Was Sarbanes-Oxley Costly? Evidence from Optimal Contracting on CEO Compensation[1]

George-Levi Gayle
Department of Economics, Washington University in St. Louis
and Federal Reserve Bank of St. Louis

Chen Li
Zicklin School of Business, Baruch College, CUNY

Robert A. Miller
Tepper School of Business, Carnegie Mellon University

February 8, 2016

[1]We thank seminar participants at Baruch College for their comments. Li acknowledges financial support from a PSC-CUNY Award, funded by The Professional Staff Congress and The City University of New York. Email addresses: ggayle@wustl.edu (Gayle), chen.li@baruch.cuny.edu (Li), ramiller@andrew.cmu.edu (Miller). The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Abstract

This paper develops measures of the costs and benefits of governance regulations within a dynamic principal agent model of hidden information and moral hazard following the passage of the Sarbanes-Oxley Act (SOX). We estimate the effects of changes in CEO compensation for S&P 1500 firms and find SOX increased total compensation in the primary sector, increasing both its agency and administrative components. The net effect was mainly insignificant in the consumer and service sectors, with agency costs rising (falling) but administrative costs falling (rising) in larger (smaller) firms. (JEL C10, C12, C13, J30, J33, M50, M52, M55).

1 Introduction

This paper is an empirical investigation of the changes in CEO compensation resulting from the passage of the Sarbanes-Oxley Act (SOX), a legislative response enacted in 2002 by the U.S. government after a wave of corporate governance failures at many prominent companies. Many studies have investigated how SOX has affected firm behavior, including switching earnings management methods,[1] reducing investment,[2] and delisting.[3] Several studies attempt to quantify the net benefit of SOX by investigating the stock market reaction to the approval of the SOX provisions by the SEC, but the evidence is mixed and varies across firm type.[4] Yet how shareholders have modified CEO compensation in response to this regulation change is underexplored.[5] SOX has changed the environment that CEOs confront, and compensation is the crucial mechanism for exercising corporate governance to mitigate agency problems that arise from CEOs' hidden actions and information (Bolton and Dewatripont, 2005). The consequences of SOX for CEO compensation is thus an important factor in any overall evaluation of SOX. This paper investigates along which dimensions, and to what extent, the regulatory changes wrought by SOX exacerbated or mitigated the agency problems pertaining to executive management in different types of firms.

To evaluate the effects of SOX we estimate a dynamic principal agent model of optimal contracting. The advantage of taking this approach is that changes in the CEOs' compensation and agency costs can be explicitly attributed to changes in the fundamentals defining the primitives of the model. The framework also provides

[1] Cohen et al. (2008) find that accrual-based earnings management declined after the passage of SOX but real earnings management increased at the same time.

[2] Bargeron et al. (2010) find that, compared with non-U.S. firms, U.S. firms reduced investment in R&D and capital. Kang et al. (2010) find that (i) overall firms apply a higher rate to discount the payoff of investment projects and (ii) firms with good governance, with a good credit rating, and with early compliance of SOX 404 have become more cautious about investment.

[3] Engel et al. (2007) find that small firms chose to go private to avoid the cost of SOX. Leuz et al. (2008) show that the increased deregistration is mainly driven by firms that go dark, rather than private.

[4] Zhang (2007) finds a negative market reaction, Jain and Rezaee (2006) a positive one. Livtak (2007) finds that the decline in the stock price of foreign cross-listed firms was greater than for the US market index, cross-listed foreign firms not subject to SOX, and foreign firms not cross-listed. Hochberg et al. (2009) find that firms who had lobbied against SOX experienced positive abnormal returns.

[5] Cohen et al. (2007) documents a decline in pay-for-performance sensitivity after SOX. Carter et al. (2009) find the following: the weight of earnings increased for CEO bonus, upward earnings management decreased, and the cash salary components decreased in the total compensation after SOX. Nekipelov (2010) attributes an increase in salary and bonus post-SOX to increased risk aversion.

several measures of welfare costs that can be used to evaluate SOX. To understand how shareholders modify CEOs' compensation contracts to respond to SOX, we estimate the changes of agency costs embedded in CEOs' compensation from the pre-SOX period to the post-SOX period. These costs are due to two fundamental frictions in the agency relationships between shareholders and CEOs – that is, CEOs' hidden action (moral hazard) and their hidden information about firms' prospects.

Our model has four key features motivated by previous published work.[6] First, the model is based on hidden actions that create moral hazard, now widely acknowledged as the primal force explaining why the wealth of a CEO fluctuates with the value of the firm he manages. Second, the model also explicitly treats private information that CEOs directly benefit from through their holdings of financial securities in their own firms. This stylized fact is not controversial; for example, Gayle and Miller (2009a) show that following a simple portfolio strategy based on compensation schemes would have netted investors an extra 10 percent over and above holding the market portfolio. Third, accounting information is interpreted within the model as a signal that reveals the CEO's private information, reflecting a belief within the accounting profession that (i) executive management exercises considerable discretion in how they report on the firm's financial standing and (ii) nevertheless accounting reports do indeed convey information about the firm.[7] A fourth key feature of the framework we develop is that optimal contracting can be implemented as a sequence of short-term contracts, a property that is consistent with the claim by Holmstrom and Kaplan (2003) that corporate governance in the U.S. reacts quickly to legislative innovation. The optimality of short-term contracts implies there is no adjustment period between adjacent regimes regulating governance, a hypothesis we test to check the robustness of our findings.

Our analysis is organized as follows. Section 2 describes the data set, constructed from financial and accounting returns plus CEO compensation of the S&P 1500. As a precursor to the main analysis, Section 3 tests whether the distribution of financial returns and the distribution of CEO compensation both changed in a statistically

[6] See Murphy (1999, 2012) for empirical surveys of managerial compensation.

[7] In principle it is possible to treat hidden information as part of a pure moral hazard model, either because shareholders deter managers from misreporting by verifying their reports or because shareholders do not fully optimize over the contract space. See for example Peng and Röell (2008). However, Gayle and Miller (2015) find estimates of a pure moral hazard model that incorporates accounting information yield counter intuitive results. For example the model implies that in bad accounting states managers would be willing to pay shareholders to be employed.

significant sense after SOX was introduced. The results from the nonparametric tests we develop show that this is indeed the case. These twin findings motivate our model of CEO compensation, presented in Section 4, a dynamic model of optimal contracting between a risk-neutral principal (the shareholders) and a risk-averse agent (the CEO) where there is both hidden actions (of the CEO) and hidden information (about the future prospects of the firm). Section 5 defines the welfare measures we use to evaluate SOX. Then in Section 6 we explain the equilibrium for the model and show how it is related to the welfare measures we wish to compute. Identification is discussed in Section 7. Then in Section 8 we report our structural estimates of the welfare costs, summarizing our main findings in Section 9.

2 Data

Financial and accounting data on the S&P 1500 was extracted from Compustat, while data on executive compensation was taken from ExecComp. Bond prices were constructed from the yield curve using data from the Federal Reserve Data Set. Online Appendix A explains how the data were assembled. For the purposes of the study we classified each firm appearing in the S&P 1500 over the 13-year period 1993 through 2005 into one of three sectors: primary, consumer, and service. The top panel of Table 1 displays summary measures of assets, capital structure, and accounting returns by sector. Average total assets on the balance sheet of firm n at the end of annual period t, denoted by A_{nt}, are reported before and after SOX, in the first two columns for each sector, along with their standard deviations. Every third column shows the t (for means) or F statistic (for standard deviations) change between the two periods. Firms in the primary and consumer sectors are of comparable size, while those in the service sector are on average about four times as large but exhibit greater size variation. On average, A_{nt} grew significantly in every sector by roughly one third, the most in the primary sector, and so did its dispersion (as measured by the standard deviation, not shown). We define the debt-to-equity ratio, by $C_{nt} \equiv D_{nt} / (A_{nt} - D_{nt})$, where D_{nt} denotes debt at the end of the period. Average C_{nt} is almost twice as large in the service sector as the other two, but there is no discernible common trend across sectors pre and post SOX. Accounting returns are defined by $r_{nt} \equiv (A_{nt} - D_{nt} + I_{nt}) / (A_{n,t-1} - D_{n,t-1})$, where I_{nt} denotes the total value of dividends (and stock repurchases) paid throughout the preceding financial period.

The dispersion of accounting returns declined in all three sectors after SOX, which is curious since executive management exerts considerable discretion when reporting accounting earnings.[8]

The bottom panel of Table 1 displays average compensation and their standard deviations for the pre-SOX and post-SOX era by firm type, further partitioning them by accounting state, along with the t or F statistics for testing a change between the two periods (as in the top panel, in every third column). To facilitate comparisons of compensation pre and post SOX on total expected CEO compensation, all firm-year observations are grouped according to how they fit within the pre-SOX population of firm-years. Specifically, we classify each firm by whether in each (pre- and post-SOX) year its total assets were less than or greater than the median of total assets for firms in that sector in the pre-SOX era and whether its debt–equity ratio was less than or greater than the median debt–equity ratio for firms in that sector in the pre-SOX era. Therefore firm type is measured by the coordinate-pair (A, C) with each corresponding to whether that element is above (L) or below (S) its median of that industry in the pre-SOX era. For example (S, L) denote lower total assets a higher debt–equity ratio than the median debt–equity ratio for firms in that sector. Likewise Bad_{nt} means the accounting return r_{nt} is lower than the average for all firms within the same sector, size, and capital structure categories, and $Good_{nt}$ means the reverse. Following Antle and Smith (1985, 1986), Hall and Liebman (1998), and Margiotta and Miller (2000), our measure of total compensation includes salary, bonus, options, promised retirement benefits and restricted stocks, as well as the change in wealth attributable to holding financial securities in the firm rather than a fully diversified portfolio. In this way executive compensation depends directly upon the abnormal returns of the firms they manage, which we denote by $x_{nt} \equiv \pi_{nt} - \pi_t$, where π_{nt} denotes financial returns on equity in firm n at t, and π_t is the financial return from holding the market portfolio. On average CEO compensation is highest in the service sector and lowest in the primary sector.

Estimated mean CEO compensation significantly increased in all firm types within the primary sector after the SOX legislation was introduced and in all but one firm type even when conditioning on accounting state. However, with one exception,

[8] For example, the reporting of accruals, defined as the difference between realized cash flow and reported earnings, is one area where management may exercise considerable discretion. See Table A-1 for more details on these statistics.

estimated mean CEO compensation did not change significantly in the other two sectors. The dispersion of compensation, as measured by its standard deviation, fell in 18 out of the 24 sectors, and did not change significantly in the remaining 6. Broadly speaking, SOX compressed managerial compensation. The table also shows that accounting states matter: CEO compensation depends in part on what they themselves report, conditional on firm type and size. This is evident in two respects. Controlling for firm characteristics, average compensation is substantially lower in bad states than that in good states; these states are in part defined by CEOs exercising their considerable discretion about reporting unverifiable events. Moreover, compensation exhibits more variation in good states than bad states, as measured by their estimated standard deviations.

3 Testing for structural change

Mean CEO total compensation in every type of firm classification within the primary sector significantly increased after SOX was introduced, but it did not significantly change in any of the firm types in the other sectors (with a handful of exceptions). However this does not imply CEO compensation in the consumer and service sectors was unaffected by SOX. CEO compensation depends on abnormal returns. Therefore a structural shift occurs if the distribution of abnormal returns changes and/or the relationship between abnormal returns and CEO compensation changes. Here we test for equality, between the pre- and post-SOX eras, of the probability density functions for gross abnormal returns and shape of the compensation schedule.

Change in the distribution of abnormal returns: Denote the set of 24 categorical variables (formed from 3 sectors, 2 firm sizes, 2 capital structures, and 2 accounting states) by Z, and let $f_{pre}(x_{nt}|z_{nt})$ denote the probability density function of abnormal returns in the pre-Sox era conditional on $z_{nt} \in Z$. Also define $f_{post}(x_{nt}|z_{nt})$ in a similar manner. Under the null hypothesis of no change $f_{pre}(x|z) = f_{post}(x|z)$ for all $(x, z) \in \mathcal{R} \times Z$. Li and Racine (2007, page 363) propose a one-sided test for the null, in which the test statistic is asymptotically distributed standard normal. Panel A of Table 2 reports the test outcome for the 24 cases. (Online appendix B gives a detailed explanation of both the tests we conducted in this section.) Aside from the bad state of (S, L) in the consumer sector, the values of the statistic lie above the critical value of 5 percent confidence level (1.64); nearly all exceed the 1 percent level (2.33). Consequently we reject, for practically all firm types in both accounting states, the null hypothesis of no change in the abnormal returns density from pre-SOX to post-SOX eras.

