = (\rho^2 \Phi_2 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_2^*}{1 - d_2^*} \right) = \kappa D^*, \tag{A.29}$$

where $D^* = d_1^* + d_2^*$. To prove this, note first that $d_1^* = d_2^* = 0$ cannot be an equilibrium because in this case, the regulator would incur a zero MC of detection but a strictly positive MB of detection of $\rho^2 \Phi_1 + \sigma_\varepsilon^2 > 0$. In addition, $d_1^* = 0$ and $d_2^* > 0$ cannot be an equilibrium. To prove this, assume by contradiction that $d_2^* > 0$. The interior solution of d_2^* must solve:

$$\Phi'_2 = (\rho^2 \Phi_2 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_2^*}{1 - d_2^*} \right) = \kappa d_2^*. \tag{A.30}$$

In addition, at $d_1^* = 0$, the MC of increasing d_1^* is κd_2^* whereas the MB is $\rho^2 \Phi_1 + \sigma_\varepsilon^2$. Note that

$$\begin{aligned}
\rho^2 \Phi_1 + \sigma_\varepsilon^2 &> \rho^2 \Phi_2 + \sigma_\varepsilon^2 \\
&> (\rho^2 \Phi_2 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_2^*}{1 - d_2^*} \right) \\
&> \kappa d_2^*.
\end{aligned} \tag{A.31}$$

As a result, the regulator would not choose $d_1^* = 0$. A contradiction. Lastly, we prove that $d_1^* > 0$ and $d_2^* = 0$ cannot be an equilibrium. Assume by contradiction that $d_1^* > 0$, which must solve:

$$\Phi_1' = (\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) = \kappa d_1^*. \quad (\text{A.32})$$

In addition, at $d_2^* = 0$, the MC of increasing d_2^* is κd_1^* whereas the MB is $\rho^2 \Phi_2 + \sigma_\varepsilon^2$. Since condition (A.23) is violated under the case of $\Phi_1 \leq H(\Phi_2)$, it must be that

$$\kappa d_1^* = (\rho^2 \Phi_1 + \sigma_\varepsilon^2) \left(1 - (1 - \delta\rho) \frac{cd_1^*}{1 - d_1^*} \right) < \rho^2 \Phi_2 + \sigma_\varepsilon^2. \quad (\text{A.33})$$

As a result, the regulator would not choose $d_2^* = 0$. A contradiction.

Lastly, we prove the comparative statics of $\{d_1^*, d_2^*\}$. Since the left-hand sides of the two first-order conditions (A.28) and (A.29) of $\{d_1^*, d_2^*\}$ are increasing in Φ_1 and Φ_2 , respectively, applying the implicit function theorem gives that D^* is strictly increasing in Φ_1 and Φ_2 . In addition, applying the implicit function theorem gives that d_1^* is increasing in Φ_1 and decreasing in Φ_2 whereas d_2^* is increasing in Φ_2 and decreasing in Φ_1 . ■

A.4 Multi-Firm Full Model

In this appendix, we extend the multi-firm myopic model to the full model with forward-looking agents. This is the model we will estimate to fit the data. We increase the number of firms from two to three, and formally state the three-firm equilibrium characterized as the solution to a dynamic programming problem with the Bellman equations below. For ease of notation, we again omit the time subscript and denote variables of the next period with a prime.

Proposition 4 *Consider a three-firm full model. Given the levels of accumulated past fraud*

at the three firms $\{\Phi_1, \Phi_2, \Phi_3\}$, the manager in firm 1 chooses manipulation

$$m_1^*(\Phi_1, \Phi_2, \Phi_3) = \frac{\rho^2 \Phi_1 + \sigma_\varepsilon^2}{q(1-q)} \left(\frac{1-d_1^*}{cd_1^* + \mu} \left(\frac{1}{1-\delta\rho} + \frac{\beta_F \rho q(1-q)}{\rho^2 \Phi_1 + \sigma_\varepsilon^2} \times E[m_1^{*'}(cd_1^{*'} + \mu)] \right) - 1 \right), \quad (\text{A.34})$$

