MANAGEMENT SCIENCE
Vol. 56, No. 6, June 2010, pp. 924–937
ISSN 0025-1909 | EISSN 1526-5501 | 10 | 5606 | 0924

informs

DOI 10.1287/mnsc.1100.1166

The Auditor's Slippery Slope: An Analysis of Reputational Incentives

Carlos Corona
McCombs School of Business, The University of Texas at Austin, Austin, Texas 78712, carlos.corona@mccombs.utexas.edu

Ramandeep S. Randhawa
Marshall School of Business, University of Southern California, Los Angeles, California 90089,
ramandeep.randhawa@marshall.usc.edu

Reputational concerns have commonly been perceived to have a positive effect on auditing firms' execution of their monitoring and attesting functions. This paper demonstrates that this need not always be the case by studying a two-period game of repeated interaction between a manager and an auditor under the assessment of the market for audit services. Regarding reputation as the sole motivator for the auditor, we illustrate how reputational concerns induce an auditing firm to misreport. We investigate the reasons and circumstances under which such misreporting takes place. In particular, a strategic manager can induce the audit firm down a slippery slope, wherein the managerial fraud increases as the tenure of the audit firm progresses, whereas the auditor's fraud reporting probability decreases.

Key words: reputation; auditing; game theory; sequential equilibrium
History: Received May 23, 2008; accepted January 26, 2010, by Stefan Reichelstein, accounting. Published online in *Articles in Advance* April 23, 2010.

1. Introduction

In the late 1990s/early 2000s, an outburst of financial scandals (including those of Enron, WorldCom, and other major corporations), all involving the presence of accounting irregularities, disrupted the confidence investors placed in accounting institutions. The causes deemed behind this sudden concentration of misconduct have been multiple and interrelated. Nevertheless, the failure of gatekeepers like auditors and boards of directors seems most troubling, because they are essential elements needed to preserve credibility in the markets (Coffee 2004). The Sarbanes-Oxley Act of 2002 strengthened auditor independence by banning audit firms from providing certain nonaudit services, enforcing audit partner rotation every five years, and requiring an audit committee to supervise the audit function in the firm. Yet, the adequacy of the current audit regulations in providing auditors with incentives to perform their role is still the subject of debate (Kinney 2005). An implicit conjecture in this debate is that reputation concerns already supply auditors with some of those incentives. This conjecture is particularly relevant because audit fees are restricted to be noncontingent by law, leaving reputation and legal liability as the only alternative motivating levers.[1]

In this paper, we argue that reputation concerns alone cannot be relied upon to provide an adequate motivation for auditors. We begin by considering a single interaction between a manager and an auditor, and we show that in this setting reputation concerns provide the auditor with the right incentives to report detected fraud. However, in a repeated interaction with the same manager, reputation concerns may in fact lead the auditor into a "slippery-slope" situation. In particular, if fraud is detected only in a later period, a fraud report at this stage may highlight the auditor's failure to detect previous fraud, and thus damage the auditor's reputation. The presence of a strategic manager can, in some circumstances, lead to an extreme form of this effect that we call "testing the waters." Here, the manager's actions fully reveal any private information about the auditor's type that the manager may have acquired in previous interactions. Therefore, the auditor is even more reluctant to report fraud because this reveals unfavorable information about his type. A well-known real-life

[1] When we refer to audit fees as being noncontingent we mean that the law forbids audit fees to be explicitly contingent on the outcome of the audit. However, because audits are a mechanism through which firms increase the credibility of their financial statements, the market's perception of the auditor's ability is indeed an asset that both firms and investors value. Accordingly, in this paper we assume that current audit fees are implicitly contingent on past audit reports because these reports change the market's perception of the auditor's ability. This contingency, however, is not contractually explicit and does not include the outcome of the current audit.

example is the case of ESM Government Securities, where an auditor knowingly concealed fraud to protect his reputation (see Maggin 1989). When the auditor discovered fraudulent activity and realized that it had been occurring for the previous two years, the manager made the following comment to the auditor: "How's it going to look? It's going to look terrible for you, and you just got promoted to partner . . . ." The slippery-slope phenomenon in general results in a cycle of intertemporally increasing likelihood of fraud and decreasing likelihood of fraud reports. We provide insights as to why, and discuss the circumstances in which, reputation concerns may or may not be adequate incentives for auditors to perform their functions properly.

The model examines a two-period interaction between a manager and an auditor. In each period, fees for audit services are determined competitively by the market. These fees are based solely on the market's perception of the auditor's fraud detection capability or, in other words, the auditor's reputation. The auditor can be one of two types—high ability or low ability. The high-ability auditor detects fraud with a strictly greater probability than the low-ability auditor. We assume that the auditor shares the public prior belief about his ability. The manager can also be one of two types—honest or dishonest. The honest manager never commits fraud, whereas the dishonest manager rationally decides whether or not to commit fraud. In each period, the manager first decides whether to commit fraud, then the auditor performs an audit. If fraud is detected, the auditor must decide whether to report it or not.

The repeated interaction between the manager and the auditor leads to two kinds of equilibria depending on the initial manager reputation (of being honest). When the manager's reputation is high, the resulting equilibrium depicts a manager behavior that produces the effect of "handcuffing" the auditor in the second period. In particular, the dishonest manager commits fraud in the first period with some probability. If in the next period the auditor detects fraud and reports it, he reveals that the manager is dishonest and, with that, the likely presence of undetected fraud in the first period. Hence, the auditor has a reduced incentive to report fraud in the second period, providing the manager with the opportunity to safely commit fraud with certainty in the second period.

When the manager's reputation is low, the market expects fraud to be committed and silence has a high cost for the auditor. The auditor is therefore more willing to report detected fraud. This cautions the manager, who now only deems it safe to commit fraud in the second period if he gains some additional "knowledge" that the auditor may be of low ability. Hence, the manager "tests the waters" by first committing fraud with a small probability. If fraud is committed and remains undetected, then the manager commits fraud in the second period with certainty. If, on the other hand, the manager does not commit fraud in the first period, he obtains no additional knowledge on the auditor type, and thus does not commit fraud in the second period. In this scenario, if the auditor detects fraud in the second period, he feels trapped because he is forced to choose between revealing fraud, thereby revealing the certain presence of missed fraud in the first period, and concealing fraud, thereby letting the market raise the suspicion of fraud being undetected in both periods. The resolution of this trade-off depends on the amount of the manager's penalty and the audit technology. For high penalty and high audit technology, the auditor may conceal fraud with some probability, but for all other cases, the auditor always reports any detected fraud. This result may appear counterintuitive in the sense that one would expect high penalty and high audit technology to be an environment in which reputational incentives should work. However, it is precisely the low likelihood of fraud in this environment that allows the auditor to remain silent without harming his reputation. On the other hand, for low penalty, one would expect a higher likelihood that the manager commits fraud, and hence the auditor only hurts his reputation by remaining silent. In the case of low audit technology, the market does not expect the auditor to detect fraud with a high likelihood, and thus if the auditor does detect fraud, reporting it only increases his reputation. Thus, in cases with low penalty or low audit technology, reputational incentives work perfectly and the auditor always reports any detected fraud.

The results thus far do not consider any possible revelation of fraud by anyone other than the auditor; that is, there is no whistle-blower or regulator to detect fraud when the auditor fails to do so. One may expect the presence of such an agent to wash away the inadequacies of reputational incentives. However, we demonstrate that as long as the probability of revelation of fraud by this agent is below a threshold, reputational incentives remain imperfect, and our results continue to hold. Furthermore, we find that this threshold can be close to one in some cases.

The model in this paper focuses on reputation as the sole motivator for the auditor and does not consider other disciplinary mechanisms. Interestingly, we find that reputation incentives can work perfectly in some scenarios even in the absence of the disciplining mechanisms. However, we do find cases in which reputational concerns do not suffice and lead the auditor down a slippery slope. It is precisely in these scenarios that additional disciplining mechanisms are needed to ensure that the auditor performs his job properly. Thus, our model sheds light on situations in

which additional measures may be called for by the regulatory authorities.

This paper is organized as follows. In the next section, we compare this paper with the extant audit literature. In §3, we describe the model setup, and in §4, we characterize the equilibria. In §5, we discuss the equilibria and performs comparative statics and a robustness study incorporating whistle-blowing. We conclude in §6. All proofs are contained in the online appendix (provided in the e-companion).[2]

2. Comparison with Related Literature

In the existing audit analytical literature, the auditor's reticence to perform his responsibility stems either from the associated cost that such a task entails (see, e.g., Simunic 1980, Morton 1993, Newman et al. 2001, Beyer and Sridhar 2006) or from some form of collusion with the manager.[3] The auditor is then motivated to perform his function by either a contingent contract with the firm,[4] the threat of a legal liability,[5] or a concern for his reputation. That is, reputation has been regarded as a motivator, with more or less relevance, but invariably with a positive effect on auditors' execution of their monitoring and attesting functions.