Change in the shape of the contract: Let $w_{pre}(x_{nt}, z_{nt})$ denote CEO compensation as a function of (x_{nt}, z_{nt}) in the pre-SOX era and similarly define $w_{post}(x_{nt}, z_{nt})$ in the post-SOX era. A straightforward way of testing whether the two mappings are equal is to include an indicator variable for the post-SOX regime in nonparametric regressions of compensation on the gross abnormal return x_{nt} for each z_{nt}. The one-sided test of the null hypothesis of equality is asymptotically standard normal. Panel B in Table 2 reports the test statistics for a change in the shape of the compensation schedule for each of the 24 cases. In all but five cases, the value of the statistic exceeds 2.33, the 1 percent level, implying the null hypothesis of no change in the compensation contract shape is strongly rejected. Moreover in all five exceptions, Panel A shows we reject the null hypothesis that abnormal returns density function was unaffected, which implies that the probability distribution of managerial compensation in those cases did change when SOX was implemented.

Illustrating the differences: To convey some sense of what lies behind rejecting the null hypothesis of no change, Figure 1 shows how the shape of the gross abnormal returns probability density function and the estimated compensation schedule adjusts for small, low leveraged firms in the consumer goods sector, controlling for the state of the firm (bad versus good), and the two periods (pre-SOX versus post). The two top panels show that in both states density for abnormal returns shifted to the right in both states and became more concentrated about the mean. Comparing Panel A with B, mean returns are, not surprisingly, higher in the good state. The bottom panels show that in both periods the compensation schedule is steeper in the good state than the bad. In addition both plots in the post-SOX period, Panel D, tend to be flatter than in the pre-SOX period, Panel C. The overall effect of concentrating the abnormal returns distribution and flattening the extremes of the compensation schedule is to reduce the dispersion of compensation between the pre and post-SOX period, as reported in Table 1.

4 Model

The results from the first test show SOX had an impact on abnormal returns in all 3 sectors, over and above a common displacement effect on the returns to all firms. Conducting the second test showed that executive compensation committees also reacted to the changes wrought by SOX. But these tests cannot be used to decide whether the reaction was simply in response to the new distribution of abnormal returns or whether CEO functions changed. Answering that question requires a model of CEO compensation incorporating information asymmetries between the CEO and his shareholders, with primitives as parameters that might change with the implementation of SOX.

To this end we now lay out a dynamic principal agent model of optimal contracting between risk-neutral shareholders and a risk-averse CEO, based on Gayle and Miller (2015), in which the CEO has information and also takes actions that cannot be directly observed by shareholders. An important feature of this model is that it treats accounting information as a nonverifiable statement by the CEO, whose credibility depends on the incentives that determine his payoff as a function of what he reports.

At the beginning of period t, the CEO is paid compensation denoted by w_t for his work the previous period, denominated in terms of period-t consumption units. He makes his consumption choice, a positive real number denoted by c_t, and the board proposes a new contract. The board announces how CEO compensation will be determined as a function of what he will disclose about the firm's prospects, denoted by $r_t \in \{1, 2\}$, and its subsequent performance, measured by abnormal returns x_{t+1}, revealed at the beginning of the next period. We denote this mapping by $w_{rt}(x)$, the subscript t designating that the optimal compensation schedule may depend on current economic conditions, such as bond prices. Then the CEO chooses whether to be engaged by the firm or not. Denote this decision by the indicator $l_{t0} \in \{0, 1\}$, where $l_{t0} = 1$ if the CEO chooses to be engaged outside the firm and $l_{t0} = 0$ if he chooses to be engaged inside the firm.

If the CEO accepts employment with the firm, $l_{t0} = 0$, the prospects of the firm are now fully revealed to the CEO but partially hidden from the shareholders. There are two states, $s_t \in \{1, 2\}$, and we denote the probability that state s_t occurs by $\varphi_{st} \in (0, 1)$. We assume that CEOs privately observe the true state, $s_t \in \{1, 2\}$ in period t, gaining information that affects the distribution of the firm's next-period abnormal returns, and reports r_t to the board. If the CEO discloses the second state, meaning $r_t = 2$, then the board can independently confirm or refute it; thus, if $s_t = 1$, he reports $r_t = 1$. If $s_t = 2$, the CEO then truthfully declares or lies about the firm's prospects by announcing $r_t \in \{1, 2\}$, effectively selecting one of two schedules, $w_{1t}(x)$ or $w_{2t}(x)$, in that case.

The CEO then makes his unobserved labor effort choice, denoted by $l_{stj} \in \{0, 1\}$ for $j \in \{1, 2\}$ for period t which may depend on his private information about the state. There are two possibilities, to diligently pursue the shareholders objectives of value maximization by working, thus setting $l_{st2} = 1$, or to accept employment with the firm but follow the objectives he would pursue if he were paid a fixed wage by setting $l_{st1} = 1$, called shirking. Let $l_{st} \equiv (l_{t0}, l_{st1}, l_{st2})$. Since leaving the firm, working and shirking are mutually exclusive activities, $l_{t0} + l_{st1} + l_{st2} = 1$.

At the beginning of period $t+1$, abnormal returns for the firm, x_{t+1}, is drawn from a probability distribution that depends on the true state, s_t, and the CEO's action, l_{st}, in period t. We denote the probability density function for abnormal returns when the CEO works diligently and the state is s by $f_{st}(x)$. Similarly, let $f_{st}(x)g_{st}(x)$ denote the probability density function for abnormal returns in period t when the CEO shirks. Thus, for both states $s_t \in \{1, 2\}$:

$$\int x f_{st}(x) g_{st}(x) \mathrm{d}x \equiv E_{st}[x g_{st}(x) < E_{st}[x] \equiv \int x f_{st}(x) \mathrm{d}x, \tag{1}$$

the inequality reflecting the shareholders' preference for diligent work over shirking. Since $f_{st}(x)g_{st}(x)$ is a density, $g_{st}(x)$ is positive and integrating $f_{st}(x)g_{st}(x)$ with respect to x demonstrates $E_{st}[g_{st}(x)] = 1$. We assume the likelihood of shirking declines to zero as abnormal returns increase without bound:

$$\lim_{x \to \infty} [g_{st}(x)] = 0 \tag{2}$$

for each $s \in \{1, 2\}$. We assume the weighted likelihood ratio of the second state occurring relative to the first given any observed value of excess returns, $x \in R$ converges to an upper finite limit as x increases, such that

$$\lim_{x \to \infty} [\varphi_{2t} f_{2t}(x) / \varphi_{1t} f_{1t}(x)] \equiv \lim_{x \to \infty} [h_t(x)] = \sup_{x \in R} [h_t(x)] \equiv \overline{h}_t < \infty. \tag{3}$$

The CEO's wealth is endogenously determined by his consumption and compensation. We assume a complete set of markets for all publicly disclosed events effectively attributes all deviations from the law of one price to the particular market imperfections under consideration. Let b_t denote the price of a bond that pays a unit of consumption each period from period t onward, relative to the price of a unit of consumption in period t; to simplify the exposition, we assume b_{t+1} is known at period t. Preferences over consumption and work are parameterized by a utility function exhibiting absolute risk aversion that is additively separable over periods and multiplicatively separable with respect to consumption and work activity within periods. In the model we estimate, lifetime utility can be expressed as

$$-\sum_{t=0}^{\infty}\sum_{j=0}^{2}\beta^t \alpha_{jt} l_{tj} \exp\left(-\gamma_t c_t\right), \tag{4}$$

where β is the constant subjective discount factor, γ is the constant absolute level of risk aversion, and α_j is a utility parameter that measures the distaste from working at level $j \in \{0, 1, 2\}$. We assume working is more distasteful than shirking, meaning $\alpha_{2t} > \alpha_{1t}$, and normalize $\alpha_{0t} = 1$.

5 Welfare Measures

Figure 2 is a schema for the welfare measures we investigate. Total expected compensation in the pre-SOX era, defined as $\sum_{s=1}^{2}\varphi_{st} E_{st}\left[w_{st}(x)\right]$, can be decomposed into administrative costs in the pre-SOX era, denoted by τ_{1t}, and agency costs, τ_{2t}. Agency costs are further divided into τ_{3t}, which arises from pure moral hazard or the costs of hidden actions, and τ_{4t}, the extra cost from hidden information when there is moral hazard. Changes in τ_{it} from the pre- to post-SOX eras are denoted by $\Delta\tau_{it}$. We now explain how each of these measures appears in our model and why SOX might affect their values.

Absent agency considerations, shareholders would pay the CEO in the pre-SOX era an amount τ_{1t}, which we interpret as an administrative wage to work for the firm instead of pursuing an outside option – in other words the certainty equivalent of being employed as a CEO. In our model $\tau_{1t} \equiv b_{t+1}\left[(b_t - 1)\gamma_t\right]^{-1} \ln \alpha_{2t}$, where b_t denotes the bond price in period t. SOX imposed additional responsibilities on executive management that make the job more onerous. For example, Section 302 of SOX holds the principal executive officer(s) and the principal financial officer(s) responsible for establishing and maintaining internal controls. Denoting the change in the administrative wage from pre-SOX to post-SOX by $\Delta\tau_{1t}$, the increased regulations lead us to speculate that $\Delta\tau_{1t} > 0$.

The difference between expected total compensation and the administrative wage a CEO would receive, τ_2 in the pre-SOX era, is the risk premium of accepting employment that pays uncertain compensation rather than a fixed wage, which shareholders pay because of agency problems, and as such represents the amount shareholders are willing to pay for perfect monitoring. We define $\tau_{2t} \equiv \sum_{s=1}^{2} \varphi_{st} E_{st}\left[w_{st}(x)\right] - \tau_{1t}$. The provisions of SOX induced the firm to be more transparent about its future profitability. For example, Section 302 requires the principal executive officer(s) and the principal financial officer(s) to certify in each annual or quarterly report filed or submitted that the financial statements and other financial information fairly present financial conditions and results and refrain from making misleading statements. Using legal machinery to enforce the truthful revelation of financial conditions may remove or ease the burden of the compensation committee in designing incentives that resolve the agency issues. For these reasons we might expect $\Delta\tau_{2t} < 0$.

The component of agency costs solely attributable to pure moral hazard, or the amount shareholders would pay to eliminate private information, is the difference between the expected compensation under the current optimal contract and expected compensation in the pure moral hazard case, which for the pre-SOX era we denote by $y_{st}(x)$. Thus $\tau_{3t} \equiv E_{st}\left[w_{st}(x) - y_{st}(x)\right]$. Many of the requirements mandated by SOX apply whether or not CEOs pursue their own objectives subject to their legal obligations or receive compensation that induces them to work in the interests of shareholders. The intention of SOX seems to have been to reduce the moral hazard problem, so if the legislation succeeded on those terms $\Delta\tau_{3t} < 0$. For example Section 304 of SOX requires the CEO and chief financial officer to reimburse the firm for any compensation received during the 12-month period following equity issue filing if there was misconduct in filling a financial statement for that equity issue. This regulation makes CEO compensation less liquid and so can mitigate incentives to make myopic investments and to take opportunistic advantage by misrepresenting financial states. Consequently, deviating from shareholder interests is less attractive, possibly aligning the interests of shareholders and CEOs more closely and mitigating the moral hazard problem.

By construction $\tau_{4t} \equiv \tau_{2t} - \tau_{3t}$. If shareholders could observe the CEO's effort, then a first-best constant-wage contract would be paid regardless of whether the CEO had private information or not (Gayle and Miller, 2009b). Therefore the only reason why hidden information might be costly, meaning $\tau_{4t} > 0$, is that it exacerbates rather than ameliorates the pure moral hazard model, which we show is an empirical question since theory does not give a decisive answer. One purpose of SOX was to enhance the independence of auditors and boards conducting monitoring functions, making shareholders more informed, presumably to reduce the role of hidden information. Consequently we might expect $\Delta\tau_{4t} < 0$.