and the regulator chooses to detect fraud at firm 1 with probability

$$\begin{aligned} d_1^*(\Phi_1, \Phi_2, \Phi_3) = & \frac{1}{\kappa} \{ \Phi_1' + \beta_R(1-d_2^*)(1-d_3^*) [W(0, \Phi_2', \Phi_3') - W(\Phi_1', \Phi_2', \Phi_3')] \\ & + \beta_R(1-d_2^*)d_3^* [W(0, \Phi_2', 0) - W(\Phi_1', \Phi_2', 0)] \\ & + \beta_R d_2^*(1-d_3^*) [W(0, 0, \Phi_3') - W(\Phi_1', 0, \Phi_3')] \\ & + \beta_R d_2^* d_3^* [W(0, 0, 0) - W(\Phi_1', 0, 0)] \} - (d_2^* + d_3^*), \end{aligned} \quad (\text{A.35})$$

where

$$\Phi_i'(\Phi_1, \Phi_2, \Phi_3) = \frac{m_i^* q(1-q)(\rho^2 \Phi_i + \sigma_\varepsilon^2)}{\rho^2 \Phi_i + \sigma_\varepsilon^2 + m_i^* q(1-q)}, \quad (\text{A.36})$$

$$\begin{aligned} E[m_1^{*'}(cd_1^{*'} + \mu)] = & (1-d_2^*)(1-d_3^*)m_1^*(\Phi_1', \Phi_2', \Phi_3')(cd_1^*(\Phi_1', \Phi_2', \Phi_3') + \mu) \\ & + d_2^*(1-d_3^*)m_1^*(\Phi_1', 0, \Phi_3')(cd_1^*(\Phi_1', 0, \Phi_3') + \mu) \\ & + (1-d_2^*)d_3^*m_1^*(\Phi_1', \Phi_2', 0)(cd_1^*(\Phi_1', \Phi_2', 0) + \mu) \\ & + d_2^*d_3^*m_1^*(\Phi_1', 0, 0)(cd_1^*(\Phi_1', 0, 0) + \mu), \end{aligned} \quad (\text{A.37})$$

$$\begin{aligned}
W(\Phi_1, \Phi_2, \Phi_3) = & -(1 - d_1^*)(1 - d_2^*)(1 - d_3^*)[\Phi'_1 + \Phi'_2 + \Phi'_3 - \beta_R W(\Phi'_1, \Phi'_2, \Phi'_3)] \\
& - (1 - d_1^*)d_2^*(1 - d_3^*)[\Phi'_1 + \Phi'_3 - \beta_R W(\Phi'_1, 0, \Phi'_3)] \\
& - (1 - d_1^*)(1 - d_2^*)d_3^*[\Phi'_1 + \Phi'_2 - \beta_R W(\Phi'_1, \Phi'_2, 0)] \\
& - (1 - d_1^*)d_2^*d_3^*[\Phi'_1 - \beta_R W(\Phi'_1, 0, 0)] \\
& - d_1^*(1 - d_2^*)(1 - d_3^*)[\Phi'_2 + \Phi'_3 - \beta_R W(0, \Phi'_2, \Phi'_3)] \\
& - d_1^*d_2^*(1 - d_3^*)[\Phi'_3 - \beta_R W(0, 0, \Phi'_3)] \\
& - d_1^*(1 - d_2^*)d_3^*[\Phi'_2 - \beta_R W(0, \Phi'_2, 0)] \\
& + \beta_R d_1^*d_2^*d_3^*W(0, 0, 0) - \frac{\kappa}{2}(d_1^* + d_2^* + d_3^*)^2.
\end{aligned} \tag{A.38}$$

The manipulation choices $\{m_2^*, m_3^*\}$ and the detection choices $\{d_2^*, d_3^*\}$ at firms 2 and 3 can be analogously derived and given in Appendix A.6.

The proof of this proposition is provided in Appendix A.4.1 below.

A.4.1 Proof of Proposition 4

Proof. of Proposition 4: We only derive the manipulation decision by manager 1 as the manipulation decisions by the other managers can be derived analogously.

