

# **Hand in the Cookie Jar: An Experimental Investigation of Equity-based Compensation and Managerial Fraud**

David Bruner\*  
Michael McKee\*\* (corresponding author)  
Rudy Santore\*\*\*

**RH: Equity-based Compensation and Managerial Fraud**

\* Department of Economics, University of Calgary, Calgary, AB, T2N 4N1, Canada. Phone: 757-651-3540

\*\* Department of Economics, Walker College of Business, Appalachian State University, Boone, NC 28608, USA. Phone: 828-262-6080

\*\*\* Department of Economics, College of Business Administration, University of Tennessee, Knoxville, TN, 37996, USA. Phone: 865-974-6081

JEL classification: G34, C91

An earlier version of this paper was presented at the Economic Science Association meeting in Montreal, Canada, June 2005. A later version was presented at Appalachian State University. We thank several participants, Tim Perri, and two referees for helpful comments. Of course, all errors are our responsibility. The J. Fred Holly Endowment to the Department of Economics at the University of Tennessee provided funds for this research.

**Hand in the Cookie Jar: An Experimental Investigation of  
Equity-based Compensation and Managerial Fraud:**

**Abstract**

The use of equity-based compensation is an increasingly popular means by which to align the incentives of top management with that of the shareholders. However, recent theoretical and empirical research suggests that the use of equity-based compensation has the unintended consequence of creating the incentive to commit managerial fraud of the type being reported in the press. This paper reports experimental evidence showing that the amount of fraud committed by subjects is positively correlated with the level of equity, as is the level of effort. As well, the amount of fraud that is committed is negatively correlated with the probability of detection and subjects' risk aversion. The experimental design permits the identification of causal relations in the directions just noted.

## I. Introduction

It has long been recognized that the managers and owners of a firm have different incentives. The owner seeks to maximize the value of the firm while the manager may derive utility from additional factors such as the firm's market share, total output and, total employment. Equity-based compensation has been an increasingly popular means by which to align the incentives of top management with those of the shareholders. By providing a share in the ownership of the firm, equity provides the manager with a greater incentive to maximize the value of the firm. However, recent theoretical and empirical work suggests that the use of equity-based compensation has the (unintended, presumably) consequence of creating an incentive to commit fraud. Management's ability to manipulate information regarding the firm's actual performance raises the possibility that *reported* output (profits) will be overstated. This paper provides behavioral evidence, gleaned from a laboratory experiment, that increasing the level of equity causes both the level of effort and the amount of fraud to increase.

Several recent empirical studies have cited the increasing use of equity-based compensation for top-level executives (Itner, Lambert, and Larcker (2003), Anderson, Banker, and Ravindran (2000), Murphy (2003), Hall and Murphy (2002), and Hall (2003)). In 1984, for example, stocks and options comprised less than one percent of total CEO compensation for the median firm for U.S. publicly traded corporations. By 2001 stocks and options accounted for nearly two-thirds of total executive compensation for the median firm. This phenomenon is even more pronounced in "new economy" firms, defined as companies in the computer, software, Internet, telecommunications, and networking industries. Fama (1980) alludes to this phenomenon in arguing that it is the market for executive labor that demands the use of performance-based compensation. It is ironic that the very market creating the incentive to use

equity-based pay may be a victim of the incentive equity-based pay creates to improve accounting and financial statements, fraudulently if necessary. Erickson, Hanlon, and Maydew (2006), Johnson, Ryan, and Tian (2006), and Denis, Hanouna, and Sarin (2006) find that executives in firms *accused* of corporate malfeasance relied significantly more on equity-based compensation than those in firms that had not been accused of fraud. Furthermore, Chen et al (2006) find evidence that weaker corporate governance, as measured by board characteristics, are associated with a higher incidence of fraud.<sup>1</sup>

Recent theoretical models emphasize management's ability to manipulate the reported earnings of the firm. Goldman and Slezak (2006) and Robison and Santore (2006) derive agency models in which a key element is the agent's ability to provide false information to the principal concerning the outcome of the agent's effort. While these models differ in the details, in both, equity compensation provides the incentive for the agents to overstate the value of the firms they manage. Thus, increasing equity compensation is predicted to increase managerial effort as well as fraudulent reporting, though this latter effect will be damped by increased auditing (enforcement) and sanctions for fraudulent activity. The purpose of this paper is to empirically test of these theoretical predictions with behavioral evidence from a laboratory experiment.

There are obvious limitations to the use of laboratory investigations of managerial malfeasance.<sup>2</sup> Despite the many insights of the empirical literature utilizing field data, there are several issues that are difficult to address with such data. While the component of executive compensation that is based on the equity value of the firm is public knowledge, the effective enforcement (audit) effort is not. Further, absent an explicit policy intervention, the field data typically do not contain specific changes in the enforcement levels. The theoretical predictions of management behavior are often predicated on the fact that the managers are fully aware of the

probability of an audit by a regulatory agency and of the effectiveness of such audits but, as we have noted, there is often considerable uncertainty and the analyst working with field data must make inferences of managements' perceptions of the regulatory processes. Of course, with field data by necessity one can only measure detected fraud. Finally, the reactions to changes in equity compensation levels and to the probability of fraud being detected depend on individual risk attitudes, which are not easily observed in the field.<sup>3</sup>

The laboratory offers the researchers considerable control via the construction of the institution and the use of induced values (Smith, 1982). This control affords us an opportunity to test the predictions of the recent theoretical models of managerial malfeasance through varying parameters predicted to affect the level of such malfeasance. In the controlled environment of the lab we are able to collect data on the actual effort and fraud choices of human subjects, and observe how these choices are affected by a change in the level of equity-based compensation and in the likelihood of fraud detection. Unlike the field data, the laboratory allows us to observe the amount of fraud committed when it goes undetected. Also, by manipulating a single variable and holding all other factors constant, we are able to observe causation rather than simply correlation.<sup>4</sup> Our design introduces orthogonal variation in the equity-based compensation and the probability that fraud will be detected. Further, our design allows us to control for risk attitudes since we are able to elicit individual risk attitudes over the domain of the payoffs provided in the effort and fraud decision setting.