The remaining symbols, ρ_{1t} through ρ_{3t}, are summary measures of the channels SOX flowed through, changing the values of the primitives in our model to affect the welfare measures. Specifically, $\rho_{1t} \equiv E_{st}[x - xg_{st}(x)]$ is the loss shareholders would incur from a CEO shirking instead of working; $\rho_{2t} \equiv b_{t+1}\left[(b_t - 1)\,\gamma_t\right]^{-1}\ln(\alpha_{2t}/\alpha_{1t})$ is the difference between a CEO's pecuniary cost of working and that of shirking; ρ_{3t} measures how much the loss from pure moral hazard changed because of the shift in the signal.

6 Equilibrium

The welfare measures, τ_{1t} through τ_{4t}, and the summary measures of the driving forces of the agency problem, ρ_{1t} through ρ_{3t}, are functions of the utility parameters and the parameters determining the distribution of abnormal returns. Yet the state s and a sample analogue to φ_{st} (the probability of each state) is not directly observed, and the parameters defining utility, γ_t, α_{1t}, and α_{2t}, and the likelihood ratio $g_{st}(x)$ cannot be estimated for either state $s \in \{1, 2\}$ without making behavioral assumptions about shareholders and CEOs. We now assume shareholders have diversified portfolios and are expected value maximizers, while CEOs are expected utility maximizers.

6.1 Optimization

In this framework, there are no gains from a long-term arrangement between shareholders and the CEO: the optimal long-term contract between shareholders and the CEO decentralizes to a sequence of short-term one-period contracts. (Both lemmas in this section are proved in the Appendix.)

Lemma 1 *Denote by* $\bar{\varsigma}$ *the date the CEO retires. The optimal long-term contract can be implemented by a* $\bar{\varsigma}$*-period replication of the optimal short-term contract.*

The next lemma solves the optimal consumption and savings plan for a CEO about to retire. It proves in our model, given the CEO's reporting about the state of the firm and the true state of the firm, that his employment and effort choices depend on his preference parameters $(\alpha_{1t}, \alpha_{2t}, \gamma_t)$, the distribution of abnormal returns when he shirks $f_{st}(x)g_{st}(x)$ and when he works $f_{st}(x)$, and aggregate economic conditions as reflected in the bond prices (b_t, b_{t+1}). However, the employment and effort choices do not depend on his current (outside) wealth. To state the lemma, let $r_t(s)$ denote the CEO's disclosure rule about the state when the true state is $s_t \in \{1, 2\}$.

Lemma 2 *If the CEO, offered a contract of* $w_{rt}(x)$ *for announcing* r*, retires in period* t *or* $t+1$ *by setting* $(1 - l_{t0})(1 - l_{t+1,0}) = 0$*, upon observing the state* s *and reporting* $r_t(s)$*, he optimally chooses* $l_{st} \equiv (l_{t0}, l_{st1}, l_{st2})$ *to minimize*

$$\sum_{s=1}^{2} \varphi_{st} \left\{ l_{t0} + (\alpha_{1t} l_{st1} + \alpha_{2t} l_{st2})^{1/(b_t - 1)} E_{st} \left[\exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) [g_{st}(x) l_{st1} + l_{st2}] \right] \right\}. \tag{5}$$

The optimal short-term contract for shareholders is found by minimizing the expected compensation subject to four constraints that the CEO prefers (i) to work for a period rather than leave the firm, (ii) to be truthful rather than lie, (iii) working to shirking, and (iv) to be truthful and working diligently than lying and shirking. Suppressing for expositional convenience the bond price b_{t+1} and recalling our assumption that b_{t+1} is known at period t, we now let $v_{st}(x)$ measure how (the negative of) utility is scaled up by $w_{st}(x)$:

$$v_{st}(x) \equiv \exp\left(-\gamma_t w_{st}(x) / b_{t+1}\right). \tag{6}$$

First, to induce an honest, diligent CEO to participate, his expected utility from employment must exceed the utility he would obtain from retirement. Setting $(l_{t2}, r_t) = (1, s_t)$ in (5) and substituting in $v_{st}(x)$, the participation constraint is thus

$$\sum_{s=1}^{2} \int \varphi_{st} v_{st}(x) f_{st}(x) \mathrm{d}x \leq \alpha_{2t}^{-1/(b_t - 1)}. \tag{7}$$

Second, given his decision to stay with the firm one more period and to truthfully reveal the state, the incentive-compatibility constraint induces the CEO to prefer working to shirking for $s_t \in \{1, 2\}$. Substituting the definition of $v_{st}(x)$ into (5) and comparing the expected utility obtained from setting $l_{t1} = 1$ with the expected utility obtained from setting $l_{t2} = 1$ for any given state, we obtain the incentive-compatibility constraint for work:

$$0 \leq \int \left(g_{st}(x) - (\alpha_{2t}/\alpha_{1t})^{1/(b_t - 1)} \right) v_{st}(x) f_{st}(x) \mathrm{d}x. \tag{8}$$

Information hidden from shareholders further restricts the set of contracts that can be implemented. Comparing the expected value from lying about the second state and working diligently with the expected utility from reporting honestly in the second state and working diligently, we obtain the truth-telling constraint:

$$0 \leq \int \left[v_{1t}(x) - v_{2t}(x) \right] f_{2t}(x) \mathrm{d}x. \tag{9}$$

An optimal contract also induces the CEO not to understate and shirk in the second state, behavior we describe as sincere. Comparing the CEO's expected utility from lying and shirking with the utility from reporting honestly and working diligently, the sincerity condition reduces to

$$0 \leq \int \left[(\alpha_{1t}/\alpha_{2t})^{\frac{1}{b_t - 1}} v_{1t}(x) g_{2t}(x) - v_{2t}(x) \right] f_{2t}(x) \mathrm{d}x, \tag{10}$$

where $(\alpha_{1t}/\alpha_{2t})^{1/(b_t - 1)} v_{1t}(x)$ is proportional to the utility obtained from shirking and announcing the first state and $f_{2t}(x) g_{2t}(x)$ is the probability density function associated with shirking when the second state occurs. Minimizing expected compensation amounts to choosing $v_{st}(x)$ that maximizes

$$\sum_{s=1}^{2} \int \varphi_{st} \ln \left[v_{st}(x) \right] f_{st}(x) \mathrm{d}x. \tag{11}$$

Noting $\ln v_{st}$ is concave increasing in v_{st}, the expectation operator preserves concavity, so the objective function is concave in $v_{st}(x)$ for each x. Each constraint is a convex set and their intersection is too. Therefore, we can appeal to the Kuhn Tucker theorem which guarantees there is a unique positive solution to the equation system formed from the first-order conditions augmented by the complementary-slackness conditions.

6.2 Comparing the pure and hybrid model contracts

The optimal contract for a parameterization of the hybrid model is plotted in Panel A of Figure 3. This parameterization follows Margiotta and Miller (2000) in assuming that abnormal returns are drawn from a truncated distribution for abnormal returns, with a common lower bound for all states and independent of the effort level.[9] For comparison purposes we also plotted in Panel B optimal compensation for the analogous two-state pure moral hazard model (where there are hidden actions but the state is known), denoted by $y_{st}(x)$. Details explaining the solution and computation of $y_{st}(x)$ are provided in Online appendix C.

The figure illustrates four important features. Since compensation in both models is a function of the likelihood ratio between the densities of the abnormal return for working and shirking, not the abnormal return itself, the wage contract is not necessarily monotonically increasing in abnormal returns. For example, in the bad states of both models of the illustrated parameterization, pay optimally declines with marginal increments to abnormal returns when they are less than -0.5. The same explanation applies to compensation levelling out at high levels of abnormal returns; the likelihood ratio converges to a constant, zero, under the assumption of a truncated normal distribution.

The other two noteworthy features relate to differences between the pure and hybrid contracts. The slope of the hybrid compensation schedule is everywhere greater in the good state than the bad, whereas in the pure moral hazard model the slope in the bad state is greater than in the good over the intermediate range where much of the probability mass of both abnormal distributions lie. Thus the point where the schedules cross is higher in the pure moral hazard model than in the hybrid model. The figure also illustrates two analytical results: in the hybrid model expected utility of the agent is greater in the good state than the bad, but in the pure moral hazard model, expected utilities are equalized across states. Intuitively, the argument is that in the hybrid model the principal induces the agent to truthfully reveal the good state by promising (i) more expected utility in the good state and (ii) a flatter compensation profile in the bad state.

Finally, because the constraints in the pure moral hazard optimization problem are

[9] If the lower bound depends on whether the agent works or shirks, a first-best solution is attained by imposing a sufficiently harsh penalty on the agent when abnormal returns can be attained only by shirking, and otherwise paying the agent the first-best fixed wage. See Mirrlees (1975).

not a subset of those in the hybrid model, there is no presumption that the expected compensation in the pure moral hazard case is lower than in the hybrid model. In other words the principal may find it cheaper not to know the private information if he can optimally spread the utility the agent receives across both states, rather than meet the participation constraint in each state.[10] Indeed our parameterization illustrates an instance where the agency cost in the pure moral hazard model is greater than in its hybrid counterpart. The parameterization demonstrates a paradox: to the extent it succeeds in making corporate governance more transparent, SOX may have perverse consequences in some sectors of industry.

7 Identification and Estimation

The parameters defining the model are characterized by $f_{st}(x)$ and $g_{st}(x)$ for $s_t \in \{1, 2\}$, which together define the probability density functions for revenue in each state, and φ_{st}, the probability of each state occurring. CEO preferences are defined, relative to the normalized utility from taking the outside option, by their distaste for working, α_{2t}, and shirking, α_{1t}, as well as their risk-aversion parameter, γ_t. In equilibrium, CEOs truthfully reveal the state, implying $s_t = r_t(s)$, so $r_t = s_t$ is observed in the data. Hence the reported state is sampled from a Bernoulli distribution with parameter φ_{st}, and the data on returns are generated by $f_{st}(x)$, implying those parameters are identified, the latter nonparametrically. Aside from observing returns from working, we assume that compensation, $w_{st}(x)$, is also observed for different values of (x, s). This leaves only $g_{st}(x)$ plus $(\alpha_{1t}, \alpha_{2t}, \gamma_t)$ to identify from the first-order conditions, the complementary slackness conditions, plus a constraint that working is an optimal choice, from observations on (x_{nt}, s_{nt}, w_{nt}) generated from the CEO working.[11] This section explains the intuition supporting identification and estimation, while the supporting technical details are relegated to Online appendix D.

We motivate the identification of this model by comparing the equilibrium compensation schedule and the excess return density shown in Figure 3 with the sample

[10] There are assumptions guaranteeing expected compensation in the hybrid model is more expensive than in the pure moral hazard model. For example, if the two distributions for the good state are simply a shift of the distributions in the bad state by constant amount to the right, the optimal contract in the pure moral hazard model depends only on the state through the translation parameter and is therefore cheaper than the optimal contract of the hybrid model.

[11] Although $(\widetilde{w}, x, r)$ rather than (w, x, r) is observed, there is no loss in generality from assuming (w, x, r) is observed because $w_r(x) = E[\widetilde{w}|X = x, R = r]$.

estimates displayed for one of the sectors in Figure 1. Both the theoretical and the empirically estimated compensation schedules vary with excess returns, for the most part increasing, and flatten at very high rates of excess returns. These features illustrate the agency problem. Moreover in the estimated schedules for both states and in the hybrid model, but not in the model of pure moral hazard, the schedules for the good state are everywhere steeper than for the bad state, and also cross at negative abnormal returns. This suggests that hidden information, not just hidden actions, may be a part of the agency problem.

