Journal *of* Accounting Research

DOI: 10.1111/1475-679X.12436
Journal of Accounting Research
Vol. No. May 2022
Printed in U.S.A.

CHICAGO BOOTH

Was Sarbanes–Oxley Costly? Evidence from Optimal Contracting on CEO Compensation

GEORGE-LEVI GAYLE,* CHEN LI,† AND ROBERT A. MILLER‡

Received 31 January 2017; accepted 25 February 2022

ABSTRACT

This paper investigates the effects of regulatory interventions on contracting relationships within firms by examining the impacts of the Sarbanes–Oxley (SOX) Act on CEO compensation. Using panel data of the S&P 1500 firms, it quantifies welfare gains from a principal–agent model with hidden information and hidden actions. It finds that SOX: (1) reduced the conflict of interest between shareholders and their CEOs, mainly by reducing shareholder loss from CEOs deviating from their goal of expected value maximization; (2) increased the cost of agency, or the risk premium CEOs are paid to align

*Department of Economics, Washington University in St. Louis and Federal Reserve Bank of St. Louis; †New York University - Shanghai; ‡Tepper School of Business, Carnegie Mellon University

Accepted by Haresh Sapra. We thank participants for their comments at the 2020 AAA-FARS, 2020 AAA-MAS, the 2021 Cowles Conference on Models and Measurement, and in seminars given at Baruch College CUNY, China Europe International Business School, City University of London, University of Calgary, Fudan University, University of Hong Kong, the University of Nottingham, the University of Minnesota, Renmin University of China, and the Wharton School of the University of Pennsylvania,. Li acknowledges financial support from a PSC-CUNY Award, funded by The Professional Staff Congress and The City University of New York. 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. An online appendix to this paper can be downloaded at http://research.chicagobooth.edu/arc/journal-of-accounting-research/online-supplements.

ramiller@andrew.cmu.edu

their interests with those of shareholders; (3) increased administrative costs in the primary sector (which includes utilities and energy) but the effect in the other two broadly defined sectors, services and consumer goods, was more nuanced; and (4) had no effect on the attitude of CEOs toward risk.

JEL codes: C10, C12, C13, J30, J33, M50, M52, M55

Keywords: CEO versus shareholders; hidden information; increased agency costs; principal–agent model; reduced conflict of interest; Sarbanes–Oxley Act; set identification; structural estimation

1. *Introduction*

This paper is an empirical investigation of changes in CEO compensation resulting from the passage of the Sarbanes–Oxley (SOX) Act in 2002. To clarify, the term SOX is often used in the literature to collectively refer to both the SOX Act and contemporary listing rules changed by the NYSE and the NASDAQ. We follow this practice, adopted by Zhang [2007], Carter et al. [2009], Hart [2009], Linck et al. [2009], Bargeron et al. [2010], Cohen et al. [2013], among others. The catalyst for SOX was a failure in corporate governance that led to the dismissal of executives and, in some cases, subsequent prosecution for fraud, conviction, and imprisonment. SOX was not primarily directed toward realigning the incentives of law-abiding managers. These executives violated legal constraints that were subject to auditing. Enacting SOX brought greater accountability to financial statements, more rigorous enforcement of property rights in governance, and higher penalties for fraud, factors that discourage white-collar crime.

SOX discourages managers from breaking the law, thereby strengthening the property rights of shareholders, but the legislation also impacts internal agency issues. It affects incentives that motivate law-abiding managers: to act in the firm's interest, or engage in legal activities they prefer; to disseminate unverifiable financial information to the board accurately, or not. We focus on these agency issues within each firm. Compensation schemes are tools shareholders use to direct managerial behavior, so curbing contracts between shareholders and managers is controversial (Hart [2009]). Linck et al. [2009, p. 3290] remark "SOX signifies a turning point, and its enactment represents a significant inroad by the government into governance." Understanding the ramifications of SOX for CEO behavior and compensation is therefore essential to a thorough evaluation of this intervention.