## **II. A Model of Managerial Behavior**

Goldman and Sle Zack (2006) and Robison and Santore (2006) provide the basis for the following theoretical discussion of the effects of equity-based pay and monitoring on the amount

of fraud that is committed. In these models, the manager is compensated via a two-part contract,  $(w, \alpha)$ , where  $w$  denotes the manager's salary income and  $\alpha$  denotes the percentage of total firm equity given to the manager. For her part, the manager must make two decisions; the level of effort and, then, the value of the firm to report to the market.

More precisely, the manager first must choose an effort level,  $L$ , which adds value to the firm. There are diminishing marginal returns to effort such that the value of the firm,  $g(L)$ , is a strictly concave function of the amount of effort the manager contributes (i.e.  $g'(L) > 0$ ,  $g''(L) < 0$ ). Providing value-adding effort is costly to the manager and therefore reduces the manager's (certain) salary based income. In the following discussion the cost of effort is normalized; each unit of effort costs the manager one unit of salary income. The manager's choice of effort determines the true value of the firm, but this value is observed by the manager only; the owners do not observe the true value of the firm.

After the manager chooses effort, she must choose the value of the firm to report to the shareholders. Any value in excess of the true value reported by the manager is considered fraud. Thus, the reported value of the firm is the true value of the firm plus any additional value the manager chooses to report,  $g(L) + F$ , where  $F$  is the amount of fraud committed by the manager.<sup>5</sup> The *potential* to commit fraud is sufficient to generate a reaction from the market. The market rationally expects some level of fraud,  $F^e$ .<sup>6</sup> It is not costless for the manager to defraud the shareholders. There is a known probability,  $p$ , that the manager will be caught committing fraud and sanctions,  $s(F)$  will be imposed on the manager. The sanctions function is increasing and strictly convex in the amount of fraud (i.e.,  $s'(F) > 0$ , and  $s''(F) > 0$ ).

The manager's preferences over potentially random distributions of wealth are given by the mean-variance utility function shown in equation 1,

$$EU_M = E(W_M) - r_M \sigma_{wm}^2, \quad (1)$$

where  $W_M$  is the manager's wealth,  $\sigma_{wm}^2$  is the variance of the manager's wealth, and  $r_M \geq 0$  is a risk aversion parameter. It is straightforward to calculate the variance of  $s(F)$ , which equals the variance in the manager's wealth.<sup>7</sup>

$$\sigma_s^2(F) = p[s - ps]^2 + (1-p)[0 - ps]^2 = p(1-p)s^2$$

Recalling that  $s''(F) > 0$ , it follows that

$$\frac{d\sigma_s^2(F)}{dF} = 2p(1-p)s(F)s'(F) > 0$$

$$\frac{d^2\sigma_s^2(F)}{dF^2} = 2p(1-p)(s(F)s''(F) + (s'(F))^2) > 0$$

Choosing a greater value of fraud increases the variance of the manager's wealth. So the cost of choosing greater fraud has two costs for the risk-averse manager: the first is the increase in the expected sanction and the second is the increase in risk.

We solve for the manager's optimal choices backwards since a rational manager will anticipate her future choice of fraud when she chooses effort. Once effort has been chosen, the manager must choose a level of fraud. The manager's objective function is given by:

$$\underset{F \geq 0}{Max} \quad \alpha(g(L) + F - F^e) - L - ps(F) - r_M \sigma_s^2(F) \quad (2)$$

At an interior solution to equation 2, the first order condition with respect to  $F$  is:

$$\alpha - ps'(F) - r_M \frac{d\sigma_s^2(F)}{dF} = 0 \quad (3)$$

Given that  $s(F)$  is convex the sufficient second-order condition is satisfied:

$$-ps''(F) - r_M \frac{d^2\sigma_s^2(F; t_F)}{dF^2} < 0$$

Equation 3 implicitly defines the optimal level of fraud,  $F^* = F(\alpha, p)$ , which is independent of the level of effort. In choosing the optimal level of effort, the manager anticipates her future choice of fraud:

$$\text{Max}_L \alpha( g(L) + F^* - F^c ) - L - ps(F^*) - r_M \sigma_S^2(F^*)$$

At an interior solution, the first order condition with respect to  $L$  is:

$$\alpha g'(L) - 1 = 0 \quad (4)$$

Equation 4 implicitly defines the optimal level of effort,  $L^* = L(\alpha)$ . The only parameter that enters the manager's choice of effort is  $\alpha$ , the percentage of equity.