The positive role of reputation has been noted in extant models of auditor behavior. These models regard reputation either as the public perception of an exogenous auditor feature (DeAngelo 1981, Titman and Trueman 1986, Dye 1993) or as the manager's perception of the auditor's ability (Datar and Alles 1999). In contrast, this paper streamlines the audit context to focus on a comprehensive depiction of reputation formation. The manager, the market and the auditor all have potentially different information sets, and therefore different perceptions of the auditor's ability. Moreover, the three players also have different perceptions of the manager's honesty. The dynamic interaction among these individual beliefs is the main reason why reputation concerns might induce an auditor to misreport.

Although Datar and Alles (1999) also model reputation formation in a dynamic setting, they focus only on the manager's perception of the auditor's ability. This constrains the role of reputation as a deterrent for the manager. In our paper, we incorporate the market's perception of the auditor's ability, thereby making the auditor's salary contingent on this perception. This introduces a new dimension to the problem that may cause reputation incentives to induce misreporting. For instance, when the manager has a worse perception of the auditor's ability than does the market, the manager's actions, if reported, may reveal this perception to the market. Thus, the auditor may decide to conceal these actions.

Illustrated in a different context, some of the informational effects in this paper are reminiscent of those of Kanodia et al. (1989). Their model gives a rational explanation for the documented "sunk cost effect." Specifically, a manager might escalate investment in a project to avoid hurting his reputation by admitting he made a mistake when he initially undertook the project. In contrast, the related slippery-slope effect illustrated in this paper is an endogenous situation in which the auditor's behavior is induced by a dishonest manager. In this case, the auditor's reluctance to report a detected fraud results from strategic decisions adopted previously by the manager. This allows us to characterize situations in which the manager–auditor interaction would lead to a slippery-slope situation.

Our results are also related to the supply chain quality literature (see Baiman et al. 2000, Balachandran and Radhakrishnan 2005) and the audit quality literature (Sarath 1991, Balachandran and Nagarajan 1987), which demonstrate that as the quality of a supplier/auditor increases, the amount of penalty required also increases. This is analogous to our main result that reputation works only when the audit quality is low.

3. Model

We examine a game with two players: a manager and an auditor. The manager and the auditor repeatedly interact over the course of the auditor's tenure. In addition, there is a market for audit services (henceforth, *the market*) that determines audit fees as a function of the public perception of the auditor's ability (or reputation). The game begins with a firm hiring a manager who can be one of two types: *honest* or *dishonest*. The honest manager never commits any fraud, whereas the *dishonest* manager is a risk-neutral agent who decides rationally whether to commit fraud or not to maximize his expected payoff. Even though only the manager knows his type, there is a public prior belief that the manager is honest with probability $y_0 \in (0, 1)$.

As there is a chance that the manager is dishonest, the firm hires an auditor to serve as a deterrent, and

[2] An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

[3] The possibility of a collusive agreement and its impact on the design of contracts in an agency setting has been examined extensively (e.g, Baiman et al. 1991, Kofman and Lawarree 1993, Lee and Gu 1998).

[4] Baiman et al. (1987) and Dye et al. (1990) are examples wherein contingent audit fees have been modeled in a one-stage contracting setting.

[5] Auditor legal liability has been extensively examined as an auditor motivator (see, e.g., Simunic 1980; Dye 1993; Narayanan 1994; Schwartz 1997, 1998; Radhakrishnan 1999; Laux and Newman 2010).

as a detector of fraud. There are two types of auditors: *high ability* and *low ability*. The high-ability auditor (referred to as type H) detects fraud, when committed, with a fixed probability γ, whereas the low-ability auditor (referred to as type L) detects fraud with a strictly smaller probability δ. We assume that both types detect fraud with a positive probability, but are not perfect, i.e., $0 < \delta < \gamma < 1$. Moreover, regardless of type, the auditor is a rational player with no inherent preference for honest reporting.[6]

In this paper, we define the auditor's reputation at any given time as the market's belief that the auditor is of type H, i.e., the probability that the auditor is of type H conditional on the public information at that time.[7] The auditor type is unobservable, even by the auditor himself.[8] However, there is an initial public prior on the auditor's type. We denote this initial reputation as x_0 and assume that $0 < x_0 < 1$. When an auditor reports fraud, it becomes public,[9] and thus allows a market update on the initial prior. Because market beliefs determine future audit fees, when the auditor detects fraud, he strategically decides whether to reveal this information or not. We further assume that the auditor never detects fraud in its absence and that if he does not detect fraud, he cannot report it (i.e., he cannot falsely accuse the manager).

In the spirit of focusing on the manager–auditor interaction, we assume that the firm that hires the manager and the auditor serves essentially as a "venue" for their interaction. The only actions taken by the firm are those of firing and replacing the manager upon a negative report by the auditor and that of paying audit fees at the market rate. Specifically, we assume that the market for audit services sets the audit fees competitively at a level that is monotonically contingent on the auditor's reputation given the publicly available information. Audit fees are paid at the beginning of each period and are given by an increasing function $w(x)$ of the auditor's reputation, x, at that time. Thus, the auditor's payoff in period t is only contingent on his reputation at the end of the previous period, x_{t-1}. We assume that the auditor ends his engagement with the current firm at the end of the second period, and his payoffs from future engagements are summarized by a terminal payoff function $\tau(\cdot)$, which is an increasing function of his final reputation.

3.1. Baseline Single-Period Model

With the intention of establishing a baseline result, we first examine a one-period setting (Figure 1 displays the timeline of events). The game begins with the dishonest manager deciding whether to commit fraud or not—that is, selecting the action $m_1 \in \{f, nf\}$, where f denotes fraud and nf denotes no fraud (note that if the manager is honest, fraud is never committed and is thus never reported). The auditor then performs an audit, which may detect fraud (if committed) with some probability. If fraud is detected, the auditor then decides whether to report it or not; that is, he selects the action $a_1 \in \{r, nr\}$, where r denotes report and nr denotes no report. Note that if fraud is not detected, it cannot be reported. Let μ and α denote the probability of committing fraud by the manager and reporting fraud upon detection by the auditor, respectively. That is, these are the players' strategies.

Turning to payoffs, the auditor is paid an initial wage of $w(x_0)$ at the beginning of the game and has a terminal payoff of $\tau(x_1)$, where x_1 is the reputation obtained as a result of a Bayesian update on the auditor's report. In particular, $x_1 \in \{x_r, x_{nr}\}$, where x_r is the reputation based on a report of fraud, whereas x_{nr} is that based on no report. The auditor's decision process is then simply reduced to maximizing his terminal payoff $\tau(x_1)$ by comparing x_r with x_{nr}. Turning to the manager's incentives, because our focus is the manager's decision to commit fraud, we normalize the manager's base payoff (when fraud is not committed) to zero. If fraud is committed and not reported, the manager obtains a payoff of F. However, if fraud is reported, the manager suffers a penalty P in addition to losing the amount of fraud F. For simplicity, we model all termination concerns, including any loss of current and future salaries, as a part of the penalty.

Given the auditor's initial reputation x_0 and an inferred auditor's reporting strategy $\hat{\alpha}$, the manager

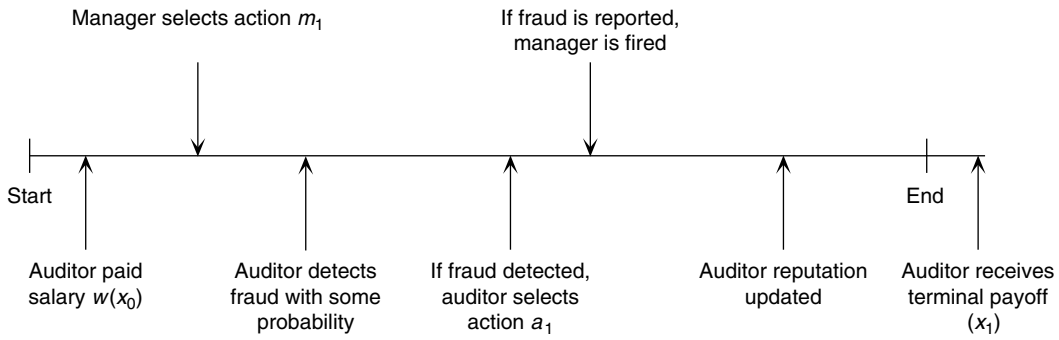
[6] Alternatively, we could consider a model in which all auditors are honest and report truthfully but must exert effort (which may be costly) to detect fraud. So, an auditor first decides to apply effort or not. Then, if effort is applied and fraud was committed, the auditor may detect fraud, in which case he always reports it. If in this alternative model effort is costless, then it can be translated to our original model so that all insights hold. If effort is costly, then the auditors' incentives not to do his job properly increase as expected, and hence fraud reporting decreases.