Taking the first-order condition of m_{1t} gives that:

$$\begin{aligned}
& c(1-q)d_{1t}^* + (1-q)\mu \\
&= \frac{1-d_{1t}^*}{1-\delta\rho} \frac{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2}{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2 + m_{1t}^*q(1-q)} (1-q) \\
&+ \frac{(1-d_{1t}^*)\beta_F\rho}{1-\delta\rho} \frac{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2}{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2 + m_{1t}^*q(1-q)} (1-q) \\
&\times E_{\Phi_{2t}, \Phi_{3t}} \left[(1-d_{1t+1}^*) \frac{m_{1t+1}^*q(1-q)}{\rho^2\Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^*q(1-q)} \right] \\
&+ \frac{(1-d_{1t}^*)\beta_F^2\rho^2}{1-\delta\rho} \frac{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2}{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2 + m_{1t}^*q(1-q)} (1-q) \\
&\times E_{\Phi_{2t}, \Phi_{3t}, \Phi_{2t+1}, \Phi_{3t+1}} \\
&\times \left[(1-d_{1t+1}^*) (1-d_{1t+2}^*) \frac{m_{1t+2}^*q(1-q)}{\rho^2\Phi_{1t+1} + \sigma_\varepsilon^2 + m_{1t+2}^*q(1-q)} \frac{m_{1t+1}^*q(1-q)}{\rho^2\Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^*q(1-q)} \right] \\
&+ \dots
\end{aligned} \tag{A.39}$$

Note that we need to take expectations over $\{\Phi_{2t}, \Phi_{3t}\}$ because $\{d_{1t+1}^*, m_{1t+1}^*\}$ depend on $\{\Phi_{2t}, \Phi_{3t}\}$. Φ_{2t} and Φ_{3t} are random because they can be either 0 or positive, depending on whether the regulator detects fraud at the two firms.

Equation (A.39) can be simplified into

$$\begin{aligned}
& \frac{(1-\delta\rho)(cd_{1t}^* + \mu)}{1-d_{1t}^*} \left(1 + \frac{m_{1t}^*q(1-q)}{\rho^2\Phi_{1t-1} + \sigma_\varepsilon^2} \right) \\
&= 1 + \beta_F\rho E_{\Phi_{2t}, \Phi_{3t}} \left[(1-d_{1t+1}^*) \frac{m_{1t+1}^*q(1-q)}{\rho^2\Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^*q(1-q)} \right] \\
&+ \beta_F^2\rho^2 E_{\Phi_{2t}, \Phi_{3t}, \Phi_{2t+1}, \Phi_{3t+1}} \\
&\times \left[(1-d_{1t+1}^*) (1-d_{1t+2}^*) \frac{m_{1t+2}^*q(1-q)}{\rho^2\Phi_{1t+1} + \sigma_\varepsilon^2 + m_{1t+2}^*q(1-q)} \frac{m_{1t+1}^*q(1-q)}{\rho^2\Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^*q(1-q)} \right] \\
&+ \dots
\end{aligned} \tag{A.40}$$

There are four possible cases of $\{\Phi_{2t}, \Phi_{3t}\}$ in period $t+1$. For each realization of $\{\Phi_{2t}, \Phi_{3t}\}$,

by induction, the first-order condition of m_{1t+1} is given by:

$$\begin{aligned} & \frac{(1 - \delta\rho) (cd_{1t+1}^* + \mu)}{1 - d_{1t+1}^*} \left(1 + \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2} \right) \\ &= 1 + \beta_F \rho E_{\Phi_{2t+1}, \Phi_{3t+1}} \left[(1 - d_{1t+2}^*) \frac{m_{1t+2}^* q (1 - q)}{\rho^2 \Phi_{1t+1} + \sigma_\varepsilon^2 + m_{1t+2}^* q (1 - q)} \right] + \dots \end{aligned} \quad (\text{A.41})$$

Multiplying both sides by $(1 - d_{1t+1}^*) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^* q (1 - q)}$ gives that:

$$\begin{aligned} & (1 - \delta\rho) (cd_{1t+1}^* + \mu) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2} \\ &= (1 - d_{1t+1}^*) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^* q (1 - q)} \\ &+ \beta_F \rho E_{\Phi_{2t+1}, \Phi_{3t+1}} \\ &\times \left[(1 - d_{1t+1}^*) (1 - d_{1t+2}^*) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^* q (1 - q)} \frac{m_{1t+2}^* q (1 - q)}{\rho^2 \Phi_{1t+1} + \sigma_\varepsilon^2 + m_{1t+2}^* q (1 - q)} \right] \\ &+ \dots \end{aligned} \quad (\text{A.42})$$

