Managerial Teams, Whistle-Blowing, and Accounting Fraud

Stephen Cotten and Rudy Santore*

ABSTRACT

We examine the effects of Sarbanes-Oxley provisions pertaining to whistle-blower protections and reporting requirements on a managerial team’s incentive to commit accounting fraud. Our analysis predicts that whistle-blowing does not occur in equilibrium, but that the whistle-blower protections combined with the reporting requirements can reduce fraud, and are most likely to do so when managers are heterogeneous in their aversion to sanctions. Interestingly, amnesty provisions have no effect on the equilibrium level of fraud. In line with previous literature, we find that equity compensation induces managerial effort, but also provides the incentive for management to fraudulently misreport the financial health of the firm.

*Department of Economics, University of Tennessee, Knoxville, TN. We would like to thank Phillip Daves, William Neilson, Christopher Snyder, Christian Vossler, and seminar participants at the 5th Annual International Industrial Organization Conference, participants in the University of Tennessee Corporate Governance Center Seminar Series, and participants in the University of Tennessee Economics Brown Bag Workshop for useful comments.
1. INTRODUCTION

Throughout the 1980s and 1990s corporations embraced the recommendations of agency theory by utilizing equity compensation to incentivize managers (Murphy, 1999). However, the recent high-profile exposure of illegal accounting practices in numerous corporations, such as Enron and Worldcom, has led both researchers and policy-makers to acknowledge that fraud may be an unexpected consequence of these compensation schemes. It is now understood that at least one major roadblock to the practical application of agency theory to the corporate setting is that managers may manipulate a firm’s stock price by misreporting the financial health of the firm. As explained by the AICPA, “Management may override controls to intentionally misstate the nature and timing of revenue or other transactions by (1) recording fictitious business events or transactions or changing the timing of legitimate transactions, particularly those recorded close to the end of an accounting period; (2) establishing or reversing reserves to manipulate results, including intentionally biasing assumptions and judgment used to estimate account balances; and (3) altering records and terms related to significant or unusual transactions.” (American Institute of Certified Public Accountants, 2005)

In response to recent accounting scandals, lawmakers passed the Sarbanes-Oxley Act of 2002. The Act attempts to reduce fraud by strengthening corporate governance and reporting requirements, and our analysis focuses on two of Sarbanes-Oxley’s provisions. The first is the protection from retaliation given to whistle-blowing employees. As discussed by Watnick (2006), “In attempting to reform American business practices, Congress impressed into service corporate officers, directors, and other corporate employees, enlisting them as ‘foot soldiers’ in the fight against corporate fraud. Congress did so by requiring those who witness corporate fraud to report what they know about it and by offering

---

2 See Watnick (2006) for a comprehensive review of the statutory language, legislative history, and regulations pursuant to the Act, as well as analysis of the effectiveness of the provisions and legal cases relating to the whistleblower provisions.
commiserate protection from retaliation under the ‘whistle-blower protection’ provisions contained within Sarbanes-Oxley.”

The second provision of interest is the requirement that management establish and affirm the adequacy of their financial reports. Specifically, “the principal executive officer or officers and the principal financial officer or officers, or persons performing similar functions” must certify that they are “responsible for establishing and maintaining internal controls” and “have designed such internal controls to ensure that material information relating to the company and its consolidated subsidiaries is made known to such officers by others within those entities.” Management must ensure, with threat of criminal penalties for noncompliance, that “information contained in the periodic report fairly presents, in all material respects, the financial condition and results of operations of the issuer.”

In this paper, we consider the implications of supporting whistle-blowers (through employee protections) and requiring complicity in misreporting (through having managers sign off on financial reports) on the incentives of a managerial team to commit fraud. We provide and analyze a model in which a team of managers receives equity compensation to run a firm. The team provides productive effort, which increases the true value of the firm, and also reports the value to the market. The reported value, however, may be fraudulently inflated by one or more managers. It is assumed that all managers know whether fraud has taken place, and thus each manager is able to blow the whistle prior to sending the report to the market. Managers differ in their sensitivity to sanctions, implying that some managers are more inclined to report fraud than others. A manager who blows the whistle may be given complete or partial amnesty for his or her individual contribution. If no manager reports the fraud, there is a chance that the fraud will later be detected by authorities, in which case every manager is treated as being complicit and receives an identical sanction.

---

5 Ibid
6 Ibid
7 18 U.S.C. 1350, “Failure of corporate officers to certify financial reports.”
8 We use the same definition of team as found in Holmstrom (1982), namely, “A group of individuals who are organized so that their productive inputs are related.”
We find that, in equilibrium, only a fraction of the (heterogeneous) managers actively participate in the fraud, yet no manager blows the whistle. This is true despite the fact that the entire team is sanctioned with positive probability. The reason that no manager blows the whistle is that every member of the team holds stock in the firm and thus benefits if the fraud goes undetected. At the same time, those managers who actively commit fraud refrain from over-reporting at a level that would cause whistle-blowing by the managers who do not actively participate in the illegal behavior. As a result, rules that protect whistle-blowing and that effectively mandate complicity among the managerial team (by having management sign off on reports created by others) do keep fraud in check, but do not eradicate it entirely.

Our analysis also helps shed light on the notable absence of amnesty provisions for whistle-blowers in the Sarbanes-Oxley Act or, more generally, federal sentencing guidelines. Surprisingly, amnesty provisions are found to have no effect on the magnitude of fraud because those who are most inclined to blow the whistle do not actively participate in illegal over-reporting. These potential whistle-blowers would not receive sanctions even if they were to report the fraud.