[7] One may wish to include the auditor's honesty as part of his reputation. However, this inclusion renders the auditor reputation two-dimensional, one on ability or probability of detection, and the other on honesty. This case is beyond the scope of this paper and left as a possibility for future work.

[8] This assumption keeps the model parsimonious. Relaxing it would increase the dimension of the game by increasing the number of information states. This increases the model's complexity without yielding additional insights.

[9] In reality, the auditor is required (as per §10A of the Securities Exchange Act of 1934) to report any detected material misstatement due to fraud to the audit committee, and if the audit committee does not take remedial action, then the auditor is required to report it to the U.S. Securities and Exchange Commission (SEC). In either case, the subsequent remedial actions (by the board of directors or the SEC) and restatement are public, and therefore are informative for updating the auditor's reputation. In this paper, we simplify reality by assuming that fraudulent activity reported by the auditor becomes public knowledge immediately.

Figure 1 Timeline of Events in Period 1



expects fraud to be reported with probability $(x_0\gamma + (1-x_0)\delta)\hat{\alpha}$. The manager's expected payoff, $\Pi_1(x_0, \hat{\alpha})$, is then the expected benefit from committing fraud minus the expected penalty:

$$\Pi_1(x_0, \hat{\alpha}) = F(1-(x_0\gamma+(1-x_0)\delta)\hat{\alpha}) - P(x_0\gamma+(1-x_0)\delta)\hat{\alpha}.$$

As the auditor and the manager interact only once, the equilibrium is straightforward. Reporting fraud whenever it is detected is a dominant strategy for the auditor. To see this, note that for any nonzero probability of fraud, the auditor's reputation after reporting fraud, x_r, is always strictly higher than his reputation after not reporting fraud, x_{nr}. The dishonest manager will commit fraud only when his expected payoff, $\Pi_1(x_0, \hat{\alpha})$, is positive. This happens only when committed fraud has a small enough probability of being detected and reported. Given that the auditor always reports any detected fraud, the manager commits fraud if and only if the probability of detection is below a threshold:

$$x_0\gamma + (1-x_0)\delta < \frac{F}{F+P}. \tag{1}$$

The following result is then immediately obtained:

Lemma 1. *In a one-period setting there is a unique equilibrium in which the auditor always reports any detected fraud and the manager commits fraud if and only if the probability of detection is below a threshold, i.e.,* (1) *holds.*

The equilibrium of the one-period setting supports the prevalent opinion that reputation concerns provide auditors with incentives to do their job properly. For instance, Datar and Alles (1999) contend that reputation formation by the auditor serves as a substitute for costly contracting, monitoring, and litigation by the owners of the firm. Indeed, if the effect of reputation incentives on auditor's decisions were fairly conveyed by the one-period setting just described, other additional incentive mechanisms would be redundant.

In the rest of the paper, we will show that when the interaction between a manager and an auditor extends beyond one period, the incentives provided by reputation are inadequate to ensure that the auditor will always choose to perform his functions correctly. It is worth noting that the manager's decisions depend on the size of the penalty relative to the fraud, P/F. Consequently, we henceforth normalize the payoff from fraud to one, i.e., $F = 1$. Furthermore, to focus on the impact of additional periods on the manager–auditor interaction, we make the following assumption:

Assumption 1. *The auditor's initial reputation x_0 is high enough to deter fraud in a one-period setting when the auditor always reports fraud, i.e., $\Pi_1(x_0, 1) < 0$. This is equivalent to assuming that the penalty is lower bounded as follows*:

$$P > P_l \equiv \frac{1}{\delta + (\gamma - \delta)x_0} - 1. \tag{2}$$

This assumption serves two purposes. First, it limits the scope of the model to the range of parameters in which it is most relevant. Extending the parameter range to allow for fraud in the one-period setting results in equilibria in the two-period model in which the manager always commits fraud in all circumstances. Second, it serves as a baseline for the existence of fraud in the two-period model. That is, it allows attributing the existence of fraud in the two-period model to the repeated interaction between the auditor and the manager.

3.2. The Two-Period Model

We consider a two-period game between the auditor and the dishonest manager. In each period t, the manager maximizes his expected payoff by selecting an action $m_t \in \{f, nf\}$, where f denotes committing fraud and nf denotes inaction or not committing fraud. We let I^m denote the information available to the manager at the beginning of the second period. In particular, $I^m \in \{k, nk\}$, where k denotes knowledge and corresponds to the information state in which the manager committed fraud in the first period but it was

not reported; in the state nk, the manager has no additional information at the beginning of period 2 because he did not commit fraud in period 1. The distinction between these cases is essential, because in the two information states the manager updates his beliefs on the auditor's ability differently. The manager's behavioral strategy is then given by $\mu = (\mu_1, \mu_{2,k}, \mu_{2,nk})$, where

1. μ_1 is the probability with which the dishonest manager commits fraud in period 1; that is, $\mu_1 = \mathbb{P}(m_1 = f \mid \text{Manager is dishonest})$.

2. $\mu_{2,k}$ is the probability with which the dishonest manager commits fraud in period 2 when he committed fraud in period 1 and was not reported. That is, $\mu_{2,k} = \mathbb{P}(m_2 = f \mid I^m = k, \text{Manager is dishonest})$. In this case, the fact that the fraud was not reported provides the manager some additional private information about the auditor's type (i.e., a Bayesian update on the auditor's reputation that is different from the market's).

3. $\mu_{2,nk}$ is the probability with which the manager commits fraud in period 2 when he did not commit fraud in period 1, and hence has no additional information about the auditor. That is, $\mu_{2,nk} = \mathbb{P}(m_2 = f \mid I^m = nk, \text{Manager is dishonest})$.

In the case that fraud committed by the manager is reported by the auditor in period 1, the manager incurs the penalty P and is fired. For convenience, we assume that the game ends at this point.[10]

We now turn our attention to the auditor. The auditor chooses the action of reporting fraud or not only when he detects fraud. We denote the action in period t by $a_t \in \{r, nr\}$, where r denotes reporting the fraud, and nr denotes not reporting. In the single-period setting, reporting detected fraud is a dominant strategy for the auditor. In a two-period setting, however, if the auditor does not report fraud in period 1 and detects fraud in period 2, his reporting decision affects the market's perception of a missed fraud in the first period. Hence, the auditor may wish to conceal the detected fraud in period 2. To focus on this effect of reputational incentives, we simplify the model by assuming that, consistent with the one-period setting, the auditor always reports detected fraud in period 1.[11] This assumption keeps the model parsimonious while allowing for a detailed analysis of the effects of reputational incentives in the auditor's second-period decision. Thus, the auditor's strategy is the probability α that the auditor reports detected fraud in period 2.

The timeline of the various events is displayed in Figure 2. The period 1 timeline is displayed in Figure 2(a). In period 2, the game follows Figure 2(b) if fraud is not reported, otherwise the game ends after period 1.

3.2.1. Payoffs. We denote the manager's *expected payoff* from committing fraud in period 2 as a function of the auditor's reputation and the manager's inference of the auditor's strategy $\hat{\alpha}$ as $\Pi_2(x, \hat{\alpha})$. We can write

$$\Pi_2(x, \hat{\alpha}) = F(1 - (x\gamma + (1-x)\delta)\hat{\alpha}) - P(x\gamma + (1-x)\delta)\hat{\alpha}.$$

Note that although $F = 1$, we retain this notation in these definitions for clarity.

The total *expected payoff* from committing fraud in period 1 can then be written as a function of the auditor's initial reputation, the manager's inference of the auditor's strategy, and the manager's strategy in period 2 as

$$\Pi(x_0, \hat{\alpha}, \mu_{2,k}) = (1-(x_0\gamma+(1-x_0)\delta))(F+\mu_{2,k}\Pi_2(x_k, \hat{\alpha})) - P(x_0\gamma+(1-x_0)\delta),$$

where $x_k = \mathbb{P}$ (Auditor is $H \mid I^m = k$) is the manager's belief that the auditor is of high ability conditioned on the fact that a committed fraud was not detected in period 1 (see Online Appendix EC.1 for the explicit expression for x_k). If, on the other hand, the manager decides not to commit fraud in the first period, he gets a payoff of zero at the end of the first period, and a payoff of $\mu_{2,nk}\Pi_2(x_{nr}, \hat{\alpha})$ at the end of the second period. By comparing these two alternative payoffs, the manager decides whether to commit fraud in the first period or not.