A CEO benefits from exercising discretion in reporting accounting information that helps determine his or her own compensation. Gayle and Miller [2009b] construct hypothetical portfolios demonstrating that if investors could have exploited information directly inferred from holdings of CEOs with respect to financial securities in their own firms: instead of following a stock market index, their returns would have been about 19%, rather than 9%, without increasing their risk exposure. It is therefore not surprising that the impact of SOX on CEO compensation has attracted the attention of empirical researchers. Carter et al. [2009] find that after SOX was introduced, more weight was placed on positive earnings changes in the CEO bonus, and less weight was placed on salary. Interpreting this shift as firms responding to improved disclosure quality, they find no evidence that CEOs were compensated for bearing more risk after SOX. Cohen et al. [2013] find a decline in pay-performance sensitivity (PPS), increased bonus, and no significant decrease in total compensation. They also find that the PPS in the post-SOX era was lower in firms with less than majority independent boards before SOX. Chhaochharia and Grinstein [2009] report decreased CEO compensation after SOX in firms with less than majority independent directors on their boards in 2002. Yet Guthrie et al. [2012] conclude the compensation committee independence requirement increased CEO compensation after SOX. To motivate our structural econometric model, we add to this body of work by conducting nonparametric tests, reported in section 2, that show the probability distribution of financial returns shifted, along with the empirical patterns relating CEO compensation to both financial and accounting returns.

We interpret these findings as indirect evidence that, when determining CEO compensation, boards use financial returns, as well as discretionary information provided by their CEOs transmitted through accounting income reports, to allay moral hazard and hidden information about the firm's state, and that SOX affected both factors. This paper quantifies, before and after the implementation of SOX: (1) how much the goals of a CEO diverge from those of the shareholders', and (2) the costs shareholders incur to incentivize their CEOs. The data for this study are drawn from observations on CEO compensation in the ExecuComp database, supplemented with data on financial returns and accounting income. We control for differences observed between firms, and for the subsample where the data exist, we distinguish between firms that had complied with SOX regulations before they were legislated versus those that had not. To capture both effects of SOX on the scope for the CEO to hide information from the board, and to act against shareholders' interests without their knowledge, we estimate an optimal contracting model borrowed from Gayle and Miller [2015], in which a risk-neutral principal (the shareholders) and a risk-averse agent (the CEO) are asymmetrically informed about both the actions of the CEO and also the firm's prospects.

Following Margiotta and Miller [2000] and Gayle and Miller [2009a], we use the estimated model to construct two measures that quantify (1), the potential conflict of interest between the goals of shareholders and their CEO, namely: the gross financial loss to shareholders if the CEO is not incentivized; the benefits to CEOs from pursuing their own goals on the job within the firm if their compensation is fixed (i.e., does not depend on firm performance). We find that SOX improved the interest alignment between shareholders and CEOs. Implementing SOX reduced the gross loss to shareholders in almost all the firm categories we investigated. Overall, noncompliant firms faced larger losses before SOX, and also experienced a bigger reduction in losses after SOX, than compliant firms. From the CEO perspective, we measure the conflict of interest by a compensating differential certainty equivalent that equalizes the expected utility from following shareholder interests rather than his or her own goals in employment. The estimated differential is orders of magnitude less than potential shareholder losses; it also varies with firm type, presumably because heterogeneity across production and information technology stimulates a variety of temptations. On this score, we find that implementing SOX tends to homogenize these opportunities; the benefits of deviating from shareholder goals become less diffuse across the diverse firm types.

Although SOX reduced the conflict of interest between shareholders and their CEOs, it was not eliminated. Aligning the goals of owners and managers still required incentives to induce CEOs to follow their interests when their actions are unobserved. On (2), the evidence on whether SOX increased the costs of compensating CEOs is mixed. One component is a compensating differential the CEO is paid relative to his or her outside option when there are no agency issues, that accounts for the nonpecuniary costs and benefits of his position (called administrative costs in this paper). In only one of the three sectors does this component increase for all firm categories. This is the primary sector, which includes energy and utilities, that tend to include highly regulated firms. The finding complements practitioners' concerns and also academic findings on the SOX impacts on regulated firms.[1] The other component, agency costs, is a risk premium measuring the maximal hypothetical amount shareholders would pay to perfectly monitor the CEO. We find that SOX increased agency costs within most firm categories of all three sectors.

We also address other issues that have been raised in connection with effects of SOX on CEOs and the companies they manage. One concern raised by directors (Cohen, Dey, and Lys [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 (Coats and Srinivasan [2014]). If their concerns had proved prescient, then presumably, risk aversion would rise with the onset of SOX. Accordingly we tested the null hypothesis that the preference for risk-taking by CEOs remained constant over the two eras. It is not rejected: attributing the impact of SOX to changes in taste does not seem to provide a fruitful basis for policy analysis. Moreover implementing SOX did not seem to impose a disproportionately greater burden on the smaller firms in our sample, another point of

[1] Filbeck et al. [2011] document that regulated firms did not outperform and even sometimes underperformed nonregulated firms after SOX. Labro and Stice-Lawrence [2020] find that the benefits of an accounting system update triggered by regulatory pressures such as SOX 404 are insignificant in hospitals.

contention about SOX.[2] Here though, we caution against reading too much into this result, because by many standards all the S&P 1500 firms are large.