Taking the expectation over $\{\Phi_{2t}, \Phi_{3t}\}$ gives that:

$$\begin{aligned} & \frac{(1 - \delta\rho) q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2} E_{\Phi_{2t}, \Phi_{3t}} [(cd_{1t+1}^* + \mu) m_{1t+1}^*] \\ &= E_{\Phi_{2t}, \Phi_{3t}} \left[(1 - d_{1t+1}^*) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^* q (1 - q)} \right] \\ &+ \beta_F \rho E_{\Phi_{2t}, \Phi_{3t}, \Phi_{2t+1}, \Phi_{3t+1}} \\ &\times \left[(1 - d_{1t+1}^*) (1 - d_{1t+2}^*) \frac{m_{1t+1}^* q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2 + m_{1t+1}^* q (1 - q)} \frac{m_{1t+2}^* q (1 - q)}{\rho^2 \Phi_{1t+1} + \sigma_\varepsilon^2 + m_{1t+2}^* q (1 - q)} \right] \\ &+ \dots \end{aligned} \quad (\text{A.43})$$

Substituting (A.43) into (A.40) yields:

$$\frac{(1 - \delta\rho) (cd_{1t}^* + \mu)}{1 - d_{1t}^*} \left(1 + \frac{m_{1t}^* q (1 - q)}{\rho^2 \Phi_{1t-1} + \sigma_\varepsilon^2} \right) = 1 + \frac{\beta_F \rho (1 - \delta\rho) q (1 - q)}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2} E_{\Phi_{2t}, \Phi_{3t}} [(cd_{1t+1}^* + \mu) m_{1t+1}^*]. \quad (\text{A.44})$$

Solving for m_{1t}^* gives that

$$m_{1t}^* = \frac{\rho^2 \Phi_{1t-1} + \sigma_\varepsilon^2}{q(1-q)} \left[\frac{1 - d_{1t}^*}{cd_{1t}^* + \mu} \left(\frac{1}{1 - \delta\rho} + \beta_F \rho q (1 - q) \frac{E_{\Phi_{2t}, \Phi_{3t}} [m_{1t+1}^* (cd_{1t+1}^* + \mu)]}{\rho^2 \Phi_{1t} + \sigma_\varepsilon^2} \right) - 1 \right], \quad (\text{A.45})$$

where

$$\begin{aligned} E_{\Phi_{2t}, \Phi_{3t}} [(cd_{1t+1}^* + \mu) m_{1t+1}^*] &= (1 - d_{2t}^*) (1 - d_{3t}^*) m_{1t+1}^* (\Phi_{1t}, \Phi_{2t}, \Phi_{3t}) (cd_{1t+1}^* (\Phi_{1t}, \Phi_{2t}, \Phi_{3t}) + \mu) \\ &\quad + d_{2t}^* (1 - d_{3t}^*) m_{1t+1}^* (\Phi_{1t}, 0, \Phi_{3t}) (cd_{1t+1}^* (\Phi_{1t}, 0, \Phi_{3t}) + \mu) \\ &\quad + (1 - d_{2t}^*) d_{3t}^* m_{1t+1}^* (\Phi_{1t}, \Phi_{2t}, 0) (cd_{1t+1}^* (\Phi_{1t}, \Phi_{2t}, 0) + \mu) \\ &\quad + d_{2t}^* d_{3t}^* m_{1t+1}^* (\Phi_{1t}, 0, 0) (cd_{1t+1}^* (\Phi_{1t}, 0, 0) + \mu). \end{aligned} \quad (\text{A.46})$$