Theoretical research exploring the effects of equity compensation on accounting fraud has mostly focused on single-agent models. Goldman and Slezak (2006) develop a model in which a manager responds to an equity contract by choosing a level of costly effort and a level of fraudulent manipulation. They find that an increase in equity compensation leads to greater incentives to commit fraud. However, an increase in the probability of detection has an ambiguous effect since the owner may increase the pay-for-performance sensitivity in response. Bar-Gill and Bebchuk (2003) use a dynamic model to analyze the decisions of firm managers over time. They include legal and illegal manipulation in their model and find that increases in exercisable equity compensation lead to greater incentives to engage in fraudulent behavior. Their model allows them to examine the impact of changes in the ability of managers to sell

---

9 Robison and Santore (2008) study an owner’s incentive to invest in both ex ante monitoring, which decreases the likelihood that fraud is possible, and ex post monitoring, which increases the probability that fraud will be detected after the fact.
10 Unlike Bar-Gill and Bebchuk (2003), Goldman and Slezak (2006) assume that all equity is exercisable.
their shares, and they find that tightening controls on the ability of managers to liquidate decreases the incentive to commit fraud.

While the papers discussed above consider managerial fraud and allow for the possibility of detection, they do not allow for the strategic decision to blow the whistle. In fact, the only other paper of which we are aware that formally models the decision to blow the whistle is by Friebel and Guriev (2005). These authors assume that upper management receives equity compensation and can inflate earnings. However, in order to prevent lower management from blowing the whistle, upper management must give lower management additional compensation. Despite some similarities, their paper differs from ours in a few important aspects. First, our managers hold identical positions within the managerial team and each is able to actively participate in the fraud. Friebel and Guriev have two levels of manager and only top management can actively commit fraud. Second, unlike Friebel and Guriev, we do not allow the team to manipulate the compensation packages in order to prevent whistle-blowing. Third, whistle-blowing in our model is motivated by the potential for team punishment when fraud is detected, which can occur even if no manager blows the whistle. However, Friebel and Guriev assume that fraud is detected only if lower management blows the whistle.

There has also been a sizable array of empirical research on the impact of executive compensation on accounting malfeasance. Johnson, Ryan, and Tian (Forthcoming) compare a sample of firms under suspicion by the SEC of engaging in fraudulent behavior with a control sample of “innocent” firms. They find that the likelihood of fraud is correlated with unrestricted stockholdings and that unrestricted stockholdings comprise the largest share of managerial incentives at firms guilty of fraud. Peng and Roell (2006) find greater pay in the form of stock options increases the probability that a firm will be involved in securities litigation. Burns and Kedia (2006) use a sample of firms that restated earnings between 1995 and 2001 to look at the relationship between CEO compensation and misreporting, and they find that there is a significant and positive relationship between the sensitivity of a CEO’s compensation package to a change in the stock price and the propensity to misreport earnings. They also find that the magnitude of misreporting is larger in firms offering more equity-based compensation, and that options provide the
strongest incentive to misreport because the risk of detection is lower than with other forms of compensation. In addition to the above field studies, Bruner, McKee, and Santore (2008) find experimental evidence that the amount of effort exerted and fraud committed are positively correlated with the level of equity compensation.\footnote{Other empirical research includes work by Bergstresser and Philippon (Forthcoming), who find that a greater share of a CEO’s compensation derived from stock and option holdings is correlated with a greater propensity to engage in earnings management. Ke (2005) also finds that managers with exercisable stock options and unrestricted stock have a greater propensity for earnings manipulation. Cheng and Warfield (2003) find a significant positive relationship between stock-based compensation and an abnormal tendency to meet or barely exceed analysts’ forecasts. Richardson, Tuna, and Wu (2003) also find that executives at restating firms have more equity-based compensation, but they do not distinguish between options and straight equity.}

2. THE MODEL

A team of \( N \) risk-neutral managers is hired to provide effort to increase the value of the firm. The size of the team depends on the specifics of the firm and is taken as exogenous. The team is compensated with both salary and equity. The managers are assumed identical so each receives the same contract. Let \( s \) represent the total salary and \( \alpha \) represent the total share of equity paid to the team. It follows that the salary and share of equity given to each manager are \( \frac{s}{N} \) and \( \frac{\alpha}{N} \), respectively. As we do not focus on the owner’s problem or the optimal contract, we assume for simplicity that the contract exceeds the reservation utility of each manager in the team.

Each manager chooses an unobservable level of effort, \( e_i \), which costs \( C(e_i) \). The true value of the firm is determined by the combined effort of all managers, \( E = \sum_{i=1}^{N} e_i \) and a stochastic component, \( \mu \), which is normally distributed with \( E(\mu) = 0 \) and variance \( \sigma^2 \). Specifically, the true value of the firm is \( V(E, \mu) = v(E) + \mu \), where \( v'(E) > 0 \) and \( v''(E) < 0 \). Once the true value of the firm is determined, this value is learned by the managers, but no one else.

With the true value determined, the managers issue a report on the firm’s value to the market, \( V^R = V(E, \mu) - s + F \), where \( F \) is a non-negative fraudulent inflation of the true value. While the

\[\textbf{\footnotesize 11}\]
contracts paid to the managers are observable, both \(V(E, \mu)\) and \(F\) are not. Each manager contributes to this report of the firm’s value, and we will refer to each manager’s decision as a choice of fraud, \(f^i\).

These choices are aggregated with the fraud of other managers, \(F_{-i} = \sum_{j \neq i} f^j\) into total fraud, \(F = \sum_{i=1}^{N} f^i\). One interpretation is that each manager is responsible for some section of the firm, and is able to report inflated earnings for that section. The reports are then aggregated to yield a reported value for the firm.