In addition to the papers cited above, we contribute to three strands of the accounting literature. The most closely related are studies on the SOX consequences for CEO compensation, which focus on the shape of compensation schemes. We complement the literature by quantifying various driving forces behind the changes of CEO compensation. We emphasize the communication role of accounting information potentially played in compensation contracts. Our findings of the pervasive increase of agency costs after SOX complement previous studies that document the direct costs from control system expenditures incurred as a result of SOX's new requirements (Coats and Srinivasan [2014]). Our findings of reduced shareholders' loss from CEO shirking after SOX within most categories of all three sectors complement previous studies reporting mixed results of the effects of SOX passage on shareholders' wealth (see Zhang [2007] and Jain and Rezaee [2006] on the stock returns and Chhaochharia and Grinstein [2007] on long-term firm performances.)

This paper is also closely related to Zakolyukina [2018] and Bertomeu et al. [2017], who estimate structural models of GAAP violations and manipulation. The most attention has been given to disclosure (Beyer et al. [2019], Bertomeu et al. [2020, 2021] Bird et al. [2019], Choi [2021], Breuer and Windisch [2019], Cheynel and Liu-Watts [2020]). On other topics, McClure [2020] studies the tax avoidance by estimating a single-agent dynamic optimization model, and Li [2021] examines the compensation design of top management teams by estimating moral hazard models of multiple agents. Our work distinguishes itself by estimating a principal–agent model of hidden action and hidden information to investigate the policy effect of a sweeping regulatory regime shift through the lens of optimal contracting.

We follow the literature when comparing a designated control group to a treatment group as a way of gauging the impact of SOX. Previous reduced form studies use firms that were compliant with the SOX provisions of board structure before SOX as the control group in their difference-in-differences (DID) design and firms that did not comply with measure until after enactment as the treatment group.[3] They examine the impact of

---

[2] Chhaochharia and Grinstein [2007] find a positive market reaction to the announcement of the new rules in firms that are less compliant with the provisions of the regulations, among which large firms earn positive excess returns, but small firms make negative excess returns. Wintoki [2007] suggests that small firms were adversely affected by SOX. Linck, Netter, and Yang (2009, RFS) also find that overall director costs are larger among small firms.

[3] Alternatively, Iliev [2010] adopts a regression discontinuity (RD) design, designating U.S. firms with a public float lower than \$75 million and foreign firms with a public float lower than \$700 million as control groups. Our sample consists of S&P 1500 firms that are much larger, and there is no obvious cutoff value to separate firms in our sample for the purposes of

SOX on corporate governance (Linck et al. [2009]), CEO compensation (Chhaochharia and Grinstein [2009], Cohen et al. [2013], Guthrie et al. [2012]), investment (Cohen et al. [2013], Banerjee et al. [2015], Lu and Wang [2015]), firm performance (Duchin et al. [2010], Bargerson et al. [2010]), and earnings management (Chen et al. [2015], Joo and Chamberlain [2017]). If SOX affects the primitive in the contracting environment of CEO compensation mainly through changing board structure, then we conjecture noncompliant firms experience more changes in our measures after SOX than the compliant firms.[4] Our empirical analysis demonstrates how a DID research design can be adapted to a structural econometrics framework.

2. *Data*

We motivate the analysis by providing some institutional and legal background for SOX, explaining how the variables for the study were constructed, and showing how SOX affected the distribution of financial returns and compensation.

2.1 INSTITUTIONAL AND LEGAL BACKGROUND

Passed in 2002, SOX is a legislative response to a wave of corporate governance failures at many prominent companies and "the most extensive regulation of the securities markets since the Securities Act of 1933 and the Securities Exchange Act of 1934 (Ball [2009, p. 290])." SOX operated through several channels to change the duties of CEOs and their compensation. First, SOX prescribes some of the CEO's duties. A senior officer of the U.S. Chamber of Commerce, Michael Ryan, expressed concerns that the "time and energy required by SOX can be a distraction . . . [because the] amount of time management is spending on the process to comply with SOX takes them away from running the business, increasing sales and developing new products" (Farrell [2007]). The legislation demands a greater degree of internal control and accountability. Section 302 makes the CEO and CFO responsible for establishing, maintaining, assessing, and disclosing the quality of internal controls of the firm. Section 404(A) requires disclosure of these responsibilities. Section 302 requires the CEO and CFO to certify the integrity of financial statements. For example, the CEO and the CFO are required to certify in each annual or quarterly report filed or submitted that the financial statements and other financial

---

this study. Similarly, the corporate governance and compensation provision of foreign firms could differ substantially from U.S. firms, so they are not used for our control groups either. In addition, RD designs typically face the challenge of generalizing beyond local estimates (Gow et al. [2016]).