Turning to the auditor, the audit fees are an increasing function of the auditor's reputation. Thus, the auditor decides to report or conceal fraud in period 2 to maximize his terminal reputation; that is, the auditor reports fraud if and only if his reputation after reporting fraud in period 2, $x_{nr,r}$, exceeds that obtained by not reporting, $x_{nr,nr}$ (see Online Appendix EC.1 for detailed updating formulae).

Now that we have set up the model and discussed the payoffs for the two players, we characterize the equilibrium of the game in the following section.

4. Equilibrium Analysis

We begin by discussing some properties of the equilibrium. Defining

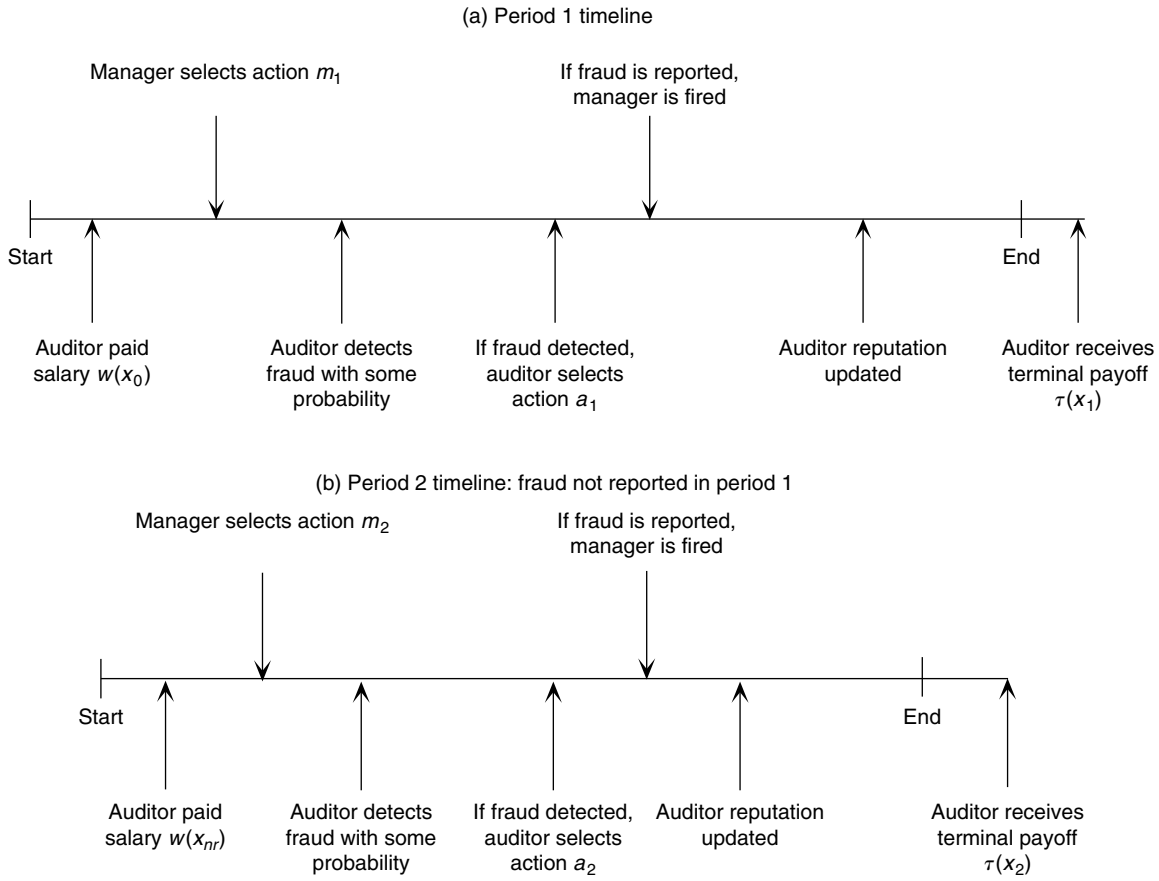
$$P_h \equiv P_l + (1+P_l)^2(1-x_0)x_0(\gamma-\delta)^2, \tag{3}$$

we note that, for penalty levels beyond P_h, the manager never commits fraud.

[10] The firm would in fact replace the manager, and the auditor would then interact with the new manager in a renewed version of the original game, albeit with a higher initial reputation. If the ensuing interaction occurs only for a single period, then, as per the analysis in §3.1, in equilibrium the manager would not commit fraud, whereas the auditor would report any detected fraud.

[11] Reporting detected fraud at the end of the first period leads to a higher reputation than concealing it. Thus, this assumption is equivalent to assuming that auditors with higher reputations should have higher future payoffs.

Figure 2 Timeline of Events



Lemma 2. *For $P > P_h$, in any equilibrium the manager never commits fraud.*

Because the manager never commits fraud for these penalty levels, the auditor's actions are reduced to an out of equilibrium threat, with an inherent dependence on market beliefs. Because our goal is to study the effect of reputational incentives on the manager-auditor interaction, we focus on the range of parameters in which the manager commits fraud with at least a positive probability. That is, henceforth we assume

Assumption 2. *The penalty imposed on the manager when reported by the auditor satisfies $P < P_h$, where P_h is given in* (3).

In the following subsections, we will demonstrate that for each set of parameters there is a unique equilibrium. We divide the equilibrium characterization based on the manager's initial reputation y_0 relative to a threshold $\bar{y} = (1-\gamma)(1-\delta)/\gamma\delta$ that depends on the audit technology. Section 4.1 portrays the equilibrium for the case in which the manager has a high initial reputation for honesty, i.e., $y_0 > \bar{y}$, and §4.2 focuses on the case in which the manager has a low initial reputation for honesty, i.e., $y_0 < \bar{y}$.

4.1. The Reputation-Handcuff Effect: $y_0 > \bar{y}$

The following result characterizes the equilibrium for the case in which the manager has a high initial reputation for honesty:

Proposition 1 (Handcuff Equilibrium: Auditor May Conceal Fraud). *If $y_0 > \bar{y}$ and $P_l < P < P_h$, there is a unique mixed strategy equilibrium given by*

$$\mu^* = \left\{ \frac{2}{\gamma + \delta + \sqrt{(\gamma-\delta)^2 + 4\gamma\delta y_0}}, 1, 1 \right\} \quad \textit{and} \quad (4)$$

$$\alpha^* = \frac{(2+P)(\delta(1-x_0)+\gamma x_0) - 1}{(1+P)(\delta^2(1-x_0)+\gamma^2 x_0)}. \quad (5)$$

This is a mixed strategy equilibrium in which the "mixing" occurs between the manager's first-period fraud decision and the auditor's second-period reporting decision.[12] Notice that this equilibrium extends over the whole range of penalties under consideration,

[12] The reader may note that multiperiod games with mixed strategies do not typically involve mixing across periods. However, in such games every period marks the start of a new subgame. The game in this paper has no subgame: If the auditor uncovers fraud in period 2, he does not know whether the manager committed fraud in the first period or not. It is this lack of information that allows mixing to occur across periods.

and that the manager commits fraud with certainty in period 2 regardless of his action in period 1.

The intuition behind the auditor's reporting decision is as follows. Examine the situation in which the auditor does not detect fraud in period 1 but detects fraud in period 2. By reporting fraud, the auditor also reveals that the manager is dishonest. This information is used by the market to reinterpret the absence of a previous report. If the market believes the manager committed fraud in the first period with high probability, then it interprets the reported fraud as an indication of a likely missed fraud in the past. This leads to a loss in reputation for the auditor. However, if the auditor remains silent when the manager's reputation is high, the market will be inclined to interpret the auditor's silence as a confirmation of the manager's honesty. Thus, the combination of a high manager's reputation and a high frequency of fraud makes the auditor less inclined to report.

The strategic manager is able to exploit the auditor's reluctance to report in the second period. Even though, by Assumption 1, the manager's expected payoff in the first period alone is negative on committing fraud, he still commits fraud with positive probability to set himself up to make full use of the fact that the auditor cannot always report truthfully in period 2. In this sense, the manager essentially "handcuffs" the auditor.