After the report is made, the managers sell their equity stakes at the market value.\(^{12}\) Afterwards, the true value is learned by the market with certainty, causing damage to the firm, \(D(F)\), commensurate with the difference between the reported value and true value. The market is assumed to anticipate a level of fraud, \(F^e\), and adjusts the market value of the firm accordingly. In this case, the market value of the firm is \(V^M = V^R - F^e - D(F^e)\).

While the firm absorbs damage from the fraud with certainty, there is only a positive probability, \(\rho\), that the deviation between the reported value and true value will be determined fraudulent.\(^{13}\) Let \(X()\) denote the convex sanction function that determines the punishment to the manager if the fraud is detected or reported, where \(X' > 0, \ X'' > 0, \ \text{and} \ X'(0) = 0\). This sanction represents seizure of pay, jail time, reputation losses, etc., examples of which can be seen in several recent federal prosecutions.\(^{14}\)

All managers receive disutility from being sanctioned, but some managers receive greater disutility than others.\(^{15, 16}\) We assume that there are two types of manager and let \(t = L, H\) denote a manager’s type. A type \(L\) manager (referred to as a low) incurs disutility \(\eta_L X()\) from being sanctioned while a type \(H\)

\(^{12}\) All managers must sell their equity stakes. If this were not the case there would exist a classic adverse selection problem (Akerlof, 1970); the mere fact that a manager sold his or her shares would signal that that fraud had occurred.

\(^{13}\) One can interpret \(\rho\) as the probability that authorities have sufficient evidence to build a case against the managerial team.


\(^{15}\) Schmidt (2005) provides a non-technical discussion of the merits of rewarding external whistle-blowing and of the impacts encouragement of whistle-blowing has on different types of managers.

\(^{16}\) For example, a young manager may suffer greater future pay losses from establishing a criminal record early in his or her career than an older manager nearing retirement. Furthermore, as discussed by Cooter and Porat (2001), behavior that is inconsistent with social norms can lead to nonlegal sanctions that can vary across individuals.
manager (referred to as a *high*) incurs disutility $\eta_H X()$ from being sanctioned, where $\eta_H > \eta_L$. There are $N_L \geq 1$ lows and $N_H \geq 1$ highs in the team, so that $N_L + N_H = N$.

In the case of a managerial team, where all members are working for a single entity and can be considered negligent if fraud occurs, team punishment is assumed. We assume that the sanctions applied to each of the managers is determined by the average level of fraud, $\frac{F}{N}$.\(^{17}\) This is consistent with the notion that, regardless of each manager’s individual contribution to fraud, each member of the team was complicit and should be held accountable. Hence, if the managerial team is found to have committed fraud, the utility cost for a manager of type $t$ is $\eta_t X \left( \frac{F}{N} \right)$.

One or more managers may choose to report fraud to external authorities before the report is issued to the market; that is, a manager may blow the whistle. If whistle-blowing occurs, fraud is reduced to zero, and managers are held responsible for their individual choices of (attempted) fraud, regardless of the fraud chosen by others. Consistent with statutes such as Sarbanes-Oxley, managers are able to protect themselves from being punished for the actions of others if they find the risk to be too great. We allow for the possibility that, in exchange for revealing fraud, managers receive some degree of immunity from prosecution in addition to the protections from retaliation stated in the Act itself. To model this partial immunity, we let $\theta$, represent the proportional reduction in penalty for the whistleblower, where $0 \leq \theta \leq 1$. To summarize, the utility cost for manager $i$ of type $t$ who blows the whistle is $(1 - \theta) \eta_t X(f_t^i)$, and the utility cost for manager $i$ of type $t$ when another manager blows the whistle is $\eta_t X(f_t^i)$.

It follows from the above discussion that the expected utility for manager $i$ of type $t$ ultimately depends on $i$’s choice of effort and fraud, the effort and fraud choices of other team members, and whether someone blows the whistle.

---

\(^{17}\) This avoids a Becker-type outcome where there is less fraud simply because the sanction is greater.
Team punishment provides an incentive to monitor and report on the behavior of other members of the team. If punishment were not applied to the entire team, a manager would never blow the whistle in equilibrium. Indeed, not blowing the whistle and not committing fraud dominates blowing the whistle since in the former case the manager may receive an inflated price for his or her equity holdings and has no chance of being sanctioned. When the choice of fraud is made individually but the entire team is punished, it is possible that total fraud could reach levels unacceptable to some managers. Faced with this possibility, with no mitigating instrument, highs may choose not to play the game, and an adverse selection problem would result. Here, whistle-blowing serves as a mitigating instrument.

To summarize, the timing of events are as follows:

**Stage 1:** Each manager simultaneously and individually chooses a level of effort.

**Stage 2:** The true value of the firm, \( V(E, \mu) - s \), is realized and observed by all managers.

**Stage 3:** Each manager simultaneously and individually chooses a level of fraud.

**Stage 4:** Each manager observes the fraud chosen by the other managers and decides whether or not to engage in “whistle-blowing”, revealing to authorities that fraud is about to occur.

**Stage 5:** Each manager sells his or her equity stake in the firm, regardless of whether or not whistle-blowing has occurred.

**Stage 6:** There is a positive probability, \( \rho \), that any difference in reported and true values is determined to be fraudulent, in which case all managers are punished based on the total amount of fraud. (Note: If whistle-blowing has occurred then the reported value equals the true value, so no sanctions are applied.) Regardless of whether or not the difference between the reported and true value is proven fraudulent, the value of the firm is reduced by \( D(F) \).
3. RESULTS

In our analysis, we focus on a symmetric pure-strategy equilibrium in which all managers of the same type choose the same level of fraud. There is a multiplicity of equilibria, but the total amount of fraud is the same in each, so restricting our discussion to the symmetric equilibrium does not alter the qualitative nature of our results.