[4] Linck et al. [2009] document increased frequency of board meetings, increased independence of board committees, and more financial experts included in boards after SOX. The effects are more pronounced in the firms that were not compliant with the new governance rules before SOX.

information fairly present financial conditions and results and refrain from making misleading statements.

The second channel are the vehicles for CEO compensation. SOX prohibits option backdating and perks, Section 402 also prohibits personal loans to the CEO, option backdating is curbed by Section 304 that enforces the "clawback" penalty, and Section 409(A) restricts deferred compensation in various ways.

Another two channels stem from changes SOX wrought to the relationship between the CEO and the board, which selects the CEO, monitors and advises him in oversight roles, and assists in determining his compensation. SOX places constraints on the composition of the board of directors, by mandating a greater degree of independence: this decree has ramifications for the CEO. Section 301 and the NYSE/NASDAQ listing rules require a majority of the firm board members to be independent, and that the audit, compensation, and nomination committees be fully independent. Finally, indirectly and more subtly, enacting SOX removed some of the board's power to exercise discretion in the detection and punishment of fraud. The criminal justice system concomitantly assumed a more prominent role in these matters, partially replaced internal incentive mechanisms, and made the CEO more accountable to the law.

2.2 CONSTRUCTING THE VARIABLES

To undertake this empirical analysis financial and accounting data on the S&P 1500 were extracted from Compustat, data on executive compensation were taken from ExecuComp and data on the independent members of board and committees were takfrom RiskMetrics. Bond prices were constructed from the yield curve of Treasury bills from the Federal Reserve Economic Dataset. The *extended* sample covers 13 years, 1993–2005 inclusive. The ending year is chosen to make results comparable with previous studies who stop around 2006 to mitigate the effects of later events, including the changes in ExecuComp disclosure in 2006, the financial crisis starting in winter 2007, and the enactment of FAS 123R on the accounting treatment of stock options effective after June 15, 2005 (Linck et al. [2009], Banerjee et al. [2015], Chhaochharia and Grinstein [2009], Carter et al. [2009], Cohen et al. [2013], Merz [2017]).[5] Within the text, our empirical analysis omits data on the two years when legislation was in a flux (2002 and 2003). The decline in interlocking directorships in the years preceding SOX, documented in Gayle et al. [2021], suggests that the legislation might have been anticipated by companies before its implementation. Thus, the *main* sample covers the pre-SOX era as the years 1993–2001 and the

[5] Chhaochharia and Grinstein [2007] provide a detailed time line of reform following the bankruptcy of Enron in December 2001 through the passage of SOX in mid-2002 to the approval by the SEC of the proposals of the NYSE and NASDAQ in November 2003 following internal reviews of their respective corporate governance requirements.

post-SOX era as the years 2004 and 2005. Online appendix A explains how the data were assembled.

As profiles of returns and compensation are likely to differ across firm type, we partition firms along three observed dimensions into twelve categories, by controlling for two firm sizes, two levels of capital structure and three sectors, based on GICS code, comprising the primary, consumer goods, and service sectors. Two practical reasons for limiting the partition to twelve categories stem from the number of observations and literary precedence: our results are comparable with previous studies using similar methods and data.[6] Industries in the primary sector include energy, materials, industrials, and utilities. The consumer goods sector includes consumer discretionary and consumer staples. The service sector includes health care, financial intermediation, information technologies, and telecommunication. Roughly speaking, the consumer goods sector is the least regulated of the three sectors. More generally, the nature of the production technology, including the capital labor ratio, varies across these sectors. These differences are reflected in the nature of the CEO's job, and the complementary firm and sector specific human capital they develop over their careers; this in turn affects job turnover and compensation (Gayle et al. [2012, 2015]).