As equation 5 shows effort is increasing in  $\alpha$ .

$$\frac{dL^*}{d\alpha} = \frac{-g'(L^*)}{\alpha g''(L^*)} > 0 \quad (5)$$

The optimal level of fraud is increasing in the share of the firm and decreasing in the probability of detection and degree of risk aversion:

$$\frac{\partial F^*}{\partial \alpha} = \frac{1}{ps'' + r_M \frac{d^2 \sigma_S^2(F; t_F)}{dF^2}} > 0 \quad (6)$$

$$\frac{\partial F^*}{\partial p} = \frac{-s'}{ps'' + r_M \frac{d^2 \sigma_S^2(F; t_F)}{dF^2}} < 0 \quad (7)$$

$$\frac{\partial F^*}{\partial r_M} = \frac{-\frac{d\sigma_S^2(F)}{dF}}{ps'' + r_M \frac{d^2 \sigma_S^2(F)}{dF^2}} < 0 \quad (8)$$

Equation 6 shows that as the share of the firm the manager owns increases, there is also an increased incentive to artificially inflate the firm's market value. On the other hand,

equations 7 and 8 show that the optimal amount of fraud is a decreasing function of the probability of the fraud being detected and the degree of the agent's risk aversion.

### **III. Experimental Design**

The experimental setting assigns human subjects the role of manager while the shareholder role is computerized. We implement the basic elements of the theoretical model presented above to investigate the effects of equity-based pay and monitoring on the amount of fraud committed. In the first stage of a period, subjects choose the level of effort and in the second stage subjects choose a level of fraud. Finally, the results of a random audit process are revealed to the subjects and the outcomes and payoffs for the period is summarized.

The experiment was programmed and conducted with the software Z-Tree (Fischbacher, 2007). In the experiment subjects interacted with a computer interface where the instructions were presented and their decisions were recorded. Subjects were volunteers recruited through announcements in undergraduate classes at the University of Tennessee. When they arrived at the lab, the subjects were seated in individual privacy carrels and entered all of their decisions via the computer mouse. They were not permitted to communicate with other subjects and they proceeded through the experiment at their own pace. Sessions lasted approximately 60 minutes and subjects earned an average of \$15 US dollars (actual range of earnings is \$8 to \$19) based on their decisions. After the subjects read the instructions and the experimenter answered clarifying questions, the experiment began. Subjects completed five practice periods and twenty periods for actual money.

Both the Robison and Santore (2006) and the Goldman and Slezack (2006) models are based on the assumption that the manager is risk averse. In order to control for the risk

preferences of subjects we employ a measure similar to Holt and Laury (2002).<sup>8</sup> For each of ten individual gambles subjects choose between a lottery paying either 1000 or 0 lab dollars, or a guaranteed payoff of 500 lab dollars. The gambles vary in their respective probabilities of winning the large prize.<sup>9</sup> Figure 1 presents the screen image for the gamble choice exercise.<sup>10</sup> A subject's pattern of responses results in an independent measure of risk preference. If preferences are as described in equation (1), the subject's risk aversion parameter determines the option at which they choose to switch from the guaranteed amount (Choice B) to the gamble (Choice A). Thus, the subject will choose Choice B if:

$$r_i > \frac{1000p_j - 500}{1000^2 p_j(1 - p_j)}$$

Since subjects make 10 decisions we are able to construct bounds on the implied risk aversion parameter. The ranges of the implied risk aversion parameter are given in column 2 of Table 1. The observed frequencies are given in column 4.

### **Table 1 Here**

As can be seen from Table 1, subjects are fairly symmetrically distributed around risk neutrality. Employing a random utility framework first introduced by McFadden (1974) allows us to estimate the average risk aversion parameter from equation (1). In order to account for the panel structure of our data we estimate a random effects probit model.<sup>11</sup> The estimated population average risk aversion parameter is  $-0.0002$  but is not statistically different from zero. Thus our subjects appear to be risk neutral on average but to exhibit some heterogeneity.

After entering their decisions for the gamble exercise, subjects receive written instructions, which they retain throughout the experiment, and also proceed through more thorough instructions on the computer. Here they are presented with examples of the relevant

information screens, definitions and descriptions of the information being provided on those screens, and the calculations that will determine their payoffs.

In the first stage of a period the computer displays the two-part contract to the subject; the endowment and the share. Subjects are told that the first part of their payoff for the period is simply their endowment (salary) less their contribution (effort).<sup>12</sup> The second part of their payoff is their share (equity) multiplied by their reported level of output. After observing the contract the subject must choose the level of contribution (effort) by selecting a radio button on the screen. Choice of contribution ranges from 0 to 200 in increments of 8 units.<sup>13</sup> A payoff table containing the first part of their payoff and the amount of output that will result from their choice is displayed on the screen. Figure 2 shows the subject's screen image for the first decision.

In the second stage of the experiment the computer displays the true level of output that results from their choice of contribution, their share (equity) of reported output, and the probability that they are checked (audited). If they are checked and they have reported additional output, a penalty is subtracted from the second part of their payoff. Subjects are informed that the penalty is an increasing function of additional output reported.<sup>14</sup> Another payoff table is displayed reporting the second part of their payoff and the possible penalty associated with each amount of additional output. The amount of additional output that can be reported ranges from 0 to 500 in 20 unit increments. Subjects are then asked to choose how much additional output to report. Additional output is reported by selecting a radio button on the screen. Figure 3 shows the subject's screen image for this second decision.

After the two decisions have been completed, the subjects are informed of the audit outcome and shown a summary page with the parameter values, their decisions, and their payoffs

for the round. The subjects are informed whether they were checked or not and the penalties that have been incurred as a result. The period summary screen is shown in Figure 4.