Gayle and Miller (2015) have shown in a static analogue to our dynamic model that if γ, the risk parameter, is known, then the other two preference parameters α_{1t} and α_{2t} are point identified, along with the likelihood ratios $g_{1t}(x)$ and $g_{2t}(x)$, with both mappings being nonparametrically identified. In terms of our dynamic framework, suppose γ_t is the value of the risk aversion parameter. Then it follows directly from the first-order condition of the compensation contract, the participation and incentive compatibility constraints (both of which are binding), and the regularity conditions for $g_{st}(x)$ and $h_t(x)$ that for every period t:

$$\begin{aligned}
\alpha_{2t} &= E\left[v_{st}(x,\gamma_t)\right]^{1-b_t} \\
\alpha_{1t} &= \alpha_{2t}\left\{\overline{v}_{2t}(\gamma)^{-1} - E_{2t}\left[v_{2t}(x,\gamma_t)^{\ -1}\right]\right\}^{b_t-1}\left\{\overline{v}_{2t}(\gamma)^{-1} - E_{2t}\left[v_{2t}(x,\gamma_t)\right]^{-1}\right\}^{1-b_t} \\
g_{2t}(x) &= \frac{\overline{v}_{2t}(\gamma_t)^{-1} - v_{2t}(x,\gamma_t)^{-1}}{\overline{v}_{2t}(\gamma_t)^{-1} - E_{2t}\left[v_{2t}(x,\gamma_t)^{-1}\right]} \\
g_{1t}(x) &= \frac{\left(\frac{\alpha_{2t}}{\alpha_{1t}}\right)^{1-b_t}\left\{\overline{v}_{1t}(\gamma_t)^{-1} - v_{1t}(x,\gamma_t)^{-1} + \eta_{3t}\left[\overline{h}_t - h_t(x)\right]\right\} - \eta_{4t}g_{2t}(x)h_t(x)}{\overline{v}_{1t}(\gamma_t)^{-1} - E_t\left[v_{st}(x,\gamma_t)\right]^{-1} + \eta_{3t}\overline{h}_t},
\end{aligned} \tag{12}$$

where $\overline{v}_{st}(\gamma_t) \equiv \lim v_{st}(x,\gamma_t)$ as $x \rightarrow \infty$ and the Kuhn Tucker multipliers η_{4t} and η_{3t} have the representation:

$$\begin{aligned}
\eta_{4t} &= \frac{\frac{E_{1t}[v_{1t}(x,\gamma_t)]}{E_t[v_{st}(x,\gamma_t)]} - E_{1t}\left[v_{1t}(x,\gamma_t)h_t(x)\right]\left\{E_{2t}\left[v_{2t}(x,\gamma_t)\right]^{-1} - E_t\left[v_{st}(x,\gamma_t)\right]^{-1}\right\} - 1}{\left(\frac{\alpha_{2t}}{\alpha_{1t}}\right)^{1-b_t} E_{1t}\left[v_{1t}(x,\gamma_t)g_{2t}(x)h_t(x)\right] - E_{1t}\left[v_{1t}(x,\gamma_t)h_t(x)\right]} \\
\eta_{3t} &= E_{2t}\left[v_{2t}(x,\gamma_t)\right]^{-1} - \eta_{4t} - E_t\left[v_{st}(x,\gamma_t)\right]^{-1}
\end{aligned} \tag{13}$$

As a practical matter, if γ_t is known or can be estimated then the remaining parameters of the model can be estimated by taking sample analogues of their population counterparts that appear in the expressions above.

Combining the nonparametric estimates of the density of excess return and relationship between excess return and compensation presented in Figure 1 with the formulae for the structural parameters in equation (12), one can glean the sources of variation in the data that identify the other structural parameters given a known γ_t. The first three equations fully apply to a pure moral hazard model in which the good state occurs with probability one. First, α_{2t} is identified from the ex-ante expected discounted utility derived from the compensation schedule.[12] The identification of α_{1t}/α_{2t}, and hence α_{1t}, comes from the concavity of the compensation schedule relative the maximum compensation in the good state. The likelihood ratio in the (verifiable) good state, $g_{2t}(\cdot)$, is identified off the slope of the compensation schedule in the good state. The last equation identifies $g_{1t}(\cdot)$ given all the other parameters in the model. As with $g_{2t}(x)$, it also depends on the slope of the compensation schedule in the same state (unverifiable in this case), but $g_{1t}(x)$ also depends on the slope of the compensation schedule in the other state, as well as $h_t(x)$, the likelihood ratio of either state given x. Since $h_t(x) \to \overline{h}_t$ as $x \to \infty$, given a compensation schedule, the equation identifying $g_{1t}(x)$ shows that we would infer abnormal returns under shirking tend to be lower in the hybrid model than in the corresponding pure moral hazard model.

Other restrictions, in the form of equalities and inequalities implied by the model, can be used to obtain bounds for admissible values of γ_t. At least one of the truth-telling constraints and the sincerity constraint should be binding. The three other Kuhn-Tucker multipliers are non-negative. Similarly the complementary-slackness conditions for the truth-telling and sincerity constraints yield two more equalities. We impose an exclusion restriction that α_{1t} does not depend on the private states. The likelihood for the bad state, $g_{1t}(\cdot)$, is positive with unit mass.[13] Value maximization implies another three inequalities reflecting that shareholders prefer the CEO working in both private states rather than shirking in either or both of them. Finally we impose the restriction that the risk aversion does not depend on bond price. Gayle and Miller (2015) obtained sharp and tight bounds for the set of observational equivalent risk

[12] This expectation is taken before the realization of the hidden information states to the manager.

[13] One can show that $g_{2t}(\cdot)$ is positive without imposing any restrictions on γ_t. See Gayle and Miller (2015).

aversion parameters, and we can adopt their methods and results to our framework. Accordingly let $\Gamma_t \equiv \{\gamma_t : Q_t(\gamma) = 0\}$ denote the Borel set of admissible values of γ_t for data generated in period t, where $Q_t(\gamma_t)$ is a quadratic form of the minus-norm of equalities and inequalities implied by the model.[14] Thus Γ_t denotes all the values of γ that are observational equivalent given the probability distributions generating the data at period t. Imposing additional restrictions that arise from multiple time periods is straightforward. For example if γ is time invariant across the first two periods, labelled 1 and 2, then the set of admissible risk aversion parameters is the intersection $\Gamma_1 \bigcap \Gamma_2$. We denote by Γ the identified set that arises from imposing all the relevant restrictions for the different time periods and denote its defining quadratic from by $Q(\gamma)$.

Since γ is not identified point-wise but only up to the set Γ, it follows that the other taste parameters, the likelihood ratios, and the measures of agency cost are only set identified too. We derive a confidence region that covers the identified set of observational equivalent parameters, any element of which could have generated the data. For example, we can write $\alpha_{1t}(\gamma)$ as the value of α_1 identified from data that are generated at t by the hybrid model when the agent's risk aversion parameter is γ. Thus α_1 is identified up to the set $\{\alpha_{1t}(\gamma) : \gamma \in \Gamma\}$. Further restrictions obtained from the panel are imposed in the same way that admissible values of γ are restricted.

We estimated a confidence region for Γ by exploiting the fact that approximations to $Q(\gamma)$ formed from the data only deviate from zero because of differences between expectations (or, in some equations, population limits) and their sample analogues. Accordingly let $Q^{(NT)}(\gamma)$ denote a sample analogue to $Q(\gamma)$ and define $\Gamma^{(NT)} \equiv \left\{\gamma : Q^{(NT)}(\gamma) \leq c_{0.95}\right\}$ where $c_{0.95}$ is the critical value, below which $Q^{(NT)}(\gamma)$ falls 95 percent of the time under the null hypothesis that $\gamma \in \Gamma$. Once $\Gamma^{(NT)}$ has been numerically determined (by subsampling in our application), we can deduce that the estimated confidence region for $\{\alpha_{1t}(\gamma) : \gamma \in \Gamma\}$, for example, is $\left\{\alpha_{1t}(\gamma) : \gamma \in \Gamma^{(NT)}\right\}$. Estimated confidence regions for the other primitive parameters and the welfare measures are derived following exactly the same procedure.

[14] The minus-norm of q, denoted $\|q\|_-$, is the norm of the maximun of $-q$ and 0, i.e., $\|q\|_- = \|\max(-q, 0)\|$.

8 Empirical Findings from the Model

We obtained estimates for two sample sizes. The estimates for the full sample cover 1993 through 2002 for the pre-SOX (containing 18,855 observations), and 2003 though 2005 for post-SOX years (5,670 observations). As a robustness check, we also report estimates for a restricted number of years by omitting the two years bordering on the SOX legislation, namely, 1993 through 2001 (16,894 observations) and 2004 through 2005 (3,781 observations). We could not reject the null hypothesis that φ_{st}, $f_{st}(x)$, α_{1t}, α_{2t}, and γ_t are time invariant within the two samples. Thus bond prices were sufficient to capture all aggregate variation within each era. To achieve comparability between the two eras we estimated the confidence region for each of the two bond prices that occurred in both eras, 16.4 and 16.8. This permits us to attribute all changes in social welfare costs to the changes in primitives rather than aggregate factors in the macroeconomy. Since the differences between the two sets of estimates are negligible, we report only those for 16.4. To account for heterogeneity in the data, the estimation is also conditional on the accounting state and firm type (by sector, assets, and capital structure).

Risk preferences: The procedure used to obtain a confidence region for changes in the welfare measures depends somewhat on whether the risk aversion parameter γ changes between the pre- and post-SOX eras. If the null hypothesis, that the risk parameter was constant over this period, is maintained, then it is straightforward to compute confidence regions for $\Delta\tau_{it}$ for $i \in \{1, \ldots, 4\}$ by substituting those values of $\gamma_t \in \Gamma^{(NT)}$ into the formulas for $\Delta\tau_{it}$ (that come from evaluating τ_{it} in the pre- and post-SOX eras). If the confidence regions for $\Delta\tau_{it}$ contain only positive (negative) values, then we cannot reject the null hypothesis that the change is positive (negative). On the other hand, if the null hypothesis is rejected, the confidence region computed for $\Delta\tau_{it}$ is based on admissible values of two parameters γ_{pre} and γ_{post}, not just one, and a further component in the decomposition is introduced, which measures the contribution of the change in risk attitude to $\Delta\tau_{it}$.

For these reasons we first tested whether the null hypothesis, that risk aversion did not change, is rejected, and only then did we construct the appropriate confidence regions for the welfare measures. The test is of independent interest. One concern raised by directors (Cohen et al., 2013) and bankers such as Alan Greenspan and William Donaldson (former SEC chairman) is that CEOs would overreact to SOX provisions and exercise undue caution in investment decisions thus destroying shareholder value (see Coats and Srinivasan, 2014).

The 95 percent confidence region of the risk-aversion parameter for the full sample of the post-SOX period covers a wider range, (0.0616, 0.2335), than that of the pre-SOX period, (0.0784, 0.2335), for which there are more years (and hence more restrictions) and coincides with the confidence region obtained under the null hypothesis of no change. Therefore we do not reject the null hypothesis that no significant change in risk attitude occurred after SOX.[15] The estimated confidence region for the restricted sample is common to both periods (0.0695, 0.6158), similar to the unrestricted sample; indeed the confidence region for risk aversion parameters for both periods in the full sample is a proper subset of the corresponding region in the restricted sample. Thus adding the restrictions for the years 2002 and 2003 to the sample yields more precise results and provides some evidence that they are robust.[16]

To give economic meaning to our estimates of risk aversion, we also computed the amount a CEO would pay to avoid an equiprobability gamble with losing or winning \$1,000,000. At the right boundary of the confidence region for the full sample in the pre-SOX period, the risk-aversion parameter is 0.2335, implying the CEO would pay \$115,704 to avoid the gamble, but at the left boundary of 0.0784 would pay only \$39,160. These estimates are in line with previously published work.[17] We conclude that changes to the distribution of abnormal returns and its mapping to CEO compensation documented in Section 3 did not arise because the implementing SOX legislation induced CEOs to think differently about risk. Other factors in our model caused the changes. We now investigate these other factors under the maintained hypothesis that the risk-aversion parameter was constant over the entire period.

Administrative costs: Administrative costs, denoted by τ_{1t}, is the premium that a CEO would be paid over the inclusive annuitized value of his outside option if there were no agency problems. Column A of Table 3 shows that these vary greatly by sector and firm type, but most of the variation is explained by the firm categories.