Firm size is measured by the total assets of firm i at the end of the annual period t, denoted by $Asset_{it}$. We classify each firm by whether its total assets averaged in the pre-SOX era were less than (S) or greater than (L) the median of the averaged total assets for firms in the same sector and whether its debt-to-equity ratio averaged in the pre-SOX era was less or greater than the median of the averaged debt-to-equity ratio for firms in the same industry in the pre-SOX era. There are many reasons for conditioning on firm size. Firm size plays an important role in explaining the evolution of CEO compensation due to moral hazard cost (Gayle and Miller [2009]). Engel et al. [2007] and Leuz et al. [2008] indicate that the SOX provision may have different implications for firms with different sizes. Linck, Netter, and Yang [2009] find that increase of board independence after SOX differs by firm size. Chhaochharia and Grinstein [2007] find a positive market reaction to the announcement of the new rules in firms that are less compliant with the rules' provisions. Large firms earn positive excess returns, but small firms earn negative excess returns. Wintoki [2007] suggests that small firms were adversely affected by SOX. Finally, the cost of information acquisition affects the efficiency of the board, which is correlated with firm size (Duchin et al. [2010]).

The capital structure is captured by the debt-to-equity ratio:

$$C_{it} \equiv Debt_{it}/(Asset_{it} - Debt_{it}),$$

where $Debt_{it}$ denotes debt at the end of the period. As with size, firms are classified by whether C_{it} lies above (L) or below (S) the median for that

[6] For example, Gayle and Miller [2009a, 2009b, 2015], Gayle, Golan, and Miller [2015], and Li [2021].

sector in the pre-SOX era. Theory suggests that capital structure influence the agency problems within firms (Stulz [1990], Zwiebel [1996]).

Another input of our estimation is a proxy for the CEO's private information about his firm's profitability. We define accounting returns by:

$$\widetilde{s}_{it} \equiv (Asset_{it} - Debt_{it} + Dividend_{it}) \big/ \big(Asset_{i,t-1} - Debt_{i,t-1}\big),$$

where $Dividend_{it}$ denotes the total value of dividends (and stock repurchases) paid throughout the preceding financial period. To capture the idea that the CEO can exercise discretion in reporting accounting returns in order to communicate private information about his firm's profitability to the shareholders (if he wishes to do so), we construct a binary variable, $s_{it} \in \{1, 2\}$, where $s_{it} = 1$ (the firm is in a bad state) if the accounting return $\widetilde{s}_{it}$ is lower for firm i than the average for all firms within the same sector, size, and capital structure categories in period t, while $s_{nt} = 2$ (the firm is in a good state) if $\widetilde{s}_{it}$ exceeds the average.[7]

Our measure of total compensation follows Antle and Smith [1985, 1986], Hall and Liebman [1998], and Margiotta and Miller [2000]. It includes salary, bonus, options, promised retirement benefits, restricted stocks, as well as the change in wealth attributable to holding financial securities in the firm rather than a fully diversified portfolio.[8] In this way, executive compensation depends directly on the excess returns of the firms they manage, denoted by x_{it}, and computed net of the financial return from holding the market portfolio but before CEO compensation is deducted.

Table 1 summarizes firm characteristics and compensation by type, the latter by accounting state as well. All dollar measures are scaled to $U.S. 2006. The top panel shows firms in the service sector have more valuable assets and are more highly leveraged than the other two, all three sectors experienced significant growth, while the debt-to-equity ratio rose (significantly) in the primary sector but fell in the service sector. Turning to the bottom panel, average compensation is: higher in large firms, although the difference is fully explained by the risk premium (Gayle et al [2015]); higher in the service sector than the other two for most types and accounting states; negative for several firm types in the bad state, reflecting losses CEOs incur when the value of the stocks and options they hold in their own firms suffer sufficiently steep losses; lower in the bad state (for each firm type), and with the implementation of SOX rose significantly in the primary sector but not the other two. With a single exception, highly leveraged large firms in the service sector in the good state, the variance of the

[7] Treating accounting earnings as a continuous variable for modeling how a CEO communicates unverifiable information to the shareholder, rather than a binary variable, might allow future research to extract more information from these data, and prove useful in explaining how the CEO reveals private information.

[8] Options are valued using the Black–Scholes formula. This construct overlooks the difference between granting and vesting financial securities, in effect assuming that all granted securities will be vested.