The experiment treatment structure yields a 2 X 2 factorial design resulting in four experiment treatments. The treatment parameters in each session are the value of  $\alpha_i$ , the manager's percentage equity, and the value of  $p$ , the probability of an audit. The equity parameter  $\alpha_i$  takes two levels, 30%, and 50% and the probability of an audit  $p$  also takes two values, 15% and 25%. Table 2 summarizes the design.

**Table 2 here**

The theoretical model directly lends itself to some testable hypotheses and the following hypotheses will be tested using the data from our experiments:

H1: *Effort is strictly increasing in the level of equity, ceteris paribus.* According to equation (5), the optimal choice of effort is an increasing function of the level of equity. Thus, the level of effort is predicted to increase from T1 to T3 and from T2 to T4.

H2: *Effort is independent of the probability of detection, ceteris paribus.* According to equation (4), the optimal choice of effort is independent of the probability of detection. Thus, the level of effort chosen is predicted to be the same for T1 and T2 and for T3 and T4.

H3: *Fraud is increasing in the level of equity, ceteris paribus.* According to equation (6), the optimal choice of fraud is an increasing function of the level of equity. Thus, the level of fraud is predicted to increase from T1 to T3 and from T2 to T4.

H4: *Fraud is decreasing in the probability of detection, ceteris paribus.* According to equation (7), the optimal choice of fraud is a decreasing function of the probability of detection. Thus, the level of fraud is predicted to decrease from T1 to T2 and from T3 to T4.

H5: *There is an interactive effect of equity and the probability of detection on the level of fraud chosen, ceteris paribus.* According to equation (6), the response of the optimal choice of fraud to a change in the level of equity is dependent upon the probability of detection. Thus, the change in the level of fraud chosen from T1 to T3 is predicted to be greater than the change in the level of fraud chosen from T2 to T4.<sup>15</sup>

H6: *Fraud is decreasing in the degree of risk aversion, ceteris paribus.* According to equation (8), the optimal choice of fraud is a decreasing function of the agent's degree of risk aversion. Thus, across all treatments individuals who are more risk averse are predicted to choose lower levels of fraud.

Twenty subjects participated in each treatment and they faced the decision setting for 20 periods. This is a between subject design; subjects face a single set of parameters in a given session and the subjects are only permitted to participate in one session (set of parameters).<sup>16</sup>

#### **IV. Data Analysis and Results**

Our dataset constitutes a panel of 1600-pooled observations. Observations for any given subject are correlated across periods, thus any analysis must account for such correlation. To account for the structure of the data we utilize a panel estimation approach, which accounts for the repeated observations on each individual and allows us to incorporate the potential for learning and feedback during the experiment session. The variables used in our analysis are defined in Table 3 with descriptive statistics included. The first two variables are the dependent variables in the subsequent estimations. There are four treatment dummy variables. Since we use a between subjects design all treatment parameters are time invariant.<sup>17</sup> We construct a dummy variable that indicates whether or not a subject exhibited risk averse preferences in the gamble

exercise.<sup>18</sup> There are lagged feedback variables that include a dummy indicator as to whether a subject was audited in the previous period, the amount of the penalty the subject incurred in the previous period, and the amount of accumulated earnings at the beginning of a decision period.<sup>19</sup> We also use the inverse of the decision period as a proxy for learning.<sup>20</sup>

**Table 3 here**

In Table 4 we report the results from generalized least squares estimation of the level of effort chosen. Three alternate specifications are reported here. In order to account for the panel structure of the data, all three specifications are estimated with panel specific heteroskedastic error terms. The first model is the most parsimonious, containing only treatment effects. The second model incorporates learning by including the inverse of the period (*inverse period*). The static model being tested does not account for learning effects; thus they must be controlled for in the analysis. The third model adds to the learning effects the possible feedback effects such as whether the subject was audited in the previous period (*lagged audit*), the amount of any penalties in the previous period (*lagged penalty*), and the amount of accumulated earnings in the experiment (*lagged wealth*). Again, since we are testing a static model, there are no theoretical predictions about the sign or magnitudes of these effects. However, in order to appropriately test the theoretical predictions these effects must be controlled for in the analysis. The third model is used for hypotheses tests. Thus, the estimated effort equation can be written as:

$$\begin{aligned} \text{effort} = & \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + \beta_4 T_4 + \beta_5 \text{averse} + \\ & \beta_6 \text{inverse\_period} + \beta_7 \text{lagged\_audit} + \beta_8 \text{lagged\_penalty} + \beta_9 \text{lagged\_wealth} + e_i \end{aligned}$$

**Table 4 here**

Recall the first two hypotheses presented in the previous section are related to the response of effort to changes in the level of equity compensation and the probability of detection respectively. We formally test these two hypotheses using a Wald test:

H1: *Effort is strictly increasing in the level of equity.*

$$B_3 - B_1 > 0$$

$$\chi^2 = 342.40 \quad \text{p-value of 0.000}$$

$$B_4 - B_2 > 0.$$

$$\chi^2 = 332.14 \quad \text{p-value of 0.000}$$

For either level of the probability of detection, we reject the null hypothesis that effort is independent of the level of equity. Thus, we find support for our first hypothesis.