We begin by solving the game backwards and determining when it is optimal for a manager to blow the whistle. If a manager blows the whistle, he or she sells the stock at the true value and faces a possibly-reduced sanction for the fraud he or she committed. If a manager does not blow the whistle, he or she benefits from selling the stock at the artificially inflated price, but has a positive probability of facing a sanction based on the average amount of fraud. Comparing the payoffs given in equation 1, we find that it is optimal for a manager of type \( t \in L(\text{low}), H(\text{high}) \) to blow the whistle if, for a given level of total fraud, the following is satisfied for any manager in Stage 4:

\[
\frac{\alpha}{N} F - \rho \eta_t X\left(\frac{F}{N}\right) \leq -\eta_t (1 - \theta) X\left(f_t^*\right), \quad t = L, H \tag{2}
\]

If (2) holds with equality, the manager is indifferent between blowing the whistle and not. In such circumstances, we assume that the manager does not blow the whistle.

**Lemma 1:** Whistle-blowing does not occur in equilibrium.

If the total level of fraud is such that it is optimal for a manager to blow the whistle, any manager who committed positive fraud receives negative net benefits. Thus, it can never be an equilibrium for managers to commit fraud if doing so will cause some manager to blow the whistle. Nevertheless, as shown below, the mere threat that another member of the team will go to authorities may be sufficient to reduce the total level of fraud.

Given that we can rule out any equilibrium in which whistle-blowing occurs, the next step is to determine the equilibrium level of fraud. Although the fear that someone may blow the whistle can affect the equilibrium, it is convenient to first determine the preferred level of fraud assuming whistle-blowing
were not possible. *Highs* and *lows* have different preferred levels of total fraud, which are the solutions to the following maximization problem.

$$\max_{F} EU_t = \frac{s}{N} + \alpha \left( v(E) - s + F - F^e - D(F^e) \right) - C(e_i) - \rho \eta_t X \left( \frac{F}{N} \right), \quad t = L, H$$

The first order condition for an interior solution is:

$$\frac{\alpha}{N} - \frac{\rho \eta_t}{N} X \left( \frac{F}{N} \right) = 0, \quad t = L, H \quad (3)$$

The assumption that $X'(0) = 0$ rules out a corner solution. The second-order condition for maximization is also satisfied:

$$-\frac{\rho \eta_t}{N^2} X'' \left( \frac{F}{N} \right) < 0, \quad t = L, H \quad (4)$$

Let $P^t > 0$ be implicitly defined by (3), where the arguments are suppressed. This value represents the preferred level of total fraud for a manager of type $t$, assuming that whistle-blowing does not occur. This preferred level is that which maximizes the expected net benefits from fraud, which we define as the expected utility derived from fraud a manager receives given that, as Lemma 1 indicates, whistle-blowing does not occur. With the preferred level of fraud determined for each type, we now determine the level of fraud which will cause whistle-blowing for a given manager.

As shown, in equilibrium no manager will blow the whistle, yet the threat of whistle-blowing may alter the equilibrium level of fraud. The level of fraud that will be tolerated by manager $i$ depends on the manager’s type and the level of fraud that he or she has chosen. Define $W^t(f^i)$ as the maximum total amount of fraud that does not cause manager $i$ of type $t$ to blow the whistle when $i$ has chosen $f^i$. From (2), $W^t(f^i)$ is the non-zero amount of fraud, $F$, that solves:

$$\frac{\alpha}{N} F - \rho \eta_t X \left( \frac{F}{N} \right) + \eta_t (1 - \theta) X \left( f^i \right) = 0, \quad t = L, H \quad (5)$$

Figure 1 provides a graphical representation of the preferred and whistle-blowing levels of fraud.
The preferred level of fraud, $P_t^*$, maximizes the increase in the value of a manager’s equity due to the fraud, $\frac{a}{N}F$, less the expected sanctions, $\rho \eta_t \left( \frac{F}{N} \right)$. Because punishment is levied against the entire team if no manager blows the whistle, the preferred level of fraud is unaffected by an individual manager’s contribution to that fraud. The whistle-blowing level of fraud, $W_t^* \left( f_t^i \right)$, is the non-zero level of fraud at which a manager is indifferent about blowing the whistle (i.e. when the expected net benefits if whistle-blowing does not occur are equal to the benefit of blowing the whistle). If a manager is responsible for some of the fraud ($f_t^i > 0$), as in Figure 1, then indifference occurs at negative expected benefits as described by (5), because the manager will be punished for his or her individual contribution if whistle-blowing occurs. As should be clear from Figure 1, it is straightforward to show that $W_t^* \left( f_t^i \right)$ is increasing in $f_t^i$. If $f_t^i = 0$, then indifference occurs at the horizontal axis. Similarly, if full amnesty is granted to whistle-blowers, $\theta = 1$, indifference occurs at the horizontal axis.

**Lemma 2:** The preferred level of fraud for a given type of manager is less than the level that causes whistle-blowing, regardless of the manager’s individual contribution to total fraud. Formally, $P_t^* < W_t^* \left( f_t^i \right)$ for all $f_t^i \geq 0$.

Whistle-blowing occurs when there is too much fraud for a manager to accept the risk of external punishment, so it is not surprising that this amount must be larger than the preferred level.

**Lemma 3:** Highs prefer less fraud than lows. Formally, $P_H^* < P_L^*$.

The above lemma states the intuitive result that those managers with a higher sensitivity to sanction prefer less fraud than managers with a lower sensitivity. This follows from the fact that both types receive the same benefit from any given level of fraud, but highs receive greater punishment, if detected, than the lows.

**Lemma 4:** If the equilibrium level of fraud exceeds the preferred level of some type $t$, then all managers of type $t$ must commit zero fraud. Formally, if $F^* > P_t^*$ for some $t = L, H$, then $f_t^* = 0$. 