Analogously, dropping the time subscript, the manipulation choice m_i^* by the manager at firm i can be derived as:

$$m_i^* = \frac{\rho^2 \Phi_i + \sigma_\varepsilon^2}{q(1-q)} \left(\frac{1 - d_i^*}{cd_i^* + \mu} \left(\frac{1}{1 - \delta\rho} + \frac{\beta_F \rho q (1 - q)}{\rho^2 \Phi_i' + \sigma_\varepsilon^2} \times E[m_i' (cd_i' + \mu)] \right) - 1 \right), \quad (\text{A.47})$$

where

$$\begin{aligned} E[m_1' (cd_1' + \mu)] &= (1 - d_2^*) (1 - d_3^*) m_1^* (\Phi_1', \Phi_2', \Phi_3') (cd_1^* (\Phi_1', \Phi_2', \Phi_3') + \mu) \\ &\quad + d_2^* (1 - d_3^*) m_1^* (\Phi_1', 0, \Phi_3') (cd_1^* (\Phi_1', 0, \Phi_3') + \mu) \\ &\quad + (1 - d_2^*) d_3^* m_1^* (\Phi_1', \Phi_2', 0) (cd_1^* (\Phi_1', \Phi_2', 0) + \mu) \\ &\quad + d_2^* d_3^* m_1^* (\Phi_1', 0, 0) (cd_1^* (\Phi_1', 0, 0) + \mu), \end{aligned} \quad (\text{A.48})$$

$$\begin{aligned}
E[m'_2(cd'_2 + \mu)] &= (1 - d_1^*)(1 - d_3^*)m_2^*(\Phi'_1, \Phi'_2, \Phi'_3)(cd_2^*(\Phi'_1, \Phi'_2, \Phi'_3) + \mu) \\
&\quad + d_1^*(1 - d_3^*)m_2^*(0, \Phi'_2, \Phi'_3)(cd_2^*(0, \Phi'_2, \Phi'_3) + \mu) \\
&\quad + (1 - d_1^*)d_3^*m_2^*(\Phi'_1, \Phi'_2, 0)(cd_2^*(\Phi'_1, \Phi'_2, 0) + \mu) \\
&\quad + d_1^*d_3^*m_2^*(0, \Phi'_2, 0)(cd_2^*(0, \Phi'_2, 0) + \mu), \tag{A.49}
\end{aligned}$$

$$\begin{aligned}
E[m'_3(cd'_3 + \mu)] &= (1 - d_1^*)(1 - d_2^*)m_3^*(\Phi'_1, \Phi'_2, \Phi'_3)(cd_3^*(\Phi'_1, \Phi'_2, \Phi'_3) + \mu) \\
&\quad + d_1^*(1 - d_2^*)m_3^*(0, \Phi'_2, \Phi'_3)(cd_3^*(0, \Phi'_2, \Phi'_3) + \mu) \\
&\quad + (1 - d_1^*)d_2^*m_3^*(\Phi'_1, 0, \Phi'_3)(cd_3^*(\Phi'_1, 0, \Phi'_3) + \mu) \\
&\quad + d_1^*d_2^*m_3^*(0, 0, \Phi'_3)(cd_3^*(0, 0, \Phi'_3) + \mu). \tag{A.50}
\end{aligned}$$

Dropping the time subscript, the regulator's objective function can be rewritten recursively as:

$$\begin{aligned}
W(\Phi_1, \Phi_2, \Phi_3) &= \max_{d_1, d_2, d_3} - (1 - d_1)(1 - d_2)(1 - d_3)[\Phi'_1 + \Phi'_2 + \Phi'_3 - \beta_R W(\Phi'_1, \Phi'_2, \Phi'_3)] \\
&\quad - (1 - d_1)d_2(1 - d_3)[\Phi'_1 + \Phi'_3 - \beta_R W(\Phi'_1, 0, \Phi'_3)] \\
&\quad - (1 - d_1)(1 - d_2)d_3[\Phi'_1 + \Phi'_2 - \beta_R W(\Phi'_1, \Phi'_2, 0)] \\
&\quad - (1 - d_1)d_2d_3[\Phi'_1 - \beta_R W(\Phi'_1, 0, 0)] \\
&\quad - d_1(1 - d_2)(1 - d_3)[\Phi'_2 + \Phi'_3 - \beta_R W(0, \Phi'_2, \Phi'_3)] \\
&\quad - d_1d_2(1 - d_3)[\Phi'_3 - \beta_R W(0, 0, \Phi'_3)] \\
&\quad - d_1(1 - d_2)d_3[\Phi'_2 - \beta_R W(0, \Phi'_2, 0)] \\
&\quad + \beta_R d_1 d_2 d_3 W(0, 0, 0) - \frac{\kappa}{2}(d_1 + d_2 + d_3)^2, \tag{A.51}
\end{aligned}$$