H2: *Effort is independent of the probability of detection, ceteris paribus.*

$$B_1 - B_2 = 0$$

$$\chi^2 = 16.01 \quad \text{p-value of 0.000}$$

$$B_3 - B_4 = 0$$

$$\chi^2 = 9.76 \quad \text{p-value of 0.002}$$

We reject the null hypothesis that the level of effort is the same for both levels of the probability of detection, holding equity constant. Thus, we do not find support for our second hypothesis.

We do find weak evidence of some feedback effect given the negative and significant coefficient on *Lagged Audit*.<sup>21</sup> We also find evidence of learning on the part of subjects given the positive and significant coefficient on *Inverse Period*. A dynamic model is required to incorporate the possibility of such effects and is beyond the scope of this paper.

Table 5 reports the results of our analysis of the fraud decision. The estimations utilize the random-effects tobit specification to account for subject heterogeneity and the censoring of the level of fraud.<sup>22</sup> Our approach is similar to that of the estimation of the effort equation. We start with the most parsimonious model first, which includes the treatment indicator variables along with the subject specific dummy variable indicating whether the subject exhibited risk averse preferences in the gamble choice exercise.<sup>23</sup> We then control for learning effects by

including the inverse of the period (*inverse period*) in the second model. The third model adds to the second feedback effects such as whether a subject was audited in the previous period (*lagged audit*), the amount of penalties in the previous period (*lagged penalties*), and the amount of accumulated earnings (*lagged wealth*). The third model is used for hypotheses tests. Thus, the estimated effort equation can be written as:

$$\begin{aligned} \text{fraud} = & \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + \beta_4 T_4 + \beta_5 \text{averse} + \\ & \beta_6 \text{inverse\_period} + \beta_7 \text{lagged\_audit} + \beta_8 \text{lagged\_penalty} + \beta_9 \text{lagged\_wealth} + e_i \end{aligned}$$

Recall the last four hypotheses in the previous section were concerned with the response of fraud to changes in the level of equity compensation and the probability of detection respectively. We formally test these four hypotheses using a Wald test.

**Table 5 here**

H3: *Fraud is strictly increasing in the level of equity.*

$$B_3 - B_1 > 0$$

$$\chi^2 = 14.77 \quad \text{p-value of 0.000}$$

$$B_4 - B_2 > 0.$$

$$\chi^2 = 6.51 \quad \text{p-value of 0.011}$$

For either level of the probability of detection, we reject the null hypothesis that fraud is independent of the level of equity. Thus, we find support for our third hypothesis that fraud is strictly increasing in the level of equity.

H4: *Fraud is strictly decreasing in the probability of detection, ceteris paribus.*

$$B_1 - B_2 < 0$$

$$\chi^2 = 1.69 \quad \text{p-value of 0.194}$$

$$B_4 - B_3 < 0$$

$$\chi^2 = 7.50 \quad \text{p-value of 0.006}$$

For the low level of equity we accept and for the high level of equity we reject the null hypothesis that the level of fraud is decreasing in the probability of detection. Thus, we only find weak support for our fourth hypothesis. Furthermore, our results are indicative of an interactive effect between the level of equity and the probability of detection. This will be explored further in our next hypothesis test.

H5: *There is an interactive effect between equity and the probability of detection on the level of fraud.*

$$B_3 - B_1 > B_4 - B_2$$

$$\chi^2 = 0.74 \quad \text{p-value of 0.390}$$

We accept the null hypothesis that the response of fraud to a change in the level of equity is independent of the probability of detection. Thus, we do not find support for our fifth hypothesis of an interaction effect between equity and the probability of detection.

H6: *Fraud is decreasing in risk aversion.*

$$B_5 < 0$$

$$Z = -2.70 \quad \text{p-value of 0.007}$$

We reject the null hypothesis that fraud is independent of risk aversion. Thus, we find support for our sixth hypothesis that fraud is decreasing in risk aversion.

There is no evidence that the subjects “learned” the optimal amount of fraud as the experiment progressed given the estimated coefficient on the inverse of the period (*inverse period*) is not statistically different from zero.<sup>24</sup> However, there is evidence of feedback effects. The estimated coefficient on *lagged penalty* is 2.34 and significant at the 1% significance level. This is consistent with the often-observed gambler’s fallacy. That is, after being caught and assessed penalties, subjects increased the amount of fraud they chose in the subsequent period.<sup>25</sup>

But, the estimated coefficient of *lagged audit* is not statistically different from zero. The estimated coefficient on *Lagged Wealth* is consistent with decreasing absolute risk aversion.

Table 6 presents the point predictions for both effort and fraud from the theoretical model assuming risk neutrality. Conditional means are obtained from the estimated coefficients on the treatment dummy variables from the effort and fraud equations, which allow us to partial out any subject specific effects, as well as feedback and learning effects.

**Table 6 here**

The mean level of effort for each treatment is close to the theoretical prediction. The theoretical level of effort for treatments 1 and 2 was 32 while the mean level of effort was 31.51 and 40.70 respectively. The theoretical level of effort for treatments 3 and 4 was 80 while the mean level of effort was 89.39 and 80.53 respectively. The mean level of fraud is close to the theoretical prediction for treatments 1 and 4. The theoretical level of fraud for both treatments 1 and 4 are between 200-220 and the mean levels of fraud are 188.55 and 239.35 respectively.<sup>26</sup> Treatments 2 and 3 are predicted to generate corner solutions (in fraud) of 0 and 500. However, the design permits errors from the predicted outcomes in only one direction respectively for each treatment.<sup>27</sup> It is apparent (Table 6) that mean levels of fraud for treatments 2 and 3 reflect an inward bias due to the censoring involved with our experimental design. That is, the mean predicted level of fraud for Treatment 2 (136.17) is greater than the theoretical prediction and, conversely for Treatment 3 (338.08) is below the predicted level.