[15] These findings contrast with those of Nekipelov (2010), who finds the risk aversion of top executives in the retail apparel industry significantly increased after SOX was introduced. Three notable differences between his work and ours is that Nekipelov assumes the contract is linear, approximates compensation by salary and bonus, and of course estimates from a different sample population.

[16] Table A-3 in the online appendix reports the findings in more detail.

[17] See Gayle and Miller (2009a, 2009b, 2015) and Gayle et al. (2015)

For example, in the pre-SOX regime, the 95 percent confidence region for the administrative cost of (S, L) firms in the primary sector is covered by the interval ranging from \$1.3 to \$1.4 million; and in (L, S) firms belonging to the service sector, the corresponding region is covered by the interval ranging from \$9.8 to \$10.7 million. In both samples we cannot reject the hypothesis that $\triangle\tau_{1t} > 0$ in at least four of the categories and that $\triangle\tau_{1t} < 0$ in at least four. In both samples every firm category within the primary sector experienced increased administrative costs, of between \$1.3 and \$3.0 million in the full sample. Our estimates from the restricted sample show that every category within the service sector experienced declines of about to \$4.1. Both results broadly reflect our findings in Table 1, which shows that mean CEO compensation significantly increased in every subcategory within the primary sector following SOX, but did not significantly increase in any subcategory of the service industry.

Agency costs: Agency costs, τ_{2t}, measure the gross costs that shareholders would be willing to pay for perfect monitoring and thus avoid the penalties induced by the incentive compatibility and truth-telling constraints. Table 4 reports the 95 percent confidence region for the observational equivalent values of τ_{2t} and its change $\triangle\tau_{2t}$. In all but one case the full and restricted samples yield the same qualitative results. Agency costs are small in some firm categories, as low as \$34,000 per year within (S, L) firms within the primary sector, but within the service sector, much greater, between \$122,000 and \$1.345 million. SOX increased agency costs in 9 out of the 12 sectors in the full sample (and 10 of 12 in the restricted). For the most part, the absolute value of the changes are small to moderate, exceeding one million dollars only once. However as a proportion of the levels, they are quite substantial, the estimated upper bound on $\triangle\tau_{2t}$ is at least as large as the lower bound of τ_{2t} in several categories.

Cost of hidden actions: Shareholder costs due to hidden actions, denoted by τ_{3t}, are the difference between the expected compensation which would have been if there were only a pure moral hazard problem and the certainty equivalent wage if CEOs could be perfectly monitored. Table 5 reports the estimated 95 percent confidence region for τ_3 in the pre-SOX era and the change it heralded, $\triangle\tau_3$. The estimated bounds of the confidence intervals for the pre-SOX era range between \$6,000 and \$3.6

million depending on firm type, markedly lower in the primary sector than the other two. The effects of SOX on the estimated cost of hidden actions varied by sector and firm type; in the restricted sample all the increases occurred within the primary sector, but the evidence in the full sample is less compelling.

Cost of hidden information: Recall the costs of hidden information, τ_{4t}, is the difference between the expected compensation and what it would have been with hidden actions but not hidden information about the firm's state. Table 6 displays a property foreshadowed in Figure 2. With a single exception hidden information ameliorates pure moral hazard. In these cases adding truth telling and sincerity constraints to the principal's optimization problem is less costly than adding the extra participation constraint that would arise in a pure moral hazard model. The net benefits range from \$27,000 in (S, S) firms belonging to the primary sector to \$2.6 million in (L, L) firms in the consumer sector. In the primary sector, $\triangle\tau_{4t} < 0$ in three of the four categories: the welfare costs of hidden information declined after SOX was introduced in those categories, increasing the gap (or in one instance creating a positive gap) between the effects of a hybrid versus a pure moral hazard model by increasing the net benefits of keeping information private. Overall $\triangle\tau_{4t}$ is relatively small and imprecise, exceeding \$1 million in only one case.

Gross loss from shirking to shareholders: The sources of the agency problem arise from conflicting objectives between the shareholders, who want the CEO to work to maximize their returns, and the CEO, who prefers to shirk. The expected gross loss to shareholders that would occur each year from the CEO shirking instead of working, denoted by ρ_{1t}, is reported as a percentage of market value in Table 7. Similar to estimates found in previous studies, they range from 5.9 to 18.9 percent per year.[18] As in previous tables, variation explained by firm category far outweighs the indeterminacy from observational equivalence that arise from set, rather than point, identification. In particular, large firms tend to lose proportionately less than small firms from a shirking CEO. Also with the notable exception of (S, S) firms in the consumer sector, given size and leverage, shareholders owning firms in the service sector have the most to lose from not aligning management objectives with their own.

[18] A variety of different models and different estimators corroborate these estimates: see Margiotta and Miller (2000), Gayle and Miller (2009a, 2009b, 2015), and Gayle et al. (2015).

The most striking new result in this table is that in both samples $\triangle\rho_{1t} < 0$ for 11 out of 12 sectors. Overall the effect of SOX was to limit the expected losses a CEO would impose on his firm from not pursuing a goal of expected value maximization. This finding indirectly raises a question: whether the purpose of SOX regulation was to make firms more efficient or was devised to serve some other objective? For example, legislators supporting SOX could claim that they were partially protecting shareholders against rogue management, which in our model of rational behavior can occur only by deviating from equilibrium. Given the magnitudes involved, partial insurance against a small but strictly positive probability might seem like a small price to pay.

Benefit from shirking to CEO: The other side of the conflict driving the agency problem is the compensating differential the CEO is paid to work rather than shirk, measured by ρ_{2t}. Table 8 shows our estimates are a tiny fraction of the losses shareholders would incur, ranging between \$1.6 million and \$9.9 million annually, a result corroborated by our previous work. Depending on which sample is used in estimation, the conflict of interest CEOs faced declines in one third to a half the firm categories, but is exacerbated in the remainder. The main new finding this table shows is that when SOX was introduced the differential mostly declined in the firm categories where it was relatively high. This is most noticeable in large firms belonging to the service sector and all firm categories in the primary sector. Thus SOX had an equalizing effect on the working versus shirking compensating differential, rendering shirking a more homogeneous activity that does not depend as much on firm type. To some extent then, greater regulation of management imposed by SOX, including the attendant legal responsibilities, enforcement, and penalties, channelled the type of shirking that occurs if and when CEOs are not properly incentivized.

Signal quality: The reason why the different objectives cannot be resolved by fiat is the signal shareholders use to evaluate the actions of the CEO, abnormal returns, is an imperfect measure of CEO effort. To complete our empirical analysis we investigated whether SOX improved the signal. That is, holding the nonpecuniary benefits of the CEO constant at their pre-SOX level, what effect does changing the signal, from the pre-SOX to the post-SOX regime, have on the cost of pure moral hazard? Table 9 reports our results of ρ_{3t}. The first set of results shows the effect of substituting

α_{1post} and α_{2post} for their pre-Sox values in the construction of τ_3. For the most part the results are comparable to those reported in Table 7. In about half the firm categories the signal improves (in both samples). The magnitudes are relatively small, typically in the tens or hundreds of thousands of dollars. We conclude that changes in signal quality were not an empirically important factor in explaining the shifts in the compensation function and the probability density for abnormal returns after the introduction of SOX legislation.

9 Conclusion

SOX was a legislative response by the U.S. government to corporate governance failures at many prominent companies. This paper conducts an empirical analysis of its effects on CEO compensation using panel data constructed for the S&P 1500 firms on CEO compensation, financial returns, and reported accounting income. Our structural empirical analysis is motivated by the empirical facts that after SOX was exacted there were significant changes in (i) the relation between a firm's abnormal returns and CEO compensation and (ii) the underlying distribution of abnormal returns. The net effect of these changes was to significantly raise expected CEO compensation in the primary sector, but not in the consumer and service sectors. A third empirical regularity motivating our study is that, both before and after SOX, conditional on issuing a favorable accounting statement, a CEO receives compensation that is on average higher but also more volatile.

We develop a dynamic principal agent model to explain why. Each period a CEO agent has private information about his firm and takes hidden actions, neither of which is observed by the shareholder principal. In the model, accounting disclosures are treated as unverifiable discretionary messages sent by the agent to the principal about the state of the firm at the beginning of the period. Our data show that compensation practice quickly adapted to the new regulations, and our model reflects this feature: the optimal long-term contract can be implemented by a sequence of short-term contracts. The optimal contract does not base compensation on abnormal returns alone (as in a pure moral hazard model), but also incorporates accounting disclosures. In equilibrium, expected compensation is higher in the good accounting state than the bad one and there is also greater variation in compensation outcomes in the good state – which are two of the empirical regularities mentioned above. In the model, CEOs are paid to reveal the good state with the promise of receiving very high compensation if the firm produces abnormally high returns. This prediction contrasts with those of a pure moral hazard model, which does not predict either empirical regularity.

We identify and estimate the model using data on compensation, abnormal returns, and accounting disclosure, controlling for different firm categories. The risk-aversion parameter of the agent in our model is set identified, and the remaining parameters of the model are identified up to the value of the risk-aversion parameter for each of the firm categories. Even though our model is semiparametric and does not impose enough functional form assumptions to achieve point identification, we find much of the variation in the data is explained by the primitives of the model and the returns process.

The variation in our data can be accounted for without resorting to an explanation based on changing tastes: we do not find evidence that the preference for risk taking by CEOs changed with SOX, contradicting concerns raised by directors (Cohen et al., 2013) and politicians such as Alan Greenspan and William Donaldson (former SEC chairman) that CEOs would overreact to provisions in SOX provisions and exercise undue caution in investment decisions, thus destroying shareholder value (see Coats and Srinivasan, 2014).

Overall the main impact of SOX was to increase the administrative burden of compliance in the primary sector, but reduce costs in the service sector. These findings, of increased indirect costs from paying a higher compensating differential to CEOs, complement those of Coats and Srinivasan (2014), who document the direct costs from control system expenditures incurred as a result of SOX's new requirements. Despite the intention of SOX to make governance more transparent, we find that SOX increased agency costs within most categories of all three sectors. In the primary sector this is mainly attributable to the higher cost of hidden actions, while in the consumer goods sector the cost of hidden information tended to increase. Implementing SOX also reduced the gross loss shareholders would bear if managers shirked, evidence that legislators were concerned with the potential for large losses rather than their expected value (which takes into account the probability of their occurrence). Finally the compensating differential managers require to align their goals with shareholders increased, presumably because one benefit from shirking is to ignore the obligations imposed by the SOX legislation.

Previous studies have found that firms remaining in the U.S. stock market after SOX are larger than the firms that exit by going private (Engel et al., 2007; Kamar et al., 2009) or going dark (Leuz et al., 2008). Also, these studies found that firms managed their market capitalization to remain below a threshold to remain exempt from SOX section 404 (Gao et al., 2009). However, concerns about the exemptions for firms below the $75 million threshold and non-public companies from SOX section 404 have been raised by both investors (CAQ, 2010) and public companies (Protiviti, 2012) and a higher threshold has been proposed. Our findings lead us to concur with previous advice that there may be scope for relaxing some of the regulations that SOX introduced, either by rolling back its harsher provisions, or tailoring them to the different sectors.