11
If the equilibrium level of fraud is higher than the preferred level for that type, a manager of that type would be better off decreasing his or her individual contribution to the fraud. Therefore, the equilibrium level of fraud cannot exceed the preferred level of that type unless managers of that type commit no fraud. We can now characterize the equilibrium.

**Proposition 1:** At a subgame-perfect symmetric equilibrium

(i) total fraud equals \( F^* = \min\{P^L, W^H(0)\} \);

(ii) the highs commit no fraud, \( f^*_H = 0 \);

(iii) the lows commit a positive level of fraud, \( f^*_L = \min\left\{\frac{P^L}{N_L}, \frac{W^H(0)}{N_L}\right\} \); and

(iv) no manager blows the whistle.

At equilibrium, the lows commit either their preferred level of fraud or the greatest amount that does not trigger a high to report the illegal activity. The *highs* commit no fraud, but do not blow the whistle, implying that they are no worse off than if fraud were impossible. Important to note is that the composition of the team does not matter, as long as there is at least one of each type. The equilibrium is determined by the preferred level of fraud for the *low* type and the whistle-blowing level of the *high* type, not the proportion of each type in the team. Figures 2 and 3 provide a graphical illustration for each of the two cases where the respective values of \( W^H(0) \) and \( P^L \) differ.

<FIGURE 2>

When the whistle-blowing condition for the *high* is less than the preferred level for the *low*, given that the *high* commits no fraud (Lemma 4), then the only choices of fraud acceptable to both types is the range given by \( P^H \) to \( W^H(0) \). *Lows* push the level of fraud as high as possible without inducing the highs to blow the whistle.

In Figure 3 below, the preferred level of fraud for the *low* is below the whistle-blowing condition for the *high*, so the equilibrium level of fraud is the one that is preferred by the *lows*. Both types benefit from the fraud, although the *highs* do not contribute to it.
Proposition 2: The relative magnitudes of $\eta_L$ and $\eta_H$ determine the nature of the equilibrium.

(i) If the two types are sufficiently similar, then the equilibrium fraud equals the level preferred by the lows. That is, if $\eta_H$ is sufficiently close to $\eta_L$, then $F^* = P_L$.

(ii) If the two types are sufficiently different, then the equilibrium fraud equals the level at which the highs are indifferent between blowing the whistle and not. That is, if $\eta_H$ is sufficiently large relative to $\eta_L$, then $F^* = W^H(0)$.

The previous proposition implies that whistle-blowing may not affect the equilibrium level of fraud if the managers are not sufficiently different. As illustration, assume that all managers are identical. In a symmetric equilibrium, these managers would divide their preferred level of fraud. Whistle-blowing would be entirely irrelevant. Now, assume that there are two types, but the highs are so averse to committing fraud that they receive infinite disutility from any sanction. The threat of their whistle-blowing would result in zero fraud being committed. It is the heterogeneity of managers that allows whistle-blowing to reduce the equilibrium level of fraud from that preferred by the least sensitive type.

Proposition 3: Granting amnesty to whistle-blowers does not alter the equilibrium level of fraud.

In equilibrium, highs do not commit fraud. When whistle-blowing occurs, punishment is based on the individual choice of fraud, so highs would face zero punishment if they blew the whistle. As such, there is no penalty to be reduced, so amnesty does not change their behavior. Lows are responsible for all of the fraud, which is never greater than their preferred level. So, in equilibrium, it is not optimal for the lows to blow the whistle, even with full amnesty. The conclusion is that amnesty has no effect on the equilibrium.

The previous analysis implies that in a Subgame Perfect Nash Equilibrium, the fraud choices are independent of the effort choices. Thus, while the managers are rational and anticipate the future choices of fraud, their effort choices do not affect the fraud that will be committed. Each effort choice is a best-response correspondence with the effort choices of other managers in the team. The assumptions that the
value of the firm is a concave function of team effort and that the cost of effort is a convex function of individual effort imply that choices of effort are strategic substitutes.

The manager’s problem, where \( E_{-i} = \sum_{j \neq i} e_j \) and \( F^* = \min\{P^L, W^H\} \) is as follows:

\[
\max_{e_i} \text{EU}_i = \frac{s}{N} + \frac{\alpha}{N} (v(e_i + E_{-i}) - s + F^* - F^* - D(F^*)) - C(e_i) - \rho \eta_t X \left( \frac{F^*}{N} \right), \quad t = L, H
\]

The first-order condition is

\[
\frac{\alpha}{N} v'(e_i + E_{-i}) - C'(e_i) = 0 \quad (6)
\]

The above implicitly defines best-response function of a manager, given the share of equity and effort exerted by other managers. The second order condition for maximization is satisfied.

\[
\frac{\alpha}{N} v''(e_i + E_{-i}) - C''(e_i) < 0 \quad (7)
\]

Notice that both types of managers are identical in the characteristics that determine the choice of effort. It is, therefore, not surprising that in equilibrium the managers will choose identical effort levels.

**Proposition 4:** In equilibrium, both types of manager choose the same level of effort, \( e^* \), which is implicitly defined by:

\[
\frac{\alpha}{N} v'(Ne^*) - C'(e^*) = 0 \quad (8)
\]

As discussed in the introduction, previous research focusing on single-agent models has shown that equity compensation is a double-edged sword increasing both effort and fraud. While our focus has been on a team setting, which allows us to explore the impact of Sarbanes-Oxley’s provisions on the level of fraud, the present model nevertheless yields predictions that are consistent with previous research. As shown by the next two propositions, an increase in equity compensation yields greater team effort and greater team fraud.