Note that in this case one never obtains complete fraud concealment, i.e., a situation in which the auditor never reports in period 2. The reason for this is that if fraud concealment is certain, then the manager's payoff in the second period from committing fraud becomes very attractive. This leads the manager to hold back fraud in the first period to avoid detection and reap the benefits of "free" fraud later on. But then this cannot be an equilibrium because a less frequent fraud in the first period allows the auditor to deviate and report in the second period. In the equilibrium, thus, the manager must choose a level of fraud that just makes the auditor indifferent between reporting or not.

To summarize, Proposition 1 illustrates that when an auditor remains in the same firm for more than one period, reputation incentives might actually drive the auditor to conceal detected fraud. In fact, the action of reporting fraud itself can be detrimental for the auditor's reputation because it portrays a higher likelihood of an undetected fraud in the past.

4.1.1. Comparing the Expected Fraud Across Periods. To better understand the result in Proposition 1, we examine the expected amount of fraud from the market's point of view. To do so, we define the following two measures of the expected fraud that are computed at the end of each period:

Definition 1. (a) *Public expected fraud* (*PEF*) is the market's assessment of the probability that fraud was committed in the period. So, the public expected fraud in period 1 is $PEF_1 = \mathbb{P}(m_1 = f)$ and that in period 2 is $PEF_2 = \mathbb{P}(m_2 = f \mid a_1 = nr)$.

(b) *Conditional expected fraud* (*CEF*) is the market's assessment of the probability that fraud was committed in the period conditional on the manager being dishonest. So, the conditional expected fraud in period 1 is $CEF_1 = \mathbb{P}(m_1 = f \mid \text{Manager is dishonest})$ and that in period 2 is $CEF_2 = \mathbb{P}(m_2 = f \mid a_1 = nr, \text{Manager is dishonest})$.

The following result illustrates the time evolution of these two market perceptions of fraud.

Corollary 1. *When the manager's reputation is high, $y_0 > \bar{y}$, the public expected fraud is higher in the first period than in the second, whereas the conditional expected fraud is lower in the first period than in the second.*

The dishonest manager commits fraud in the first period only with some probability, but commits fraud in the second period with certainty. Thus, the conditional expected fraud is higher in the second period than in the first one. Even though the dishonest manager increases the conditional expected fraud with time, the public expected fraud decreases. Whenever fraud is not detected, the market updates favorably its perception about the manager's honesty. That is, the reputation of the manager improves. Moreover, this reputation improvement is always enough to compensate for the increase in fraud committed by the dishonest manager and make the overall public expected fraud smaller in the second period.

Corollary 1 thus completes the explanation for the equilibrium in Proposition 1. From a public perspective, the situation improves because the expected probability of fraud decreases with time. Nevertheless, the dishonest manager's behavior becomes more daring with time as the auditor finds himself in a reputation trap. This public perception results from an underlying equilibrium in which the manager starts small, just enough "to seed the trap and handcuff" the auditor in the next period. If the manager is lucky and does not get caught, he then increases the amount of fraud as he knows the auditor now only reports with a small frequency. Indeed, by reporting fraud in the second period, the auditor admits the possible presence of undetected fraud in the first period, making the market suspicious of his ability.

4.2. The Testing the Waters Effect: $y_0 < \bar{y}$

When the manager's reputation is lower than the reputation threshold, $y_0 < \bar{y}$, the market expects fraud to occur with high probability. A distrustful market interprets the absence of a fraud report as a likely undetected or detected but unreported fraud. This presumption exerts considerable pressure on the auditor to report. As a result, the manager becomes

cautious and more selective in committing fraud in the second period. Defining

$$P_m = \frac{P_h + \gamma\delta(\bar{y} + P_l)P_l}{1 + \gamma\delta(\bar{y} + P_l)}, \tag{6}$$

the following proposition describes this equilibrium formally.

Proposition 2 ("Testing the Waters" Equilibrium). *For the case $y_0 < \bar{y}$, we have the following unique equilibrium*:

(1) (*Mixed strategy*: *Auditor may conceal fraud*). *If $\gamma + \delta > 1$ and $P_m < P < P_h$,*

$$\mu^* = \left\{\frac{1-\bar{y}}{1-y_0}, 1, 0\right\} \quad \textit{and} \tag{7}$$

$$\alpha^* = \frac{2 - (2+P)(\gamma x_0 + (1-x_0)\delta)}{(1+P)((1-\delta)\delta(1-x_0) + \gamma(1-\gamma)x_0)}. \tag{8}$$

(2) (*Pure strategy* 1: *Manager does not commit fraud*). *If $\gamma + \delta < 1$ and $P_m < P < P_h$, $\mu^* = \{0, 1, 0\}$ and $\alpha^* = 1$.*

(3) (*Pure strategy* 2: *Auditor never conceals fraud*). *If $P_l < P < P_m$, $\mu^* = \{1, 1, 0\}$ and $\alpha^* = 1$.*

On a technical note, we remind the reader that in all of this paper with the exception of this result, we use the sequential equilibrium concept. In Proposition 2, this approach does not avoid the common problem of multiplicity of equilibria due to arbitrary beliefs in states that have a zero probability of occurrence in equilibrium. Here, we use the concept of proper equilibrium as a refinement of the sequential equilibrium (we refer the reader to Definition EC.1 in Online Appendix EC.2.5 for the definition of a proper equilibrium).

Proposition 2 describes an equilibrium characterized by two different ranges of penalties. For high penalties, $P_m < P < P_h$, and high audit technology, $\gamma + \delta > 1$, there is a unique mixed strategy equilibrium in which the manager's first-period strategy mixes with the auditor's second-period strategy. Notice also that the manager either commits fraud in both periods or never commits fraud. That is, the outcome of the manager's randomization in the first period is played again in the second period. For the case of high penalties, $P_m < P < P_h$, and low audit technology, $\gamma + \delta < 1$, there is a unique pure strategy equilibrium in which the auditor would always report any fraud detected, and in response the manager never commits fraud. Finally, for low penalties, $P_l < P < P_m$, the manager commits fraud with certainty in both periods, and the auditor responds by reporting fraud whenever he detects it. Thus, in this case the reputation incentive works perfectly.

Consider first the mixed strategy equilibrium for high penalties, $P_m < P < P_h$. The fact that the manager only commits fraud in period 2 when he committed fraud in the past makes a second-period fraud report very informative. Not only does it identify the manager as dishonest, but it also reveals the presence of undetected fraud in the first period. Once the market is certain about the presence of fraud in the first period, the market's inference about the manager's mixed strategy ceases to affect the auditor's reputation. However, in the event of no fraud being reported, the market infers the likely presence of an unreported fraud in both periods because of the low reputation of the manager. If the auditor cannot meet the market expectations with a fraud report, his reputation suffers greatly. The market's inference about the first-period manager's mixed strategy then becomes relevant. The larger the probability of fraud in the first period, the more likely is the presence of undetected fraud and the more the auditor's reputation suffers from the absence of a fraud report; that is, an increase in the probability of fraud in the first period makes the auditor more willing to report any detected fraud in the second period. In contrast to the handcuff equilibrium, here the manager's fraud in the first period motivates the auditor to report in the second period.

The manager's purpose for committing fraud in the first period is merely to "test the waters." By committing fraud with a nonzero probability, the manager incurs a negative payoff in the first period. Nevertheless, if the fraud remains undetected, he then deems the auditor to be of lower ability and that allows him to commit fraud with certainty in the second period. It is precisely the contingency of the manager's decision on the information he obtained previously that shapes this equilibrium. A fraud report here reveals the manager's private information gained by previously testing the auditor. Consequently, the auditor becomes less inclined to report fraud, and we obtain the mixed equilibrium.

Next, consider the pure strategy 1 equilibrium, i.e., the equilibrium for high penalties ($P_m < P < P_h$) and low audit technology ($\gamma + \delta < 1$). In this case, the market's expectation of fraud ever being detected is low enough so that any report is considered an indication of the high-ability auditor. In the absence of a fraud report, the auditor's reputation can be at best the initial reputation, x_0. With no additional information, the market can only infer that if fraud was committed it remained undetected. However, reporting fraud at any time raises the auditor's reputation beyond x_0. In essence, if $\gamma + \delta < 1$, then the likelihood of an undetected fraud in the first period and a detected fraud in the second period is higher for the high-ability auditor than that for the low-ability auditor, i.e., $(1-\gamma)\gamma > (1-\delta)\delta$. Therefore, the market sees a fraud report as a good sign of the auditor's ability. Thus, in this case it is always best for the auditor to report any detected fraud. Thus, in response to the auditor's strategy, the

manager's best response is to *never* commit fraud for high penalties.