A Appendix

Proof of Lemma 1. In our model, the proof of Proposition 5 in Margiotta and Miller (2000) can be simply adapted to show that Theorem 3 of Fudenberg et al. (1990) applies, thus demonstrating that the long-term optimal contract can be sequentially implemented. An induction completes the proof by establishing that the sequential contract implementing the optimal long-term contract for a CEO who will retire in $\overline{\varsigma}$ periods replicates the one-period optimal contract. In the optimal short-term contract, the participation constraint is satisfied with strict equality, which implies that at the beginning of period $\overline{\varsigma} - 1$ the expected lifetime utility of the CEO is determined by setting $t = \overline{\varsigma} - 1$ in the equation

$$-b_t \exp\left(-\tfrac{a_t+\gamma_t e_t}{b_t}\right). \tag{14}$$

Suppose that at the beginning of all periods $t \in \{\varsigma+1, \tau+2, \ldots, \overline{\varsigma}-1\}$; the expected lifetime utility of the CEO is given by Equation (14). We first show the expected lifetime utility of the CEO at ς is also given by Equation (14). From Lemma 3.1 in the main text, the problem shareholders solve at ς is identical to the short-term optimization problem solved in the text. In the solution to each cost-minimization subproblem for the four (L_{1t}, L_{2t}) choices, the CEO's participation constraint is met with equality. Consequently, the CEO achieves the expected lifetime utility given by Equation (14), as claimed. Therefore the problem of participating at time ς and possibly continuing with the firm for more than one period reduces to the problem of participating at time ς for one period at most, solved in Lemma 2. The induction step now follows. ■

Proof of Lemma 2. Let $\lambda_{t'}$ be the date-t price of a contingent claim made on a consumption unit at date t', implying the bond price is defined as $b_t \equiv E_t\left[\sum_{t'=t}^{\infty}\lambda_{t'}\right]$, and let q_t denote the date-t price of a security that pays off the random quantity $q_t \equiv E_t\left[\sum_{t'=t}^{\infty} \lambda_{t'}\left(\ln \lambda_{t'} - t' \ln \beta\right)\right]$. From Equation (15) on page 680 of Margiotta and Miller (2000), the value to a CEO with current wealth endowment e_{nt} of announcing state $r_t(s)$ in period t when the true state is s and choosing effort level l_{st2} in anticipation of compensation $w_{r_t(s)t}(x)$ at the beginning of period $t+1$ when he retires one period later is

$$-b_t \alpha_{2t}^{1/b_t} \left\{ E_t \left[\exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) \right] \right\}^{1-1/b_t} \exp\left(-\frac{q_t + \gamma_t e_{nt}}{b_{t+1}} \right).$$

The corresponding value from choosing effort level l_{st1} is

$$-b_t \alpha_{1t}^{1/b_t} \left\{ E_t \left[\exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) \left[g_{st}(x) \right] \right] \right\}^{1-1/b_t} \exp\left(-\frac{q_t + \gamma_t e_{nt}}{b_{t+1}} \right),$$

whereas from their Equation (8) on page 678 of Margiotta and Miller (2000), the value from retiring immediately is $-b_t \exp\left(-\frac{q_t + \gamma_t e_{nt}}{b_{t+1}} \right)$. Dividing each expression through by the retirement utility, it immediately follows that the CEO chooses $l_{st} \equiv (l_{t0}, l_{st1}, l_{st2})$ to minimize the negative of expected utility:

$$l_{t0} + \left\{ (\alpha_{1t} l_{st1} + \alpha_{2t} l_{st2})^{1/(b_t - 1)} E_t \left[\exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) \left[g_{st}(x) l_{st1} + l_{st2} \right] \right] \right\}^{(b_t - 1)/b_t}.$$

Since $l_{t0} \in \{0, 1\}$ and $b_t > 1$, the solution to this optimization problem also solves

$$l_{t0} + (\alpha_1 l_{st1} + \alpha_2 l_{st2})^{1/(b_t - 1)} E_t \left[\exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) \left[g_{st}(x) l_{st1} + l_{st2} \right] \right].$$

Summing over the two states $s \in \{1, 2\}$ yields the minimand in Lemma 2. ■

References

Antle, Rick and Abbie Smith. 1985. "Measuring Executive Compensation: Methods and an Application." *Journal of Accounting Research*, *23* (1): 296–325.

Antle, Rick and Abbie Smith. 1986. "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives." *Journal of Accounting Research*, *24* (1): 1–39.

Bargeron, Leonce L., Kenneth M. Lehn, and Chad J. Zutter. 2010. "Sarbanes-Oxley and Corporate Risk-taking." *Journal of Accounting and Economics*, *49* (1-2): 34–52. Conference Issue on Current Issues in Accounting & Reassessing the Regulation of Capital Markets.

Bolton, Patrick and Mathias Dewatripont. 2005. *Contract Theory*, The MIT Press.

Carter, Mary Ellen, Luann J. Lynch, and Sarah L. C. Zechman. 2009. "Changes in Bonus Contracts in the Post-Sarbanes-Oxley Era." *Review of Accounting Studies*, *14* (4): 480–506.

Center for Audit Quality, CAQ. 2010. "The CAQ's Fourth Annual Individual Investor Survey."

Coates, John C and Suraj Srinivasan. 2014. "SOX after Ten Years: A Multidisciplinary Review." *Accounting Horizons*, *28* (3): 627–671.

Cohen, Daniel A., Aiyesha Dey, and Thomas Z. Lys. 2007. "The Sarbanes Oxley Act of 2002: Implications for Compensation Contracts and Managerial Risk-Taking." *working paper.*

Cohen, Daniel A., Aiyesha Dey, and Thomas Z. Lys. 2008. "Real and Accrual-Based Earnings Management in the Pre- and Post Sarbanes Oxley Periods." *The Accounting Review*, *83* (3): 757–787.

Cohen, Jeffrey R, Colleen Hayes, Ganesh Krishnamoorthy, Gary S Monroe, and Arnold M Wright. 2013. "The Effectiveness of SOX Regulation: An Interview Study of Corporate Directors." *Behavioral Research in Accounting*, *25* (1): 61–87.

Engel, Ellen, Rachel M. Hayes, and Xue Wang. 2007. "The Sarbanes–Oxley Act and Firms' Going-private Decisions." *Journal of Accounting and Economics*, *44* (1): 116–145.

Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom. 1990. "Short-term Contracts and Long-term Agency Relationships." *Journal of Economic Theory*, *51* (1): 1–31.

Gao, Feng, Joanna Shuang Wu, and Jerold Zimmerman. 2009. "Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes-Oxley Act." *Journal of Accounting Research*, *47* (2): 459–506.

Gayle, George-Levi and Robert A. Miller. 2009. "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" *American Economic Review*, *99* (5): 1740–1769.

Gayle, George-Levi and Robert A Miller. 2009. "Insider Information and Performance Pay." *CESifo Economic Studies*, *55* (3-4): 515–541.

Gayle, George-Levi and Robert A. Miller. 2015. "Identifying and Testing Models of Managerial Compensationn." *Review of Economic Studies*, *82* (3): 1074–1118.

Gayle, George-Levi, Limor Golan, and Robert A. Miller. 2015. "Promotion, Turnover and Compensation in the Executive Labor Market." *Econometrica*, *83* (6): 2293–2369.

Hall, Brian J. and Jeffrey B. Liebman. 1998. "Are CEOs Really Paid Like Bureaucrats?" *The Quarterly Journal of Economics*, *103* (3): 653–691.

Hochberg, Yael V, Paola Sapienza, and Annette Vissing-Jørgensen. 2009. "A Lobbying Approach to Evaluating the Sarbanes-Oxley Act of 2002." *Journal of Accounting Research*, *47* (2): 519–583.

Holmstrom, Bengt and Steven N. Kaplan. 2003. "The State of U.S. Corporate Governance." *European Corporate Governance Institute Finance Working Paper*, *Sepetember* (23).

Jain, Pankaj K. and Zabihollah Rezaee. 2006. "The Sarbanes-Oxley Act of 2002 and Capital-Market Behavior: Early Evidence." *Contemporary Accounting Research*, *23* (3): 629–654.

Kamar, Ehud, Pinar Karaca-Mandic, and Eric Talley. 2009. "Going-Private Decisions and the Sarbanes-Oxley Act of 2002: A Cross-Country Analysis." *Journal of Law, Economics, and Organization*, *25* (1): 107–133.

Kang, Qiang, Qiao Liu, and Rong Qi. 2010. "The Sarbanes-Oxley Act and Corporate Investment: A Structural Assessment." *Journal of Financial Economics*, *96* (2): 291–305.

Leuz, Christian, Alexander Triantis, and Tracy Yue Wang. 2008. "Why Do Firms Go Dark? Causes and Economic Consequences of Voluntary SEC Deregistrations." *Journal of Accounting and Economics*, *45* (2): 181–208.

Li, Qi and Jeffrey Scott Rachine. 2007. *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

Litvak, Kate. 2007. "The Effect of the Sarbanes-Oxley Act on Non-US Companies Cross-listed in the US." *Journal of Corporate Finance*, *13* (2): 195–228.

Margiotta, Mary M. and Robert A. Miller. 2000. "Managerial Compensation and the Cost of Moral Hazard." *International Economic Review*, *41* (3): 669–719.

Murphy, Kevin J. 1999. "Chapter 38: Executive Compensation." in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Part B of *Handbook of Labor Economics*, Elsevier, pp. 2485–2563.

Murphy, Kevin J. 2013. "Chapter 4: Executive Compensation: Where We Are, and How We Got There." in Rene M. Stulz George M. Constantinides, Milton Harris, ed., *Handbook of the Economics of Finance: Corporate Finance*, 1 ed., Elsevier Science North Holland: Amsterdam, pp. 211–347.

Nekipelov, Denis. 2010. "Empirical Content of a Continuous-Time Principal-Agent Model: The Case of the Retail Apparel Industry." *Working paper.*

Peng, Lin and Ailsa Röell. 2008. "Manipulation and Equity-based Compensation." *American Economic Review*, *98* (2): 285–90.

Protiviti. 2012. "Sarbanes-Oxley Compliance Survey 2012."

Zhang, Ivy Xiying. 2007. "Economic Consequences of the Sarbanes-Oxley Act of 2002." *Journal of Accounting and Economics*, *44* (1-2): 74–115.

Table 1: Firm Characteristics and Compensation

		Primary			Consumer			Service		
		Pre	Post	t/F-stat	Pre	Post	t/F-stat	Pre	Post	t/F-stat
Total Asset		4602	6910	7.4	3637	5126	4.2	14009	19967	3.6
		(7713)	(11783)	0.4	(8872)	(12010)	0.5	(44521)	(76742)	0.3
Debt/Equity		1.810	2.113	4.5	1.561	1.530	-0.5	3.580	2.913	-6.2
		(1.437)	(2.548)	0.3	(1.627)	(2.033)	0.6	(5.204)	(4.368)	1.4
Accounting Return		1.100	1.126	4.0	1.115	1.085	-3.6	1.164	1.092	-11.1
		(0.236)	(0.221)	1.1	(0.282)	(0.253)	1.2	(0.343)	(0.252)	1.8
		Overall			Bad			Good		
Sector	(A,C)	Pre	Post	t/F-stat	Pre	Post	t/F-stat	Pre	Post	t/F-stat
Primary	(S,S)	2190	4047	2.8	521	1526	1.3	4096	7206	2.7
		(12452)	(13908)	0.8	(8244)	(12516)	0.4	(15743)	(14913)	1.1
	(S,L)	1467	4120	3.9	-29	2432	4.0	3453	6299	2.3
		(8678)	(8805)	1.0	(7356)	(5528)	1.8	(9835)	(11433)	0.7
	(L,S)	4771	8131	3.3	3429	6849	2.7	6501	9607	2.0
		(12224)	(15201)	0.6	(10015)	(14023)	0.5	(14415)	(16390)	0.8
	(L,L)	4546	8460	5.7	3368	6586	3.9	6015	10569	4.1
		(11860)	(15407)	0.6	(10847)	(13337)	0.7	(12867)	(17228)	0.6
Consumer Goods	(S,S)	2377	2821	0.5	-1203	-1594	-0.4	7301	7494	0.1
		(21723)	(18327)	1.4	(15811)	(11796)	1.8	(27128)	(22425)	1.5
	(S,L)	1712	2565	0.8	-567	65	0.5	4494	4528	0.0
		(13992)	(11803)	1.4	(11016)	(8268)	1.8	(16528)	(13700)	1.5
	(L,S)	6450	8580	0.9	1569	2279	0.3	12432	13560	0.3
		(31513)	(31025)	1.0	(23109)	(26279)	0.8	(38644)	(33589)	1.3
	(L,L)	7595	10984	2.0	3858	5588	0.9	12923	16759	1.4
		(26364)	(31235)	0.7	(21787)	(25096)	0.8	(31014)	(35861)	0.7
Service	(S,S)	3408	4078	0.9	454	558	0.1	6929	7782	0.7
		(20488)	(18998)	1.2	(14951)	(13127)	1.3	(25125)	(23095)	1.2
	(S,L)	3384	4416	0.7	1814	2600	0.6	5145	6497	0.5
		(16394)	(18797)	0.8	(13211)	(12876)	1.1	(19218)	(23743)	0.7
	(L,S)	10890	11573	0.3	5351	5502	0.1	18610	17536	-0.3
		(38668)	(34641)	1.2	(30923)	(29914)	1.1	(46350)	(37877)	1.5
	(L,L)	9812	10678	0.9	6666	6964	0.3	14321	16451	1.1
		(26134)	(26185)	1.0	(22752)	(19732)	1.3	(29764)	(33072)	0.8

Note: In the columns "Pre" and "Post", standard deviation is listed in parentheses below the corresponding mean. The columns "t/F-stat" report the statistics of a two-sided t-test on equal mean with critical value equal to 1.96 at 5% confidence level, and the one-sided F-test on equal variance with critical value equal to 1. Firm type is measured by the coordinate-pair (A, C), where A is assets and C is the debt–equity ratio with each corresponding to whether that element is above (L) or below (S) its industry median. Accounting return is classified as "Good (Bad)" if it is greater(less) than the industry average. Assets (Compensation) is measured in millions (thousands) $ US.