**Proposition 5:** An increase in equity compensation, \( \alpha \), increases the equilibrium level of team effort, \( E^* = Ne^* \).
The value of the firm depends on team effort, so effort is akin to a public good in our model. The benefits of effort accrue to all managers, but the cost of effort is borne privately. As such, each manager takes the effort choices of other managers as given and chooses an effort level such that the marginal benefit of additional effort equals the marginal cost. Increasing equity compensation increases the marginal benefit of each unit of effort, while the marginal cost remains the same, causing all managers to increase their effort.

**Proposition 6:** The equilibrium level of total fraud, $F^*$, decreases if either equity compensation, $\alpha$, decreases or the probability of detection, $\rho$, increases.

Within the present model, owners may control fraud by reducing the level of equity compensation or increasing auditing efforts so as to increase the probability that fraud will be detected. Both options have corresponding costs. As shown by Proposition 5, decreasing equity compensation implies less team effort and less overall firm value. Audits have direct costs as well as indirect costs, such as diverting managerial efforts from productive endeavors to assisting auditors. While firms may not have much incentive to detect fraud (Robson and Santore, 2008), the Sarbanes-Oxley Act attempts to increase the likelihood that fraud is detected.\(^{18}\)

### 4. CONCLUSION

The Sarbanes-Oxley Act was passed in an attempt to reduce accounting fraud and restore investor confidence in financial reporting. This paper has focused on the implications of two important provisions of the Act: protections for whistle-blowing employees and the requirement that management sign-off on financial statements. The latter provision essentially implies that a large-scale fraud requires complicity among members of the managerial team.

\(^{18}\) Sections 408 of the Sarbanes-Oxley Act identifies factors for the SEC staff and other regulators to use as prospective problems in financial reporting, and Section 404 requires that firms publicly disclose “material weaknesses” in their internal controls over financial reporting. Financial reporting practices now receive more scrutiny from outsiders, increasing the probability of detection. (Alexander and Weiss, 2007)
Our analysis suggests that these provisions can decrease the total amount of fraud committed by the team by inducing managers to act as monitoring agents. Although whistle-blowing is not predicted to occur in equilibrium, the threat of whistle-blowing restricts fraudulent activity. This indicates that Sarbanes-Oxley could be effective in reducing fraud without costly investment in internal monitoring if its provisions are adequately enforced.

Our findings are consistent with previous theoretical and empirical literature on the relationship between managerial fraud and equity compensation. In our team setting, greater equity compensation induces greater effect from firm managers but also encourages greater misreporting of earnings. By formally modeling whistle-blowing in teams, we are able to generalize the main results of previously studied single-agent models while examining the likely effects of policy such as Sarbanes-Oxley.

The analysis also sheds light on the importance of a managerial team’s composition. The threat of whistle-blowing has an effect on the level of fraud only when there is sufficient heterogeneity in the team. Since one can view each managerial hire as a random draw from a pool of applicants, a larger team is more likely to have at least one manager who is highly sensitive to sanctions and thus willing to blow the whistle for even minor infractions.

Finally, it is interesting to note that formal guidelines for reduced sentencing are absent from the Sarbanes-Oxley Act. Although it is expected that prosecutors would grant some degree of leniency to executives who decide to cooperate with the government, such reductions in penalty have not been codified in law as they were in the case of whistle-blowing on price-fixing cartels. Our results predict that such amnesty provisions will have no effect on behavior because those managers who are most likely to blow the whistle are not those managers who are responsible for the fraud.

---

19 The US Antitrust Division created a “Leniency Policy” in 1978. Basically, a cartelist which reports the existence of a cartel to regulators is granted immunity from punishment. Apesteguia, Dufwenberg, and Selten (2007) describe and experimentally test the effectiveness of such a policy, and also provide a concise review of the relevant literature.
REFERENCES


Appendix:

Proof of Lemma 1: Suppose, to the contrary, that $F^* > 0$ is the equilibrium level of fraud and that some manager finds it optimal to blow the whistle. In this case, a manager who chooses a positive level of fraud receives no benefit from fraud and pays a penalty with certainty (recall equation 1). Thus, any manager committing positive fraud would have a profitable deviation because this manager would be better off committing zero fraud, contradicting the assumption of equilibrium.

Proof of Lemma 2: Define $Z(F) = \frac{a}{N}F - \eta_t(\frac{F}{N})$ and rewrite (5) as

$$Z(F) + \eta_t(1 - \theta)X(f^*_t) = 0, \quad t = L, H \quad (A1)$$

Now observe that the first and second derivatives of $Z(F)$ are the expressions in (2) and (3). It follows that $Z(F)$ is concave, achieves a maximum at $F = P^t$, and is positive at all $F \leq P^t$. Since the term $\eta_t(1 - \theta)X(f^*_t)$ in (5') is non-negative at all $f^*_t \geq 0$, the solution to (A1) must occur at $F > P^t$ for all $f^*_t \geq 0$.

Proof of Lemma 3: Implicit differentiation of (3) yields

$$\frac{\partial P^t}{\partial \eta_t} = \frac{-NX'(\frac{F}{N})}{\eta_tX''(\frac{F}{N})} < 0, \quad t = L, H \quad (A2)$$

where the sign follows from $X' > 0, X'' > 0$. Since $\eta_H > \eta_L$, we have $P^H < P^L$.

Proof of Lemma 4: Suppose, to the contrary, that for some type $t$ we have $F^* > P^t$ and $f^*_t > 0$. This cannot be an equilibrium because a manager of type $t$ could increase his or her payoff by choosing a lower level of fraud.

The next two Lemmas are not stated in the text but are used by Proposition 1.

Lemma 5: There does not exist an equilibrium in which $F^* < \min\{P^L, W^H(0)\}$.