Finally, we consider the pure strategy 2 equilibrium, i.e., the equilibrium for low penalties ($P_l < P < P_m$). In this case, the dishonest manager considers it profitable to commit fraud with certainty in both periods. A lower penalty directly reduces the expected punishment from committing fraud. This increases the manager's willingness to incur a negative expected payoff in the first period to access a future positive expected payoff. As a result of the increased likelihood of fraud in the first period, the absence of a fraud report becomes more costly, inducing the auditor to always report any detected fraud.

4.2.1. Comparing the Expected Fraud Across Periods. As in the previous subsection, we can examine the expected amount of fraud from the market's point of view. We find that the public expected fraud is either higher in the first period than in the second (similar to the handcuff equilibrium) or remains the same; that is, the market expects less fraud to be committed as time progresses. Contrary to Corollary 1 though, in this case, the conditional expected fraud weakly decreases with time; that is, the dishonest manager is expected to commit the same or less fraud in the second period than in the first period. The following corollary states this result.

Corollary 2. *When the manager's reputation is low, $y_0 < \bar{y}$, we have the following results*:

(1) (*Mixed strategy*). *If $P_m < P < P_h$ and $\gamma + \delta > 1$, the public expected fraud, and the conditional expected fraud are higher in the first period than in the second*.

(2) (*Pure strategy* 1). *If $P_m < P < P_h$ and $\gamma + \delta < 1$, the public expected fraud is higher in the first period than in the second, and the conditional expected fraud is the same in both periods*.

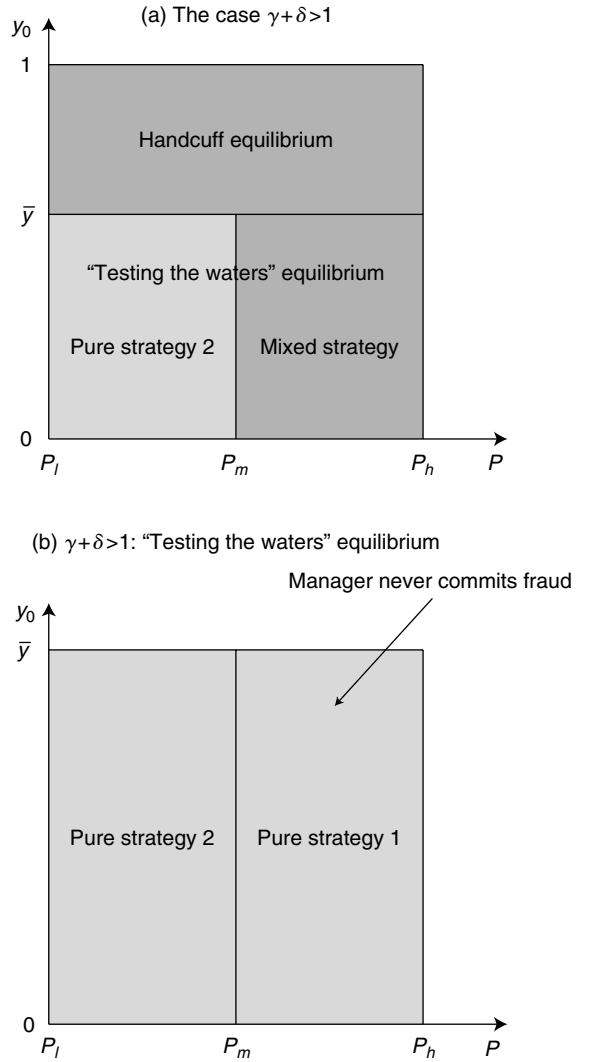
(3) (*Pure strategy* 2). *If $P_l < P < P_m$, the public expected fraud is higher in the first period than in the second, and the conditional expected fraud is the same in both periods*.

In the mixed strategy equilibrium, the dishonest manager tests the waters by committing fraud with nonzero probability in the first period. His intention is to gain information about the auditor's ability and benefit from it in the second period. However, unless he gains information by actually committing fraud and remaining undetected, the manager does not commit fraud again. The conditional expected fraud is then smaller in the second period. Only for low penalties does the dishonest manager commit fraud with certainty in both periods; therefore, the conditional expected fraud remains the same. Given this result and the fact that the reputation of the manager improves when fraud is not detected, it is no surprise that the public expected fraud is strictly smaller in the second period. In the pure strategy 1 equilibrium, the result is straightforward because fraud is never committed.

5. Discussion of Equilibria

Figure 3 summarizes the different scenarios presented in §4. The lighter shade depicts regions where the auditor always reports fraud, and the darker shade depicts regions where the auditor may conceal fraud. Note that the auditor always reports detected fraud when the manager's initial reputation is low ($y_0 < \bar{y}$) and either the penalty is low ($P_l < P < P_m$) or the audit technology is low ($\gamma + \delta < 1$). On the other hand, the auditor may conceal fraud if the manager's initial reputation is high ($y_0 > \bar{y}$), or if the penalty and audit technology are both high ($P_m < P < P_h$ and $\gamma + \delta > 1$).

Figure 3 The Unique Equilibrium Obtained as a Function of the Problem Parameters



Notes. The auditor may conceal detected fraud in the darker regions, whereas he always reports detected fraud in the lighter regions. Furthermore, when $\gamma + \delta < 1$ and penalties are high, the manager never commits fraud in equilibrium.

To better understand the equilibria, we provide an intuitive explanation.

When the manager has a high reputation, the auditor's reputation suffers little damage from not reporting fraud in the second period because the market expects little fraud to take place. The resulting equilibrium portrays the dishonest manager committing fraud with some probability in the first period. This has the effect of "handcuffing" the auditor in the second period. Indeed, by reporting in the second period, the auditor reveals that the manager is dishonest and with it, the likely presence of undetected fraud in the first period. Consequently, if the manager remains undetected, he then feels safe to commit fraud with certainty in a lucrative second period.

With a low-reputation manager, however, the market expects fraud to be committed, so silence has a high cost for the auditor. The auditor is therefore more willing to report detected fraud. This cautions the manager, who deems it safe to commit fraud in the second period only if he committed fraud previously and remained undetected. That is, the manager commits fraud in the first period mainly to "test the waters." Nevertheless, the auditor feels trapped because he is forced to choose between revealing fraud, thereby revealing the certain presence of a missed fraud in the first period, and concealing fraud, thereby letting the market raise the suspicion of fraud being undetected in both periods. The resolution of this trade-off depends on the amount of the penalty and the audit technology. For high penalty and high audit technology, the auditor may conceal fraud with some probability, but for all other cases the auditor always reports any detected fraud.

5.1. Comparison of the Auditor's Strategy in Handcuff and "Testing the Waters" Equilibria

The above discussion stresses the role of the manager's initial reputation in shaping the equilibrium. This initial reputation sets the market's perception of the likelihood of fraud; thus, a low value of this reputation should make the auditor more inclined to report detected fraud. The following result formalizes this intuition and proves that the auditor is always more likely to report detected fraud in the "testing the waters" equilibrium, where the manager's initial reputation is lower than that in the corresponding handcuff equilibrium.

Proposition 3. *For any penalty $P_l < P < P_h$, any auditor's initial reputation $0 < x_0 < 1$, and any auditor's detection probabilities $0 < \delta < \gamma < 1$ such that $\gamma + \delta > 1$, the auditor's equilibrium strategy α^* is greater in the "testing the waters" equilibrium than in the handcuff equilibrium.*

5.2. Comparative Statics for Mixed Strategy Equilibria

We now examine the effect of the different model parameters, namely, the auditor's and the manager's initial reputation, the penalty, and the detection probabilities, on the equilibrium strategies. The following proposition summarizes the comparative statics.

Proposition 4. (1) *The auditor's reporting strategy in the second period α^* in the handcuff equilibrium ("testing the waters" mixed strategy equilibrium) is increasing (decreasing) in the auditor's initial reputation x_0, the penalty P, and in the detection probabilities (γ and δ, maintaining the difference in the probabilities $\gamma - \delta$ constant*[13]).

(2) *The manager's strategy μ_1^* in the handcuff equilibrium ("testing the waters" mixed strategy equilibrium) is decreasing (increasing) in his initial reputation y_0 and in the detection probabilities (γ and δ, maintaining the difference in the probabilities $\gamma - \delta$ constant).*

The auditor's initial reputation complements the detection probability and the penalty as means of deterring fraud. With a more reputable auditor, the manager becomes more cautious in committing fraud in both periods. Nevertheless, the effect on the auditor's reporting strategy differs across scenarios. When the manager's reputation is high (handcuff equilibrium, $y_0 > \bar{y}$), the presence of fraud in period 1 is what deters the auditor from reporting in the second period. Consequently, by curbing the manager's incentives to commit fraud in the first period, a better auditor reputation allows the auditor to report with a higher probability in the second period. The opposite effect develops when the manager's reputation is low ("testing the waters" equilibrium, $y_0 < \bar{y}$). In this scenario, a lower first-period fraud probability induces a lower reporting frequency in the second period. Indeed, the perception of a lower fraud probability in the first period does not affect the market's assessment of the auditor's ability once fraud is reported, but reduces the market's suspicions in the event of no report. As a result, a higher auditor reputation allows the auditor to relax his reporting strategy in the second period.