Table 2: Nonparametric Tests

A: Test on PDF of Gross Abnormal Returns

Sector (A,C)	Primary Bad	Primary Good	consumer Bad	consumer Good	Service Bad	Service Good
(S,S)	18.05	10.34	12.51	12.39	14.25	14.55
(S,L)	5.88	5.02	1.26	2.27	14.70	5.29
(L,S)	3.29	4.16	3.74	2.03	9.01	19.69
(L,L)	29.46	8.57	9.03	8.68	71.68	29.56

B: Test on Contract Shape

Sector Firm Type	Primary Bad	Primary Good	Consumer Bad	Consumer Good	Service Bad	Service Good
(S,S)	10.06	1.58	2.89	1.09	1.54	1.47
(S,L)	6.82	6.45	3.30	1.71	4.08	6.85
(L,S)	19.67	7.34	5.51	3.52	5.66	8.74
(L,L)	10.32	23.38	3.69	6.74	7.37	10.65

Firm type is measured by the coordinate-pair (A, C), where A is assets and C is the debt–equity ratio with each corresponding to whether that element is above (L) or below (S) its industry median. Accounting return is classified as "Good (Bad)" if it is greater(less) than the industry average. Both tests are one-sided test and both statistics follow a standard normal distribution $N(0,1)$.

Table 3: Administrative Cost

(measured in thousands of 2006 US$)

		A: Full Sample	B: Full Sample		C: Restricted Sample	
Sector	(A,C)	τ_1	$\Delta\tau_1$		$\Delta\tau_1$	
Primary	(S,S)	(2191, 2330)	(1398, 1406)	+	(2285, 2455)	+
	(S,L)	(1341, 1404)	(2570, 2582)	+	(3182, 3209)	+
	(L,S)	(4466, 4599)	(2780, 2910)	+	(4113, 4648)	+
	(L,L)	(4311, 4398)	(2921, 3012)	+	(2829, 3165)	+
Consumer	(S,S)	(1764, 2258)	(-679, -555)	-	(-437, 31)	=
	(S,L)	(1235, 1363)	(949, 972)	+	(-25, 110)	=
	(L,S)	(4393, 5072)	(1724, 2113)	+	(-1041, 590)	=
	(L,L)	(6708, 7353)	(-214, -130)	-	(-767, -389)	-
Service	(S,S)	(2639, 2990)	(1036, 1065)	+	(-1473, -1153)	-
	(S,L)	(2641, 2875)	(-76, -29)	-	(-462, -112)	-
	(L,S)	(9803, 10689)	(-1756, -1624)	-	(-4129, -3888)	-
	(L,L)	(9441, 9949)	(-1240, -1218)	-	(-1738, -1262)	-

Note: $\tau_1 \equiv \gamma^{-1} \frac{b_{t+1}}{b_t - 1} \ln \alpha_{2,pre}$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price, 16.4.

Table 4: Aggregate Agency Costs
(measured in thousands of 2006 US$)

		A: Full Sample	B: Full Sample		C: Restricted Sample	
Sector	(A,C)	τ_2	$\Delta\tau_2$		$\Delta\tau_2$	
Primary	(S,S)	(72, 210)	(3, 11)	+	(20, 190)	+
	(S,L)	(33, 96)	(-18, -6)	-	(3, 30)	+
	(L,S)	(68, 201)	(68, 198)	+	(76, 611)	+
	(L,L)	(45, 132)	(47, 137)	+	(43, 379)	+
Consumer	(S,S)	(259, 753)	(-189, -65)	-	(-527, -59)	-
	(S,L)	(67, 195)	(13, 37)	+	(21, 156)	+
	(L,S)	(357, 1036)	(192, 582)	+	(182, 1812)	+
	(L,L)	(340, 985)	(55, 140)	+	(81, 459)	+
Service	(S,S)	(183, 534)	(-44, -15)	-	(-360, -41)	-
	(S,L)	(122, 356)	(24, 71)	+	(45, 395)	+
	(L,S)	(460, 1345)	(95, 227)	+	(113, 355)	+
	(L,L)	(265, 772)	(10, 32)	+	(53, 529)	+

Note: $\tau_2 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}\left[w_{s,pre}(x)\right] - \tau_1$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative). The confidence region is estimated for the single common bond price, 16.4.

Table 5: Welfare Costs of Moral Hazard
(measured in thousands of 2006 US$)

		A: Full Sample	B: Full Sample		C: Restricted Sample	
Sector	(A,C)	τ_3	$\Delta\tau_3$		$\Delta\tau_3$	
Primary	(S,S)	(172, 502)	(97, 188)	+	(228, 1532)	+
	(S,L)	(60, 144)	(-1, 38)	=	(-39, 165)	=
	(L,S)	(6, 184)	(154, 532)	+	(96, 1774)	+
	(L,L)	(301, 689)	(-118, -65)	-	(265, 380)	+
Consumer	(S,S)	(1060, 3294)	(423, 533)	+	(-4387, -600)	-
	(S,L)	(149, 479)	(-274, -121)	-	(-817, -202)	-
	(L,S)	(949, 2575)	(624, 1320)	+	(-3111, -649)	-
	(L,L)	(1478, 3606)	(-2868, -1161)	-	(-3848, -332)	-
Service	(S,S)	(358, 1017)	(125, 229)	+	(-328, -150)	-
	(S,L)	(285, 655)	(61, 150)	+	(-399, 268)	=
	(L,S)	(1356, 3381)	(-2417, -603)	-	(-6438, 433)	=
	(L,L)	(769, 2327)	(-1445, -381)	-	(-2621, 479)	=

Note: $\tau_3 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}[y_{s,pre}(x)] - \tau_1$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price, 16.4.

Table 6: Welfare Costs of Hidden Information
(measured in thousands of 2006 US$)

Sector	(A,C)	A: Full Sample τ_4	B: Full Sample $\Delta\tau_4$		C: Restricted Sample $\Delta\tau_4$	
Primary	(S,S)	(-292, -101)	(-177, -94)	-	(-1342, -208)	-
	(S,L)	(-48, -27)	(-56, -5)	-	(-135, 42)	=
	(L,S)	(17, 62)	(-334, -86)	-	(-1163, -20)	-
	(L,L)	(-558, -257)	(112, 255)	+	(-227, 87)	=
Consumer	(S,S)	(-2541, -802)	(-674, -488)	-	(540, 3860)	+
	(S,L)	(-284, -82)	(135, 312)	+	(223, 973)	+
	(L,S)	(-1539, -592)	(-738, -432)	-	(831, 4923)	+
	(L,L)	(-2621, -1139)	(1217, 3008)	+	(413, 4307)	+
Service	(S,S)	(-483, -175)	(-273, -140)	-	(-32, 217)	=
	(S,L)	(-299, -163)	(-126, 10)	=	(-218, 795)	=
	(L,S)	(-2036, -896)	(698, 2644)	+	(-320, 6788)	=
	(L,L)	(-1554, -504)	(391, 1477)	+	(-348, 3150)	=

Note: $\tau_4 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre} \left[w_{s,pre}(x) - y_{s,pre}(x) \right]$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price, 16.4.

Table 7: Gross Losses to Firms from Shirking
(measured in percentage)

Sector	(A,C)	A: Full Sample ρ_1	B: Full Sample $\Delta\rho_1$		C: Restricted Sample $\Delta\rho_1$	
Primary	(S,S)	(11.94, 12.07)	(-0.59, -0.51)	-	(-2.69, -1.96)	-
	(S,L)	(12.53, 12.90)	(-5.40, -5.08)	-	(-6.92, -4.75)	-
	(L,S)	(10.21, 10.68)	(-2.57, -2.49)	-	(-2.82, -2.10)	-
	(L,L)	(5.90, 6.09)	(-1.02, -0.94)	-	(-1.96, -1.95)	-
Consumer	(S,S)	(18.30, 18.51)	(-9.33, -9.25)	-	(-9.16, -8.72)	-
	(S,L)	(10.60, 10.61)	(11.81, 12.70)	+	(2.12, 12.21)	+
	(L,S)	(9.15, 10.03)	(-1.43, -0.95)	-	(-0.40, 1.54)	-
	(L,L)	(7.52, 8.13)	(-2.84, -2.02)	-	(-2.68, -2.11)	-
Service	(S,S)	(17.44, 17.57)	(-3.80, -3.43)	-	(-8.93, -6.34)	-
	(S,L)	(12.99, 13.74)	(-6.58, -5.92)	-	(-3.02, -1.03)	-
	(L,S)	(18.61, 18.91)	(-12.46, -11.48)	-	(-16.59, -15.37)	-
	(L,L)	(9.96, 10.73)	(-7.06, -6.80)	-	(-5.97, -5.07)	-

Note: $\rho_1 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre} \left\{ x \left[1 - g_{s,pre}(x) \right] \right\}$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price (16.4).

Table 8: Compensating Differential from Shirking versus Working
(measured in thousands of 2006 US$)

Sector	(A,C)	A: Full Sample ρ_2	B: Full Sample $\Delta\rho_2$		C: Restricted $\Delta\rho_2$	
Primary	(S,S)	(2792, 2995)	(51, 125)	+	(122, 221)	+
	(S,L)	(1587, 1668)	(31, 51)	+	(-57, -24)	-
	(L,S)	(2303, 2471)	(1085, 1163)	+	(1716, 2125)	+
	(L,L)	(1983, 2073)	(364, 472)	+	(100, 380)	+
Consumer	(S,S)	(7671, 8438)	(-1992, -1766)	-	(-3213, -2091)	-
	(S,L)	(2224, 2400)	(317, 368)	+	(287, 476)	+
	(L,S)	(5353, 6165)	(1623, 1869)	+	(18, 792)	+
	(L,L)	(4687, 5465)	(-2264, -2218)	-	(-1078, -654)	-
Service	(S,S)	(4608, 5074)	(50, 87)	+	(-780, -487)	-
	(S,L)	(2542, 2840)	(23, 97)	+	(67, 446)	+
	(L,S)	(8758, 9929)	(-6960, -6722)	-	(-7697, -5721)	-
	(L,L)	(5916, 6610)	(-3985, -3885)	-	(-2041, -1985)	-