Proof of Lemma 5: Suppose, to the contrary, that there exists an equilibrium in which $F^* < \min\{P^L, W^H(0)\}$. First, observe that for any $F' < \min\{P^L, W^H(0)\}$ neither type wishes to blow the whistle since $F' < W^H(0) \leq W^H(f^*_H)$ and $F' < P^L < W^L(0) \leq W^L(f^*_L)$. Second, observe that at any $F^* < P^L$, a low would prefer to increase his or her chosen level of fraud to $f' = f^*_L + \epsilon$ where $0 < \epsilon \leq P^L - F^*$ as long as doing so does not cause some manager to blow the whistle. However, it has already been shown that no manager will blow the whistle as long as $F^* + \epsilon < \min\{P^L, W^H(0)\}$. It follows that
a low could increase his or her payoff by increasing his or her level of fraud by some small \( \varepsilon \) without inducing whistle-blowing, contradicting the assumption of equilibrium. ■

**Lemma 6:** There does not exist an equilibrium in which \( F^* > \min \{ P^L, W^H(0) \} \).

**Proof of Lemma 6:** First, consider the case in which \( P^L \leq W^H(0) \) and suppose, to the contrary, that there exists an equilibrium in which \( F^* > P^L \). By Lemmas 3 and 4 we must have \( f^*_L = f^*_H = 0 \), which contradicts the supposition that \( F^* > P^L \). Second, consider the case in which \( P^L > W^H(0) \) and suppose, to the contrary, that there exists an equilibrium in which \( F^* > W^H(0) \). By Lemma 2 we have \( F^* > P^H \) which, along with Lemma 4, implies \( f^*_H = 0 \). Thus, a high who has not committed fraud would blow the whistle since \( F^* > W^H(0) \). However, from Lemma 1 we know that we cannot have an equilibrium in which a manager blows the whistle. ■

**Proof of Proposition 1:** By Lemma 5 we have \( F^* \leq P^L \). Lemmas 2 and 6 imply \( P^H < F^* \). By Lemma 4, \( P^H < F^* \) implies \( f^*_L > 0 \). Therefore, we must have \( f^*_L > 0 \). There are now two mutually exclusive cases to consider.

The first case is when \( P^L \leq W^H(0) \). Here we have \( F^* = P^L \), \( f^*_H = 0 \) and \( f^*_L = \frac{P^L}{N_L} \). The highs do not have a profitable deviation because \( F^* > P^H \), implying that any increase in fraud would necessarily reduce a high’s utility. Lows do not have a profitable deviation because \( P^L \) is defined as the level of fraud that maximizes their utility, so any change in fraud causes a decrease in utility for a low. (2) is not satisfied strictly for either type so whistle-blowing does not occur. Thus, each manager’s actions are a best-response to those of other managers.

The second case is when \( P^L > W^H(0) \). Here we have \( F^* = W^H(0) \), \( f^*_H = 0 \) and \( f^*_L = \frac{W^H(0)}{N_L} \). The highs do not have a profitable deviation because \( F^* > P^H \), implying that any increase in fraud would reduce utility for the highs. A low does not have a profitable deviation to reduce fraud because the level of fraud is below the preferred level for a low, nor does a low have a profitable deviation to increase fraud because that would cause whistle-blowing. (2) is not satisfied strictly for either type, so whistle-blowing does not occur. Thus each manager’s actions are a best-response to those of other managers. ■

**Proof of Proposition 2:** (i) To start, hold \( \eta_H \) fixed, so that \( W^H(0) \) remains fixed. By Proposition 1, \( W^H(0) > P^L \) implies \( F^* = P^L \), so we need to show that for \( \eta_L \) sufficiently close to \( \eta_H \) we must have \( W^H(0) > P^L \). First, equation (3) allows us to write \( P^L \) as a function of \( \eta_L \) so we can write \( P^L(\eta_L) \). It follows by inspection that \( P^L(\eta_H) = P^H \). By Lemma 2 and the fact that \( P^L(\eta_H) = P^H \) we have
\( P^L(\eta_H) < W^H(0) \). It thus follows from the fact that \( W^H(0) \) is independent of \( \eta_L \) and the continuity of \( P^L() \) that for \( \eta_L \) sufficiently close to \( \eta_H \) we must have \( P^L(\eta_L) < W^H(0) \).

(ii) To start, hold \( \eta_L \) fixed, so that \( P^L \) remains fixed. By Proposition 1, \( W^H(0) < P^L \) implies \( F^* = W^H(0) \), so it is sufficient to show that as \( \eta_H \to \infty \) we have \( W^H(0) \to 0 \). By definition, \( W^H(0) \) is implicitly defined by (5) when \( f^H_0 = 0 \), which can be rearranged to yield

\[
\frac{a}{\rho \, N \, \eta_H} = \frac{X \left( \frac{W^H(0)}{N} \right)}{W^H(0)}.
\]

The limit of the left hand side of this expression goes to zero as \( \eta_H \to \infty \), so we need to show that as \( W^H(0) \to 0 \) the limit of the right hand side of the expression goes to zero. Since \( X(0) = 0 \), we need to apply L’Hospital’s rule. Taking the derivative of the numerator and denominator with respect to \( W^H(0) \) yields

\[
\frac{X \left( \frac{W^H(0)}{N} \right)}{\frac{W^H(0)}{N}}.
\]

Now using the fact that \( X'(0) = 0 \), it follows that as \( W^H(0) \to 0 \) the right hand side of the expression goes to zero. It follows that for any \( \eta_L > 0 \) there exists a \( \eta_H \) such that \( W^H(0) < P^L \).