TABLE 1

Firm Characteristics and Compensation

		Primary Sector			Consumer Sector			Service Sector		
		Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat
Total assets		4,460	6,971	6.7	3,460	5,091	4.0	13,678	20,311	3.3
		(7,436)	(11,549)	(0.4)	(8,242)	(11,358)	(0.5)	(42,619)	(78,265)	(0.3)
Debt-to-equity		1.792	2.092	3.5	1.559	1.498	−0.9	3.651	2.798	−6.9
		(1.428)	(2.646)	(0.3)	(1.589)	(1.779)	(0.8)	(5.273)	(4.231)	(1.6)
Accounting return		1.110	1.135	3.2	1.117	1.070	−4.7	1.180	1.090	−11.7
		(0.233)	(0.231)	(1.0)	(0.283)	(0.258)	(1.2)	(0.342)	(0.258)	(1.8)
Total Compensation										
		Overall			Bad			Good		
		Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat
	(S, S)	1,837	4,704	3.6	387	2,268	2.1	3,593	8,600	3.5
		(12,126)	(13,797)	(0.8)	(8,132)	(12,283)	(0.4)	(15,480)	(15,174)	(1.0)
	(S, L)	1,099	4,675	4.7	−131	3,173	4.5	2,818	6,927	2.4
Primary		(7,970)	(8,906)	(0.8)	(7,520)	(5,679)	(1.8)	(8,267)	(11,970)	(0.5)
	(L, S)	4,300	10,074	4.4	3,225	8,525	3.3	5,772	11,924	2.9
		(11,983)	(16,369)	(0.5)	(9,827)	(15,074)	(0.4)	(14,305)	(17,717)	(0.7)
	(L, L)	4204	8362	5.2	3147	7106	4.3	5595	10115	3.2
		(11571)	(14816)	(0.6)	(10662)	(12921)	(0.7)	(12540)	(17003)	(0.5)
	(S, S)	1425	2151	0.6	−1874	−2903	−0.9	6056	7875	0.9
		(21403)	(18474)	(1.3)	(15927)	(12315)	(1.7)	(26629)	(22279)	(1.4)
	(S, L)	1702	2355	0.4	−691	−1516	−0.4	4754	5120	0.1

(*Continued*)

TABLE 1—*(Continued)*

		Total Compensation								
		Overall			Bad			Good		
		Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat	Pre	Post	t-/F-Stat
Consumer		(14362)	(13605)	(1.1)	(11263)	(9682)	(1.4)	(17079)	(15316)	(1.2)
goods	(L, S)	6074	6793	0.2	1355	1399	0.0	12012	10903	−0.2
		(32090)	(31953)	(1.0)	(23206)	(24290)	(0.9)	(39848)	(36345)	(1.2)
	(L, L)	7297	9015	1.0	3460	4590	0.6	12710	13283	0.2
		(26565)	(28798)	(0.9)	(21908)	(22864)	(0.9)	(31228)	(33068)	(0.9)
	(S, S)	3757	2149	−1.9	590	−953	−1.8	7766	5700	−1.3
		(21304)	(18728)	(1.3)	(15543)	(14180)	(1.2)	(26350)	(22352)	(1.4)
	(S, L)	3311	4318	0.6	2041	1725	−0.2	4706	7579	0.9
Service		(16672)	(19656)	(0.7)	(13242)	(12581)	(1.1)	(19696)	(25703)	(0.6)
	(L, S)	11231	7135	−1.6	6065	205	−2.1	18877	14515	−1.0
		(38738)	(32155)	(1.5)	(31249)	(25182)	(1.5)	(46727)	(36918)	(1.6)
	(L, L)	9438	9185	−0.2	6383	6221	−0.1	14114	14365	0.1
		(26040)	(24173)	(1.2)	(22689)	(19051)	(1.4)	(29875)	(30536)	(1.0)

In the columns "Pre" and "Post" indicating the pre- and post- SOX eras, 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 the 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–to-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) of 2006 U.S. dollars.

compensation within each type either declined from the pre- to post-SOX eras, or did not increase.

2.3 TESTING FOR STRUCTURAL CHANGE

CEO compensation depends on excess returns. Therefore, a structural shift occurs if the distribution of excess returns changes and/or the relationship between excess returns and CEO compensation changes. Here we test for equality between the pre- and post-SOX eras of the probability density functions for excess returns and the compensation schedule's shape. These tests make use of point-wise information and thus are expected to reflect more information in the data than a test on the mean.

Change in the distribution of excess returns. Denote by Z the Cartesian product of the 12 firm categories, formed from three sectors, two firm sizes, and two capital structures, plus the two accounting states, formed from $s_{it} \in \{1, 2\}$. Let $f_{pre}(x_{it}|z_{it})$ denote the probability density function of excess returns in the pre-SOX era conditional on $z_{it} \in Z$, and define $f_{post}(x_{it}|z_{it})$ 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, p. 363] propose a one-sided test for the null, in which the test statistic is asymptotically distributed standard normal. Panel A in table 2 reports the test outcome for the 24 cases. (Online appendix B provides a detailed explanation of both tests conducted in this section.) Aside from the bad state of (S, L) in the consumer goods sector, the statistics values lie above the critical value of the 1% confidence level (2.33). Consequently, we reject the null hypothesis of no change in the excess returns density from the pre-SOX to post-SOX eras for practically all firm types in both states.