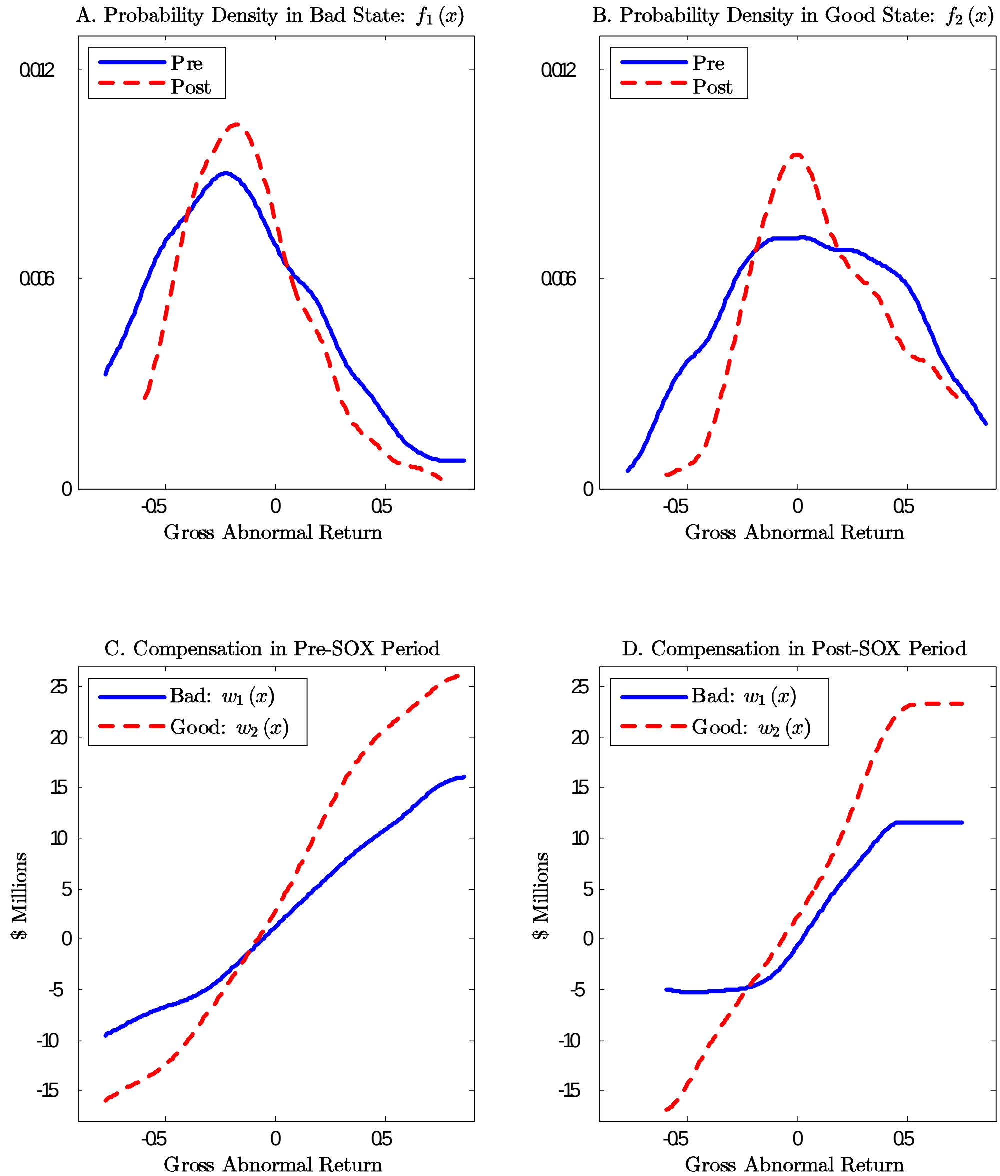
Note: $\rho_2 \equiv b_{t+1} \left[(b_t - 1)\,\gamma\right]^{-1} \ln(\alpha_{2,pre}/\alpha_{1,pre})$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price, 16.4.

Table 9: Change in Welfare Costs of Moral Hazard Caused by Signal Quality
(measured in thousands of 2006 US$)

Sector	(A,C)	A: Full Sample $\tau_3(\alpha_{j,post}, g_{s,pre})$	B: Full Sample ρ_3		C: Restricted Sample ρ_3	
Primary	(S,S)	(179, 551)	(90, 139)	+	(216, 1261)	+
	(S,L)	(62, 153)	(-3, 29)	=	(-37, 198)	=
	(L,S)	(75, 468)	(85, 248)	+	(-617, -151)	-
	(L,L)	(429, 973)	(-401, -194)	-	(59, 225)	+
Consumer	(S,S)	(548, 1723)	(936, 2025)	+	(180, 1254)	+
	(S,L)	(207, 643)	(-438, -179)	-	(-1575, -272)	-
	(L,S)	(1628, 4784)	(-889, -55)	-	(-3202, -930)	-
	(L,L)	(542, 839)	(-250, -101)	-	(-224, 1181)	=
Service	(S,S)	(366, 1061)	(117, 185)	+	(-130, 398)	=
	(S,L)	(302, 667)	(50, 133)	+	(-503, 212)	=
	(L,S)	(143, 184)	(609, 782)	+	(1503, 24282)	+
	(L,L)	(107, 250)	(282, 632)	+	(724, 4470)	+

Note: $\rho_3 \equiv \tau_3(\alpha_{j,post}, g_{s,post}) - \tau_3(\alpha_{j,post}, g_{s,pre})$. Here "+" ("-") means we cannot reject the null hypothesis that the change is positive (negative), while "=" means we cannot reject the null hypothesis of no change. The confidence region is estimated for the single common bond price, 16.4.

Figure 1: Empirical Compensation Schedule and Excess Return Density



Note: The plots present the nonparametrically estimated density of gross abnormal returns and the optimal compensation of firms with large size and high leverage in the Primary sector. The compensation of both periods is anchored at bond prices equal to 16.5 (b_t) and 16.4 (b_{t+1}).

Figure 2: Welfare Cost Summary

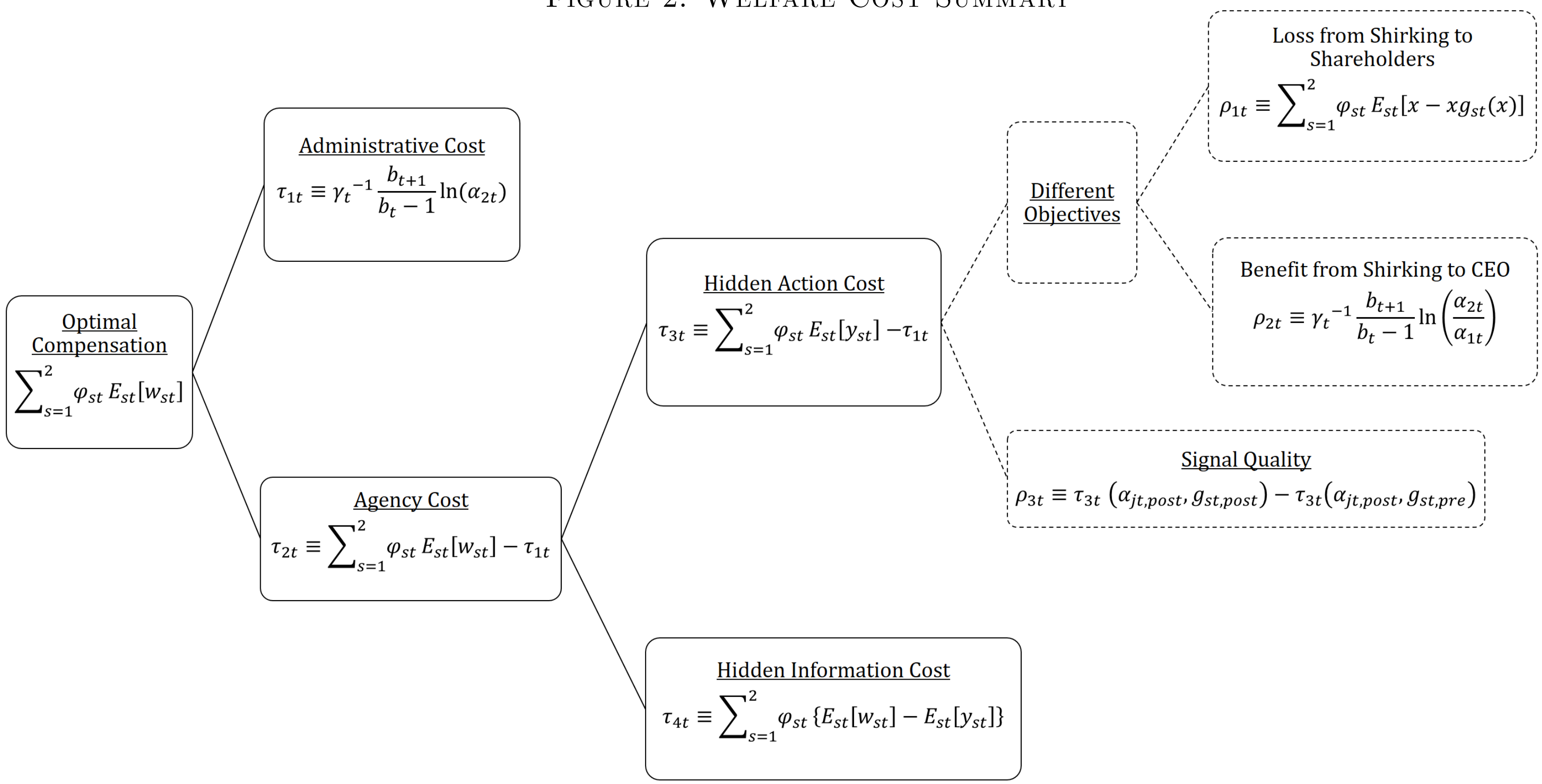
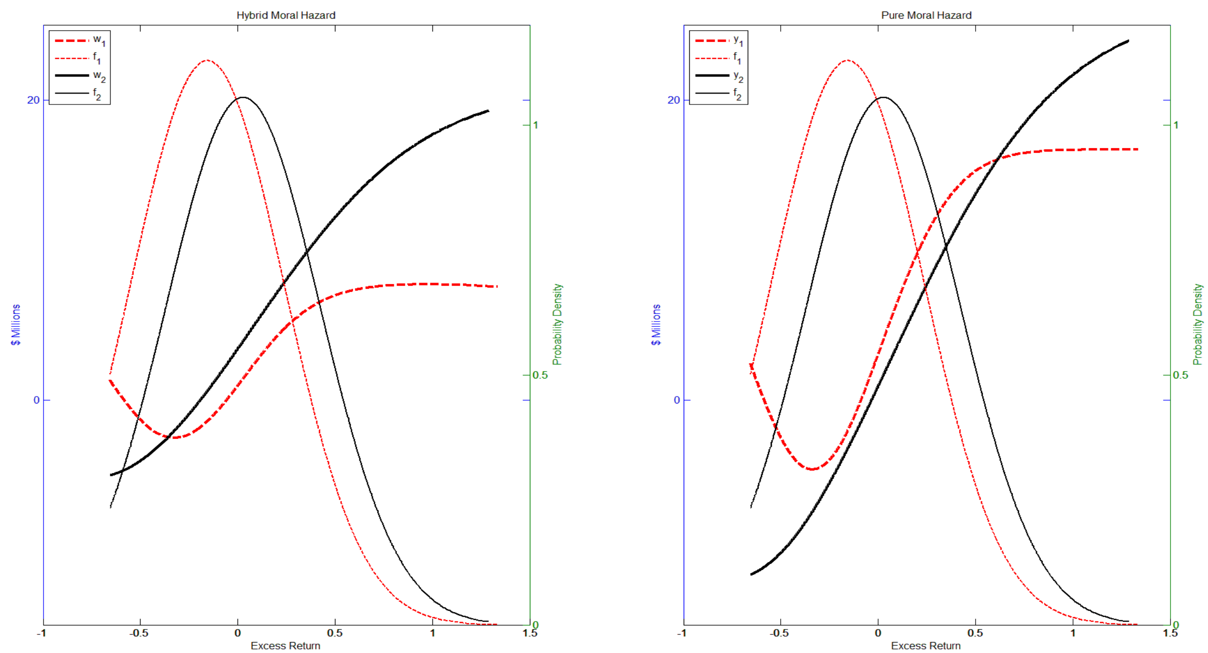


Figure 3: The Optimal Compensation Schedules



Note: The plots use the return and optimal compensation of firms with small size and low leverage in the Primary sector in the Pre-SOX period. The risk aversion parameter γ equals to 0.08. Effort cost coefficient of shirking (α_1) equals to 0.96 and effort cost coefficient of working (α_2) equals to 1.20. Bond prices are 16.5 (b_t) and 16.4 (b_{t+1}). The excess return is approximated by one-side truncated normal distribution $TN(a,\mu,\sigma)$ with truncated point on the left (a), mean (μ), and standard deviation (σ) as follows. Working in bad state: $TN(-0.66,-0.16,0.39)$. Working in good state: $TN(-0.66,0.03,0.39)$. Shirking in bad state: $TN(-0.66,-0.25,0.27)$. Shirking in good state: $TN(-0.66,-0.11,0.36)$. The probability of bad state is 0.54 and the probability of good state is 0.46.