where

$$\Phi'_i \equiv \frac{m_i^* q (1 - q) (\rho^2 \Phi_i + \sigma_\varepsilon^2)}{\rho^2 \Phi_i + \sigma_\varepsilon^2 + m_i^* q (1 - q)}. \tag{A.52}$$

Taking the first-order condition yields:

$$\begin{aligned}
d_1^* = & \frac{1}{\kappa} \{ \Phi'_1 + \beta_R (1 - d_2^*) (1 - d_3^*) [W(0, \Phi'_2, \Phi'_3) - W(\Phi'_1, \Phi'_2, \Phi'_3)] \\
& + \beta_R (1 - d_2^*) d_3^* [W(0, \Phi'_2, 0) - W(\Phi'_1, \Phi'_2, 0)] \\
& + \beta_R d_2^* (1 - d_3^*) [W(0, 0, \Phi'_3) - W(\Phi'_1, 0, \Phi'_3)] \\
& + \beta_R d_2^* d_3^* [W(0, 0, 0) - W(\Phi'_1, 0, 0)] \} - (d_2^* + d_3^*), \tag{A.53}
\end{aligned}$$

$$\begin{aligned}
d_2^* = & \frac{1}{\kappa} \{ \Phi'_2 + \beta_R (1 - d_1^*) (1 - d_3^*) [W(\Phi'_1, 0, \Phi'_3) - W(\Phi'_1, \Phi'_2, \Phi'_3)] \\
& + \beta_R (1 - d_1^*) d_3^* [W(\Phi'_1, 0, 0) - W(\Phi'_1, \Phi'_2, 0)] \\
& + \beta_R d_1^* (1 - d_3^*) [W(0, 0, \Phi'_3) - W(0, \Phi'_2, \Phi'_3)] \\
& + \beta_R d_1^* d_3^* [W(0, 0, 0) - W(0, \Phi'_2, 0)] \} - (d_1^* + d_3^*), \tag{A.54}
\end{aligned}$$

$$\begin{aligned}
d_3^* = & \frac{1}{\kappa} \{ \Phi'_3 + \beta_R (1 - d_1^*) (1 - d_2^*) [W(\Phi'_1, \Phi'_2, 0) - W(\Phi'_1, \Phi'_2, \Phi'_3)] \\
& + \beta_R (1 - d_1^*) d_2^* [W(\Phi'_1, 0, 0) - W(\Phi'_1, 0, \Phi'_3)] \\
& + \beta_R d_1^* (1 - d_2^*) [W(0, \Phi'_2, 0) - W(0, \Phi'_2, \Phi'_3)] \\
& + \beta_R d_1^* d_2^* [W(0, 0, 0) - W(0, 0, \Phi'_3)] \} - (d_1^* + d_2^*). \tag{A.55}
\end{aligned}$$

■

About ECGI

ECGI is an international non-profit membership organisation that provides a platform for debate and dialogue among academics, policymakers, and business leaders, with a focus on major corporate governance, ESG, and stewardship issues.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

ECGI Working Paper Series in Law

Editorial Board

Editor	Amir Licht , Professor of Law, Radzyner Law School, Interdisciplinary Center Herzliya
Consulting Editors	Hse-Yu Iris Chiu , Professor of Corporate Law and Financial Regulation, University College London Martin Gelter , Professor of Law, Fordham University School of Law Geneviève Helleringer , Professor of Law, ESSEC Business School and Oxford Law Faculty Kathryn Judge , Professor of Law, Columbia Law School Wolf-Georg Ringe , Professor of Law & Finance, University of Hamburg
Editorial Assistant	Asif Malik , ECGI

<https://www.ecgi.global/publications/working-papers>