Before concluding our analysis a discussion of the relative effects of our treatment variables (equity share and probability of detection) would be useful. Given we observe two distinct levels for each of our treatment variables, we can evaluate the average unconditional discrete change in fraud due to each of the treatment variables (Model 3, Table 6). For the level

of equity the average unconditional discrete change was 75.59.<sup>28</sup> For the probability of detection the average unconditional discrete change was -44.15. Thus we find for the levels of the treatment variables we observe, the increase in fraud caused by a 66 percent increase in the level of equity is almost twice as large as the decrease in fraud caused by a 66 percent increase in the probability of detection.

## **V. Conclusions**

The recent publicity of high profile cases of corporate fraud has drawn the attention of not only the media but also political agents. In response, the U.S. Congress passed the Sarbanes-Oxley Act (SOX) in an effort to restore investor confidence by appearing to attack corporate fraud. Central to the debate is the question of the root causes of fraudulent behavior. One well-received hypothesis, provided by academics as well as policy makers, is that the use of firm equity to compensate managers has had the unintended consequence of making fraud more attractive to these managers. The results of the experimental investigation reported here suggest that, while it does increase productive effort, equity compensation has the effect of increasing managerial fraud.

The multivariate analysis strongly confirms the predictions of the theoretical framework: both the level of effort chosen and the level of fraud committed are increasing in the share of output (equity) the manager subjects received. Increasing enforcement, the probability that fraud is detected, is shown to reduce the amount of fraud committed. The conditional means from the estimated equations suggest support for the theoretical model. Higher levels of equity-based compensation are correlated with higher levels of effort and as well as higher levels of fraud. Our subject managers clearly understand the incentives established by the laboratory setting.

Finally, in our experiments the level of equity compensation was a treatment variable and, therefore, exogenous. However, in the natural occurring world the level of equity compensation is determined by owners, who are influenced by external factors such as the regulatory environment and the market's potential reaction to the revelation of fraud. Given that equity compensation is expected to result in more fraud, it is natural to ask whether owners will provide greater equity compensation in response to legislation (such as SOX) that imposes greater monitoring on the managers. If so, the net effect on the actual amount of fraud may be damped. One important avenue for future research is to examine the response of the optimal compensation contract to regulation designed to decrease corporate fraud.

The usual caveat is in order regarding the use of laboratory experiments to inform our understanding of behavior in the naturally occurring world. We address the central question: does the laboratory setting provide for the necessary degree of "parallelism" to the naturally occurring world that is crucial to generalizing our experimental results beyond the setting of the lab (Smith, 1982; Plott 1987)? The experimental setting need not attempt to capture all of the variation in the naturally occurring environment, but it should sufficiently recreate the fundamental elements of the naturally occurring world if the results are to be relevant in policy debates. While our payoffs are relatively small, our experimental setting provides the computations necessary for the decision and a clear link between decisions and rewards, thus reducing the decision costs. List and Levitt (2007) describe several factors that they assert limit the experimenter's ability to generalize behavior beyond the lab. Most, if not all, of the issues raised by List and Levitt are minimized by incorporating the precepts of experimental design articulated by Smith (1982) and ensuring that payoffs are salient and that the financial rewards dominate such factors as the subjects' desires to please (or punish) the experimenter.<sup>29</sup>

## References

- Anderson, M., R. Banker, and S. Ravindran. 2000. Executive Compensation in the Information Technology Industry. *Management Science*, 46: 530 – 547.
- Andreoni, J. 1995. Cooperation in Public Goods Experiments: Kindness Versus Confusion? *The American Economic Review* 85: 891 –904.
- Bruner, D. 2006. Risk Preference Elicitation: An Empirical Analysis of Mechanism Consistency. Unpublished Paper, Department of Economics, University of Calgary.
- Chen, G., M. Firth, D.N. Gao, and O.M. Rui. 2006. Ownership Structure, Corporate Governance, and Fraud: Evidence from China. *Journal of Corporate Finance*. 12: 424 – 448.
- Croson, R. and J. Sundali. 2005. The Gambler’s Fallacy and the Hot Hand: Empirical Data from Casinos. *The Journal of Risk and Uncertainty* 30: 195 – 209.
- Davis, D.D. and C.A. Holt. 1993. *Experimental Economics*. Princeton, NJ: Princeton University Press.
- Denis, D.J., P. Hanouna, and A. Sarin. 2006. Is There a Dark Side to Incentive Compensation. *Journal of Corporate Finance* 12: 467 – 488.
- Fama, E. 1980. Agency Problems and the Theory of the Firm. *Journal of Political Economy* 88: 288 – 307.
- Fishbacher, U. 2007. Z-Tree: Zurich Toolbox for Readymade Economics Experiments. *Experimental Economics*. In press.
- Friedman, D. and S. Sunder. 1994. *Experimental Methods: A Primer for Economists*. New York, NY: Cambridge University Press.