We next study the impact of reputation on the public expected fraud committed in equilibrium. The following result summarizes this impact.

Proposition 5. *In equilibrium, the public expected fraud in periods 1 and 2 is nonincreasing in both the auditor's initial reputation x_0 and the manager's initial reputation y_0.*

[13] Relaxing this restriction introduces a contraposed effect that makes the resulting comparative statics parameter dependent.

Proposition 5 states that even though reputational incentives may lead an auditor onto a slippery slope in some equilibria, the auditor's reputation nevertheless is a consistent measure of the auditor's performance in all equilibria in the sense that auditors with higher reputations lead to lower expected fraud. Similarly, all else being equal, the manager's initial reputation is also a consistent measure of his honesty, i.e., firms with managers of higher reputation have lower expected fraud.

Proposition 5 is reassuring because it implies that reputation *does matter*. That is, even though auditors may conceal detected fraud at times to preserve their reputation, auditors with higher reputation still perform better than those with lower reputation on average.

5.3. Robustness of Equilibria: Presence of a Whistle-Blower

Thus far in this paper, the auditor is the only means of uncovering fraud. In reality, however, fraud may be revealed by a whistle-blower (or by regulatory agencies, such as the Securities and Exchange Commission, the Public Company Accounting Oversight Board, etc., or by subsequent auditors). The existence of these alternative means of revelation may make the auditor less willing to conceal detected fraud. For instance, in the extreme case that the whistle-blower uncovers *all* committed fraud with certainty, the auditor has no option but to reveal *all* detected fraud. This naturally leads to the question of the robustness of the equilibria in which the auditor conceals detected fraud. To answer this question, we focus on the two mixed equilibria in which the auditor conceals fraud with some probability, and we introduce the revelation of committed fraud by the whistle-blower with some fixed probability. Because the auditor always reports fraud in the first period and our goal is to study the robustness of the auditor's concealment decision in the second period, we restrict our attention to the case in which the whistle-blower might reveal fraud that is committed only in the second period. That is, if the auditor does not reveal fraud in the second period, the whistle-blower detects and reports it with probability p at the end of the second period. Finally, we assume that if committed fraud is revealed by the whistle-blower, the auditor's terminal reputation becomes $\underline{x}$, which is independent of p.

The following result demonstrates that the slippery-slope phenomenon described in the previous section is robust with respect to the presence of this independent means of fraud revelation.

Proposition 6. (1) (*Handcuff equilibrium*). *If $y_0 > \bar{y}$ and $P_l < P < P_h$, then there exists $\bar{p}_h > 0$ such that for all $p < \bar{p}_h$ there is an equilibrium with strategies $\mu^* = \{\mu_1^*(p), 1, 1\}$ and $\alpha^*(p)$ with $0 < \mu_1^*(p), \alpha^*(p) < 1$.*

(2) (*"Testing the waters" mixed strategy equilibrium*). *If $y_0 < \bar{y}$, $P_m < P < P_h$, and $\gamma + \delta > 1$, then there exists $\bar{p}_t > 0$ such that for all $p < \bar{p}_t$ there is an equilibrium with strategies $\mu_1^* = \{\mu_1^*(p), 1, 0\}$ and $\alpha^*(p)$ with $0 < \mu_1^*(p)$, $\alpha^*(p) < 1$.*

This result proves that all the equilibria where the auditor may conceal fraud will continue to exist even when a whistle-blower might reveal unreported fraud. This result follows from the fact that the payoffs of the auditor and manager are continuous in the probability p, and clearly holds for settings more general than the one described above, for example, when there is a whistle-blower in both periods.

To illustrate the extent to which this equilibrium is robust, we consider the case in which the auditor's terminal payoff is linear in his terminal reputation, i.e., $\tau(x_2) = a + bx_2$ for some $a, b > 0$, and in the case that fraud is revealed by the whistle-blower, the auditor completely loses his reputation, i.e., $\underline{x} = 0$. As is demonstrated by the following result, the handcuff and "testing the waters" equilibria continue to hold as long as the probability of fraud being uncovered by a whistle-blower is less than a threshold.

Proposition 7. *If the auditor's terminal payoff is linear in his terminal reputation and $\underline{x} = 0$, then*

(1) (*Handcuff equilibrium*). *If $y_0 > \bar{y}$ and $P_l < P < P_h$, then there exist probabilities $\alpha^*(p)$ and $\mu_1^*(p)$ such that $\mu^* = \{\mu_1^*(p), 1, 1\}$ and $\alpha^*(p)$ is an equilibrium for $p < \bar{p}_h$ with*

$$\bar{p}_h = \min\left(\frac{\delta(1-\delta-P\delta)+(\gamma-\delta)x_0(1+\gamma-(3+2P)\delta-(2+P)(\gamma-\delta)x_0)}{(1+P)(\gamma-\delta)^2(1-x_0)x_0}, \bar{q}_h\right), \tag{9}$$

where $\bar{q}_h$ solves

$$\bar{q}_h = \frac{\gamma\delta(\gamma-\delta)(1-x_0)(y_0-\bar{y})}{((1-\delta)\delta(1-x_0)+\gamma(1-\gamma)x_0)(1-\gamma(1-y_0)-(1-\gamma)\gamma(1-y_0)\alpha^*(\bar{q}_h))}.$$

(2) (*"Testing the waters" equilibrium*). *If $y_0 < \bar{y}$ and $P_m < P < P_h$, then there exist probabilities $\alpha^*(p)$ and $\mu_1^*(p)$ such that $\mu^* = \{\mu_1^*(p), 1, 0\}$ and $\alpha^*(p)$ is an equilibrium for $p < \bar{p}_t$ with*

$$\bar{p}_t = \min\left(\frac{\delta(1-\delta(1+P))+(\gamma-\delta)x_0(1+\gamma-(3+2P)\delta-(2+P)(\gamma-\delta)x_0)}{(1+P)(\gamma-\delta)^2(1-x_0)x_0}, \frac{(\gamma-\delta)(\gamma+\delta-1)(1-x_0)}{(1-\delta)\delta(1-x_0)+\gamma(1-\gamma)x_0}\right). \tag{10}$$

Note that the value of the threshold $\bar{p}_h$ can be as high as $\bar{p}_h = 1$ in the limit as $x_0 \to 0$, $\gamma \to 1$, $\delta = 1/2$, and $y_0 \to 1$. That is, for any probability of whistle-blowing p, there are parameters $(\gamma, \delta, x_0, y_0, P)$ such that the handcuff equilibrium holds. However, we find that the maximum value of the threshold for the "testing the waters" equilibrium is $\bar{p}_t = 0.17$ (in the limit as $x_0 \to 0$, $\gamma \to 1$, $\delta = 0.58$, and $P \to P_m$); that is,

for probability of whistle-blowing $p > 0.17$, the "testing the waters" equilibrium ceases to exist and reputational incentives work perfectly.

This robustness check illustrates that even in the presence of a whistle-blower and the threat of a total loss of reputation, the auditor may conceal fraud with positive probability. However, if one considers a model in which the auditor incurs additional liabilities that go beyond the total loss of reputation, we expect the thresholds derived in Proposition 7 to decrease as these liabilities increase, and thus reduce the instances in which the auditor conceals fraud. Because the focus of this paper is on reputational incentives, we do not investigate such liabilities.

6. Conclusion

This paper studies reputation incentives as a sole means of motivating the auditor to perform his task responsibly. We streamline the audit context to focus on a more complex depiction of reputation formation than that examined by previous auditing literature. As the manager, the market and the auditor have different information sets, we allow them to have different perceptions of the auditor's ability. Similarly, the three players also have different perceptions of the manager's honesty. We model and study the dynamic interaction among these individual beliefs in detail.