**Proof of Proposition 3:** From Proposition 1, we know that \( F^* = \text{Min}\{P^L, W^H(0)\} \). It is therefore sufficient to show that neither \( P^L \) nor \( W^H(0) \) depend on \( \theta \). Equation (3) implies that \( P^L \) is independent of \( \theta \). From (5) it follows that \( W^H(0) \) solves

\[
\frac{a}{\rho \, N \, \eta_H} \sum_{i=1}^{N} F - \rho \, \eta_H \, X \left( \frac{F^*}{N} \right) = 0,
\]

which is independent of \( \theta \).

**Proof of Proposition 4:** That (8) must be satisfied for every manager follows from the discussion in the text. Now observe that the marginal product of effort for any given manager, \( \frac{\rho}{N} \, v'(e_i + E_{-i}) \), depends only on team effort. Therefore, in equilibrium, we must have \( C'(e_i^*) = C'(e_j^*) \) which implies \( e_i^* = e_j^* \) since the marginal cost of effort is strictly increasing.

**Proof of Proposition 5:** Differentiating (8) we can calculate the change in the equilibrium effort for one manager

\[
\frac{d e^*}{d a} = -\frac{v'(N e^*)}{\alpha v'(N e^*) - C''(e^*)}.
\]

(A3)

So the change in team effort is

\[
\frac{d (N e^*)}{d a} = -\frac{v'(N e^*)}{\alpha v'(N e^*) - C''(e^*)} > 0
\]

(A4)

where the sign of the above follows from \( v' > 0, v'' < 0, \text{ and } C'' > 0 \).

**Proof of Proposition 6:** We first show the effect of a change in equity compensation. From Proposition 1, we know that \( F^* = \text{Min}\{P^L, W^H(0)\} \). We show that increasing equity compensation will increase both the preferred level of the fraud for the lows and the level that triggers whistle-blowing.
To calculate the change in $W^H(0)$, the whistle-blowing condition resulting from a change in equity, differentiate (5) evaluated at $f_H^* = 0$ with respect to $F$ and $\alpha$ to get:

$$\frac{\partial W^H(0)}{\partial \alpha} = \frac{-W^H(0)}{N} \frac{1}{\alpha - \frac{\eta_H \rho}{N} X \left(\frac{W^H(0)}{N}\right)}$$  \hspace{1cm} (A5)

Using arguments similar to those in the proof of Lemma 2, it is straightforward to show that the denominator of (A5) is negative. It follows that $\frac{\partial W^H(0)}{\partial \alpha} > 0$.

To calculate the change in $P^L$ resulting from a change in the share of equity, differentiate (3) with respect to $F$ and $\alpha$ to get:

$$\frac{\partial P^L}{\partial \alpha} = \frac{N}{\rho \eta_L X^\prime(F)} > 0$$  \hspace{1cm} (A6)

where the sign follows immediately from $X^\prime > 0$.

We next show the effect of a change in the probability of detection. From Proposition 1, we know that $F^* = \text{Min}\{P^L, W^H(0)\}$. We show that increasing the probability of detection will decrease both the preferred level of the fraud for the lows and the level that triggers whistle-blowing.

Implicit differentiation of (3) yields

$$\frac{\partial P^L}{\partial \rho} = \frac{-NX\left(F\right)}{\rho X^\prime\left(F\right)} < 0, \quad t = L, H$$  \hspace{1cm} (A7)

where the sign follows from $X^\prime > 0, X^\prime > 0$.

To calculate the change in $W^H(0)$, resulting from a change in the probability of detection, differentiate (5) evaluated at $f_H^* = 0$ with respect to $F$ and $\rho$ to get:

$$\frac{\partial W^H(0)}{\partial \rho} = \frac{\eta_H X \left(\frac{W^H(0)}{N}\right)}{\alpha - \frac{\eta_H \rho}{N} X \left(\frac{W^H(0)}{N}\right)}$$  \hspace{1cm} (A8)

Using arguments similar to those in the proof of Lemma 2, it is straightforward to show that the denominator of (A8) is negative. It follows that $\frac{\partial W^H(0)}{\partial \rho} < 0$.  \hspace{1cm} ■
Figure 1: The preferred and whistle-blowing levels of fraud for a manager who commits a positive level of fraud and where full-amnesty from punishment is not granted. (i.e. $f_t^i > 0, \theta < 1$)

Expected Wealth

\[
\text{NET BENEFITS TO MANAGER} = \frac{\alpha}{N} p - \rho \eta t \left( \frac{F}{N} \right)
\]

\[
W_t^c(f_t^i) = -\eta t (1 - \theta) X(f_t^i)
\]
Figure 2: The expected benefit of fraud to each type of manager when the managers are “sufficiently different.” (i.e. $W^H(0) < P^L$)

Expected Wealth

\[
egin{align*}
    & \text{NET BENEFITS TO LOW SENSITIVITY MANAGER} \quad \frac{\alpha}{N}F - \rho \eta X \left( \frac{F}{N} \right) \\
    & \text{NET BENEFITS TO HIGH SENSITIVITY MANAGER} \quad \frac{\alpha}{N}F - \rho \eta X \left( \frac{F}{N} \right)
\end{align*}
\]
Figure 3: The expected benefit of fraud to each type of manager when the managers are “sufficiently similar.” (i.e. $W^H(0) > P^L$)

Expected Wealth

\[
\begin{align*}
\text{NET BENEFITS TO LOW SENSITIVITY MANAGER} & : \quad \frac{\alpha}{N} P - \rho \eta X \left( \frac{F}{N} \right) \\
\text{NET BENEFITS TO HIGH SENSITIVITY MANAGER} & : \quad \frac{\alpha}{N} P - \rho \eta X \left( \frac{F}{N} \right)
\end{align*}
\]