Change in the shape of the contract. Let $w_{pre}(x_{it}, z_{it})$ denote CEO compensation as a function of (x_{it}, z_{it}) in the pre-SOX era and similarly define $w_{post}(x_{it}, z_{it})$ in the post-SOX era. To test whether the two mappings are equal, we follow Aït-Sahalia et al. [2001] by including an indicator variable for the post-SOX regime in nonparametric regressions of compensation on the excess return x_{it} for each z_{it}. 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 two cases, the value of the statistic exceeds 1.64, implying the null hypothesis of no change in the compensation contract shape is rejected at the 5% level. Moreover, in these two exceptions, panel A shows we reject the null hypothesis that the excess 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 a sense of what lies behind rejecting the null hypothesis of no change, figure 1 shows how the shape of the excess returns probability density function and the estimated compensation

TABLE 2
Nonparametric Tests

Panel A: Test on PDF of excess returns

	Primary		Consumer		Service	
(A, C)	Bad	Good	Bad	Good	Bad	Good
(S, S)	24.16	27.82	15.52	14.67	23.65	23.76
(S, L)	8.31	6.85	−0.62	2.98	14.98	6.69
(L, S)	8.59	19.36	4.66	3.02	7.84	18.29
(L, L)	43.55	17.36	9.06	12.56	61.39	22.34

Panel B: Test on contract shape

	Primary		Consumer		Service	
(A, C)	Bad	Good	Bad	Good	Bad	Good
(S, S)	10.25	1.81	2.55	1.25	1.70	1.52
(S, L)	8.24	8.28	2.16	2.30	5.09	11.78
(L, S)	28.16	7.86	3.43	1.72	5.70	3.33
(L, L)	16.28	9.62	2.26	5.02	8.90	5.75

Both tests are one-sided. This table reports the statistics of the two tests, which both follow a standard normal distribution $\mathcal{N}(0, 1)$. Firm type is measured by the coordinate pair (A, C), where A is assets and C is the debt-to-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.

schedule adjusts for small, low-leverage firms in the consumer goods sector, controlling for the state of the firm (bad vs. good) and the two eras (pre-SOX versus post-SOX). The two top panels show that in both states, density for excess returns shifted to the right and became more concentrated about the mean after SOX. Comparing panel A with B, the mean returns are not surprisingly higher in the good state. The bottom panels show that in both eras the compensation schedule is steeper in the good state than in the bad. Also, both plots in the post-SOX era (panel D) tend to be flatter than in the pre-SOX era (panel C). Overall, concentrating the excess returns distribution and flattening the compensation schedule's extremes reduces the dispersion of compensation between the pre- and post-SOX eras, as reported in table 1.

Balancing the panel. One concern when interpreting the tests and figure 1 is that the pre-SOX era is several times longer than the post-SOX eras. To address this concern, we constructed a *balanced* sample comprising the same two post-SOX years, 2004 and 2005, but only including two years in the pre-SOX era, namely, 2000 and 2001. Table S1 in the online appendix displays the test results for differences in excess returns and the shape of the compensation schedule. With a single exception all the null hypotheses is rejected. Figure S1 in the online appendix plots the estimated excess returns density and compensation schedule for the same firm type as illustrated in figure 1. Although they are not identical, figure S1 shares