The analysis in this paper does not consider any discipline mechanism for the auditor other than reputation concerns. In such a streamlined setting we find that, when the market perceives fraudulent behavior from the manager as a likely event, the auditor never conceals detected fraud to maximize his reputation; that is, when the manager is perceived as dishonest (low reputation) and either the manager's penalty or the audit technology is low, reputation concerns work without any additional disciplining mechanism. It is in the remaining scenarios, in which the market does not expect such fraudulent behavior, that the auditor may conceal fraud. It is precisely in these circumstances, in which the manager is perceived as honest (high reputation) or the manager's penalties are high, that a careful scrutiny by regulatory authorities and additional disciplining mechanisms are required to remedy the shortcomings of reputation incentives. The analysis of such disciplining mechanisms in conjunction with reputational incentives is beyond the scope of this paper, and is left for future work.

We conclude this paper by discussing some potential implications for empirical studies of audit quality. To the extent that what we denote as fraud in this paper can be interpreted more loosely as any reporting irregularity, we can use the results in this paper to discuss some proxies of audit quality. Audit quality can be thought of as the probability that the auditor detects and reports the presence of such reporting irregularities. However, audit quality is not publicly observable because it is conditional on the knowledge about the presence of such reporting irregularities. Therefore, empirical work must rely on observable proxies such as the frequency of financial restatements (see, for instance, Stanley and DeZoort 2007, Romanus et al. 2008) and the market reaction to financial reporting (estimated as the earnings response coefficient; see, for instance, Balsam et al. 2003, Ghosh and Moon 2005). As we discuss below, our paper suggests caution in using such one-dimensional proxies for audit quality.

When using financial restatements as a proxy of audit quality, there is an implicit assumption that the auditor has no option but to restate earnings when an irregularity is found. If indeed that is the case, then restatements are a good proxy for the frequency of missed past irregularities and, therefore, a proxy for poor audit quality. Nevertheless, our model shows that when restating earnings is a decision that the auditor must consider, a restatement could imply good audit quality. In fact, if the auditor decides to report an irregularity, it could be the case that his reputation increases by doing so. That is, a restatement increases the market's perceived probability that the auditor is of the high type and, therefore, that the audit in period 1 is of high quality (note that the auditor always reports irregularities detected in period 1, and thus the probability that an irregularity is reported is equivalent to the probability of detection).

The investor's perception of audit quality is, in some empirical studies, proxied with the earnings response coefficient. Using this proxy assumes that a more credible financial statement corresponds to a better audit quality. Although this seems intuitive, there are other factors that determine the credibility of financial statements that might be confounded with audit quality. Indeed, in our model the public expected fraud decreases with tenure, indicating that the market expects irregularities to occur with a lower probability, and therefore making financial reports more credible. However, the expected probability of irregularities might be lower not only because of a higher-quality audit, but also because investors regard the manager as being honest with a higher probability. In fact, we find that the market perceived probability of a report conditional on the presence of irregularities decreases with tenure, indicating that the audit quality *decreases* with tenure;[14] that is, the market's perception of the manager's honesty increases with time and dominates the decrease

[14] The market perceived probability of a report conditional on the presence of irregularities equals $(x_0\gamma + (1 - x_0)\delta) \times$ *probability that auditor reports detected irregularities*. In period 1, the auditor always reports detected irregularities, whereas in period 2, the auditor may conceal these with some probability.

in audit quality. The overall effect is an increase in the credibility of financial statements with tenure, in spite of a decrease in audit quality. This insight indicates that, when investigating the effects of tenure on audit quality, the empiricist may need to control for other confounding factors such as the manager's reputation.

This paper studies a two-period model. A possible extension would be to consider settings with more than two periods to explore audit tenure regulatory implications. We expect our insights to carry through in some form to such a setting. For instance, we expect that the dishonest manager's propensity to commit fraud would increase with audit tenure. However, the model would have to incorporate the learning curve that the auditor experiences while dealing with the same firm repeatedly. One would expect the auditor's ability to improve with time, and hence the resultant audit quality would depend on the interaction between this learning curve and the slippery-slope phenomenon. We leave a detailed investigation such a model for a future endeavor.

7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

The authors thank Frank Gigler, Chandra Kanodia, Paul Newman, and the seminar participants at University of Minnesota, Baruch College, and the University of Texas at Austin for their helpful comments and suggestions. The authors also thank the two anonymous referees and the associate editor for their constructive comments.

References

Baiman, S., J. H. Evans III, N. Nagarajan. 1991. Collusion in auditing. *J. Accounting Res.* **29**(1) 1–18.

Baiman, S., J. H. Evans III, J. Noel. 1987. Optimal contracts with a utility-maximizing auditor. *J. Accounting Res.* **25**(2) 217–244.

Baiman, S., P. E. Fischer, M. V. Rajan. 2000. Information, contracting, and quality costs. *Management Sci.* **46**(6) 776–789.

Balachandran, K. R., N. J. Nagarajan. 1987. Imperfect information, insurance, and auditors' legal liability. *Contemporary Accounting Res.* **3**(2) 281–301.

Balachandran, K. R., S. Radhakrishnan. 2005. Quality implication of warranties in a supply chain. *Management Sci.* **51**(8) 1266–1277.

Balsam, S., J. Krishnan, J. S. Yang. 2003. Auditor industry specialization and earnings quality. *Auditing* **22**(2) 71–97.

Beyer, A., S. S. Sridhar. 2006. Effects of multiple clients on the reliability of audit reports. *J. Accounting Res.* **44**(1) 29–51.

Coffee, J., Jr. 2004. Gatekeeper failure and reform: The challenge of fashioning relevant reforms. Working paper, Program in Law and Economics, School of Law, University of California, Berkeley, Berkeley.

Datar, S., M. Alles. 1999. The formation and role of reputation and litigation in the auditor-manager relationship. *J. Accounting, Auditing, Finance* **14**(4) 401–428.

DeAngelo, L. 1981. Auditor size and audit quality. *J. Accounting Econom.* **3**(3) 183–199.

Dye, R. 1993. Auditing standards, legal liability, and auditor wealth. *J. Political Econom.* **101**(5) 887–914.

Dye, R., B. Balachandran, R. Magee. 1990. Contingent fees for audit firms. *J. Accounting Res.* **28**(2) 239–266.

Ghosh, A., D. Moon. 2005. Auditor tenure and perceptions of audit quality. *Accounting Rev.* **80**(2) 585–612.

Kanodia, C., R. Bushman, J. Dickhaut. 1989. Escalation errors and the sunk cost effect: An explanation based on reputation and information asymmetries. *J. Accounting Res.* **27**(1) 59–77.

Kinney, W., Jr. 2005. Twenty-five years of audit deregulation and reregulation: What does it mean for 2005 and beyond? *Auditing: J. Practice Theory* **24** 89–109.

Kofman, F., J. Lawarree. 1993. Collusion in hierarchical agency. *Econometrica* **61**(3) 629–656.

Laux, V., D. P. Newman. 2010. Auditor liability and client acceptance decisions. *Accounting Rev.* **85**(1) 261–285.

Lee, C., Z. Gu. 1998. Low balling, legal liability and auditor independence. *Accounting Rev.* **73**(4) 533–555.

Maggin, D. L. 1989. *Bankers, Builders, Knaves, and Thieves: The $300 Million Scam at ESM*. Contemporary Books, Chicago.

Morton, S. 1993. Strategic auditing for fraud. *Accounting Rev.* **68**(4) 825–839.

Narayanan, V. G. 1994. An analysis of auditor liability rules. *J. Accounting Res.* **32** 39–59.

Newman, D. P., E. Patterson, R. Smith. 2001. The influence of potentially fraudulent reports on audit risk assessment and planning. *Accounting Rev.* **76**(1) 59–80.

Radhakrishnan, S. 1999. Investors' recovery friction and auditor liability rules. *Accounting Rev.* **74**(2) 225–240.

Romanus, R. N., J. J. Maher, D. M. Fleming. 2008. Auditor industry specialization, auditor changes, and accounting restatements. *Accounting Horizons* **22**(4) 389–413.

Sarath, B. 1991. Uncertain litigation and liability insurance. *RAND J. Econom.* **22**(2) 218–231.

Schwartz, R. 1997. Legal regimes, audit quality and investment. *Accounting Rev.* **72**(3) 385–406.

Schwartz, R. 1998. Auditors' liability, vague due care and auditing standards. *Rev. Quant. Finance Accounting* **11**(2) 183–207.

Simunic, D. 1980. The pricing of audit services: Theory and evidence. *J. Accounting Res.* **18**(1) 161–190.

Stanley, J. D., F. T. DeZoort. 2007. Audit firm tenure and financial restatements: An analysis of industry specialization and fee effects. *J. Accounting Public Policy* **26**(2) 131–159.

Titman, S., B. Trueman. 1986. Information quality and the valuation of new issues. *J. Accounting Econom.* **8**(2) 159–172.