- Erickson, M., M. Hanlon, and E. L. Maydew. 2006. Is There a Link between Executive Equity Incentives and Accounting Fraud? *Journal of Accounting Research* 44: 113-143.
- Goldman, E. and S.L. Slezak. 2006. The Economics of Fraudulent Misreporting. *Journal of Financial Economics* 80: 603 – 626.
- Goeree, J., C.A. Holt, and T.R. Palfrey. 2003. Risk Averse Behavior in Generalized Matching Pennies Games. *Games and Economic Behavior* 45: 97-113.
- Hall, B.J. 2003. Six Challenges of Equity-Based Pay Design. *Journal of Applied Corporate Finance* 15: 21-33.
- Hall, B.J. and K. Murphy. 2002. Stock Options for Undiversified Executives. *Journal of Accounting and Economics*. 33: 3-42.
- Harrison, G.W., E. Johnson, M.M. McInnes and E.E. Rustrom. 2005. Risk Aversion and Incentive Effects: Comment. *The American Economic Review* 95: 897 – 901.
- Holt, C.A. and S.K. Laury. 2002. Risk Aversion and Incentive Effects. *The American Economic Review*. 92: 1644 – 55.
- \_\_\_\_\_. 2005. Risk Aversion and Incentive Effects: New Data without Order Effects. *The American Economic Review*. 95: 902 – 904.
- Itner, C.D., R.A. Lambert, and D.F. Larcker. 2003. The Structure and Performance Consequences of Equity Grants to Employees of New Economy Firms. *Journal of Accounting and Economics*. 34: 89 – 127.
- Johnson, S.A., H.E. Ryan, and Y.S. Tian. 2006. Managerial Incentives and Corporate Fraud: The Sources of Incentives Matter. Presented at 2006 EFA Zurich Meetings and available at SSRN: <http://ssrn.com/abstract=395960>

- Khaneman, D. and A. Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*. 47: 263 – 292.
- Lange, A., J. List, and M. K. Price. 2007. Auctions with Resale when Private Values are Uncertain: Theory and Empirical Evidence. *International Economic Review*. In press.
- List, J. and S. Levitt. 2007. What do Laboratory Experiments Measuring Social Preferences Tell us About the Real World? *Journal of Economic Perspectives*. In press.
- McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*, edited by R. Zarembka. New York, NY: Academic Press, pp 105 – 142.
- Murphy, K.J. 2003. Stock-Based Pay in New Economy Firms. *Journal of Accounting and Economics*. 34: 129 – 147.
- Plott, C.R. 1987. Dimensions of Parallelism: Some Policy Applications of Experimental Methods. In *Laboratory Experimentation in Economics: Six Points of View*, edited by Alvin E. Roth. New York, NY: Cambridge University Press, pp 193-219.
- Poirier, D.J. 1980. Partial Observability in Bivariate Probit Models. *Journal of Econometrics*. 12: 209 – 217.
- Robison, D. and R. Santore. 2006. Managerial Incentives, Fraud, and Firm Value. SSRN working paper #770924.
- Smith, V.L. 1982. Microeconomic Systems as an Experimental Science. *The American Economic Review*, 72: 923 – 955.
- Smith, V.L. and J. Walker. 1993. Monetary Rewards and Decision Cost in Experimental Economics. *Economic Inquiry*, 31: 245-261.

## Endnotes

---

<sup>1</sup> The use of field data is problematic since fraud is only observed if it is detected, creating an identification problem. We employ a bivariate probit model with partial observability to address this concern. See Poirier (1980) for a discussion.

<sup>2</sup> List and Levitt (2005) discuss some limitations as they affect our ability to generalize beyond the lab. Essentially, these address the external validity of the behavior observed in the laboratory conditions and the extent to which we can achieve parallelism in the lab. We return to these issues in our concluding section.

<sup>3</sup> We are not the first to investigate corporate malfeasance in the lab. Yu (2004), for example, reports on a set of lab experiments in which the fraud takes the form of embezzling the company assets. Yu's method differs from ours in three important ways: fraud is a binary choice, not a level; subject risk attitudes are not measured; and subjects are not exposed to an audit each period. Rather, subjects are selected for an audit that covers *all* periods of the experiment; such audit being announced only at the end of the session.

<sup>4</sup> Analysis based on field data is only able to detect correlation between variables due to the potential for confounding unobservable effects.

<sup>5</sup> Contrary to the concept of earnings management in the accounting literature there is no incentive to understate the value of the firm.

<sup>6</sup> The expected level of fraud will be normalized to zero for purposes of the experiment.

<sup>7</sup> Since the sanction is the only random component of wealth, the variance in wealth is equivalent to the variance in the sanction.

---

<sup>8</sup> Employing such a mechanism to obtain a measure of subject's risk preferences is not without critics. Both Holt and Laury (2002, 2005) and Harrison, Johnson, McInnes, and Rustrom (2005) demonstrate that the measure of risk aversion obtained by such a mechanism is sensitive to the magnitude of payment. That is, scaling up real payments results in an increase in risk aversion. Furthermore, the gamble choice exercise only allows for gains while in the fraud experiment both gains and losses were possible. In the context of alternative theories for decision-making under uncertainty, such as Prospect Theory (see Khaneman and Tversky 1979), any measure for risk preference in the first setting cannot be generalized to the second. However, in our experimental setting the payment from the gamble choice exercise is comparable to that of a decision period in the actual experiment. Also, if the subjects view the guaranteed payoff as a reference point, then the gamble choice exercise involves both gains and losses. Thus, any inferred risk preference from the exercise should remain stable across the actual fraud experiment. Goeree, Holt, and Palfrey (2003) report findings that risk attitudes obtained from generalized matching pennies games are consistent with those from the Holt-Laury procedure for the same subject pool. Lange, List, and Price (2007) demonstrate the requirement to incorporate risk attitudes into the econometric specification.

<sup>9</sup> The realization of the gamble phase is not revealed to subjects until after completion of the managerial decision experiment. This prevents interaction between the decision tasks in the experiment session.

<sup>10</sup> Our Table 1 is derived in the same fashion as Table 5 in Goeree, Holt, and Palfrey (2003). We had 20 subjects exhibit inconsistent preferences by switching from the certain amount to the gamble and then back to the certain amount. We also had 4 subjects exhibit irrational preferences

---

by opting for the certain payoff in option 10, in which they could have chosen a “gamble” with a 100% chance of winning the large payoff. These subjects are coded as missing observations.

<sup>11</sup> See Bruner (2006) for details

<sup>12</sup> Neutral language was used in order to minimize possible framing effects.

<sup>13</sup> The use of radio buttons invoked a tradeoff between approaching continuous decision space and the cognitive burden associated with evaluating each choice. Thus, we divided the decision space for both effort and fraud into 25 equally spaced discrete choices.

<sup>14</sup> Again, subjects are informed via text in the instructions and can observe such in the payoff table on the decision screen.

<sup>15</sup> This is true for the values of the probability of detection,  $p$ , which we investigate since variance of sanctions is a concave function of  $p$ . A strength of employing a 2x2 experimental design is the ability to detect interaction effects between the treatment variables since the variables are perfectly orthogonal.

<sup>16</sup> The experiment required parameterization of both the  $G(L)$  and the  $S(F)$  functions. Both functions assumed the form  $\theta_i x^{\lambda_i}$ , where  $i$  is either G or S and  $x$  represents the decision variable. Specifically,  $\theta_G = 50$  and  $\theta_S = 1$  and  $\lambda_G = 0.45$  and  $\lambda_S = 1.11$ . The choice of parameters was made in order to make the objective function as concave as possible to maximize the likelihood that the gains from discovering the unique maximum surpassed the cognitive cost of searching. Risk neutrality was assumed during the choice of parameters since risk aversion requires some knowledge of individual characteristics.

<sup>17</sup> This also prevents us from running a pure treatment effects model since treatment indicator variables will be perfectly correlated with either the subject fixed or random effects.

---

<sup>18</sup> 24 subjects provided inconsistent or irrational responses (see footnote 10). These people are coded as missing observations for the risk aversion dummy variable. Thus, they are excluded from the estimated fraud equations because they include that dummy variable. Our results are robust to coding these subjects as 0 for the risk aversion dummy variable.

<sup>19</sup> Note that the minimum level of the wealth variable is negative. If subjects are detected committing fraud in the early rounds of the session, the fine is sufficient to result in negative wealth.

<sup>20</sup> The reason the inverse of the period is used is to allow for diminishing marginal effects of learning. That is, subjects are most likely to learn in the early decision periods.

<sup>21</sup> The correlation between *lagged audit* and *lagged penalty* is 0.699. Thus, multicollinearity is of concern and may be responsible for the insignificance of the estimated coefficient on *lagged audit*.

<sup>22</sup> Subjects could choose levels of fraud between 0 and 500. Thus, the observations are censored at 0 and at 500. Given our experimental design we expect 400 observations to occur at both of these levels of fraud. We actually observe 180 observations that occur at 0 and 211 observations that occur at 500. OLS estimates are inconsistent in the presence of censoring and this prohibits the use of a Hausman test to verify the appropriateness of the random-effects specification.

<sup>23</sup> Since we employ a between subjects design our treatment dummies are perfectly correlated with any subject specific fixed or random effects. Thus, the most parsimonious model we can estimate must include at least one subject specific variable so that the random effects are not perfectly correlated with the explanatory variables.

<sup>24</sup> There is an inconsistency in the sense that estimated “learning” effects are significant in the effort equation and insignificant in the fraud equation. The fraud decision was more difficult

---

than the effort decision in that it required subjects to calculate expected values. Thus, it could be there were too few periods for any “learning” effects to be estimated for the fraud decision while there were enough for such effects to be estimated for the effort decision.

<sup>25</sup> This could be considered irrational behavior. As Davis and Holt (1993, p. 509) note individuals are observed to make decisions in lottery settings which may be seen to be irrational due to errors in probability estimation such as due to subjective probability assessment (Davis and Holt, p 465). Of course, this is a plausible explanation but we cannot be certain of the estimated coefficients, given the correlation between *lagged audit* and *lagged penalty* (see footnote 22). See Croson and Sundali (2005) for a detailed discussion concerning the prevalence of the gambler’s fallacy in both laboratory and field experiments.

<sup>26</sup> The objective function yields payoffs that are very similar at the 200 and 220 choices for fraud and we have elected to represent the prediction as a range.

<sup>27</sup> Andreoni (1995) cites this issue in reference to the often higher than predicted levels of contributions in public goods experiments. In our setting, subjects can only make errors in their fraud decision greater (less) than the theoretical prediction for treatment 2 (4).

<sup>28</sup> Calculated as  $[(\Delta 3 - \Delta 1) + (\Delta 4 - \Delta 2)] / 2$ .

<sup>29</sup> As Friedman and Sunder (1994) note, one cannot “prove” parallelism through deductive reasoning. In the end, it is inductive reasoning that allows us to assert that, because we have observed regularities, these will continue. Smith and Walker (1993) argue for ensuring that the decision rewards are commensurate with the decision costs and task complexity.