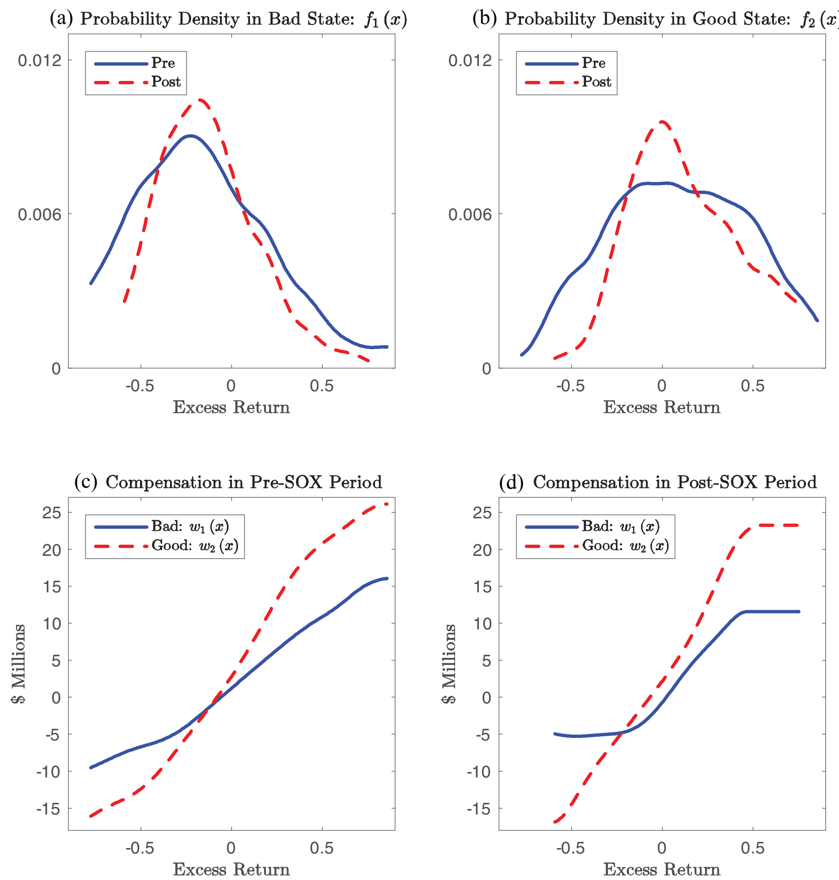

FIG 1.—Empirical compensation schedule and excess returns density.
The plots present the nonparametrically estimated density of excess returns and the optimal compensation of firms with small size and low leverage in the Consumer Goods sector. “Pre” and “Post” indicating the pre-SOX and post-SOX eras. The compensation of both periods is anchored at bond prices equal to 16.5 (b_t) and 16.4 (b_{t+1}).

many of the same features of figure 1. Summarizing both the balanced sample and the main (unbalanced) sample provide overwhelming evidence that the distribution of excess returns, and the compensation schedule, changed with the implementation of SOX.

3. Model

This section lays out a principal–agent model of optimal contracting between shareholders and their CEO, borrowed from Gayle and Miller [2015], in which the risk-averse CEO has hidden information and takes actions that risk neutral shareholders do not observe. Following Holmstrom [1979], it is well-known that including moral hazard in the model explains why CEO compensation systematically varies with the value of the firm he or she (henceforth he) manages.

Our justification for augmenting hidden actions with hidden information is threefold. First, even though accounting numbers are not financial returns to be converted to consumption, they constitute information that shareholders value. This is evident from the fact that accounting announcements sometimes have an immediate impact on the firm's market value. Second, in structural empirical work incorporating accounting information, Gayle and Miller [2015] find evidence in favor of a hybrid model (HMH) over one of pure moral hazard (PMH). The third reason for explicitly modeling hidden information is that SOX was fundamentally directed toward improving the quality of private information revealed to shareholders for assessing the firm's value.

There are two states in the model, good and bad, profiled in tables 1 and 2. If the good state is not directly verifiable, the CEO might have an incentive to falsely claim the firm is more profitable than it really is. Lying this way prompted legislators to enact SOX. We do not model this form of private information: it is hard to accurately gauge the deterrence effect of SOX from a relatively small number of detected criminal violations. Our analysis focuses on the error of not disclosing good states when they occur: SOX imposed more stringent requirements to verify a good state and increased the costs of exposure to false convictions with harsher penalties.

To make the model empirically tractable, yet capture a role for discretionary accounting information, our approach to the agency problem is minimalist. A candidate offered the position of CEO takes one of three actions: he rejects the offer and takes an outside option instead, such as retirement or alternative employment; he accepts the offer and works, by maximizing the expected value of the firm; he accepts the job and shirks, following some personal objectives instead of value maximization. Upon accepting employment the CEO observes one of two states, but only the good state can be verified directly by shareholders. Shareholders observe whether the candidate for CEO accepts or rejects their offer. If the CEO accepts their offer, they receive his report on the state of the firm and, if the CEO announces a good state, verify whether the report is truthful. Later they observe the firm's financial returns, and then pay the CEO according to contractual arrangements. They do not observe whether the CEO works or shirks; nor can they verify whether the CEO is truthful if he announces a bad state.[9]

Information and choices. At the beginning of period t, the firm realizes a gross return on its stock (i.e., before subtracting managerial

[9] For the sake of expositional convenience, we assume that a board, representing shareholders and supported by a compensation committee, designs the contract. An alternative interpretation is that candidates interview for the position of CEO by submitting proposals to the board.

compensation), from events of the previous period. It is the sum of $\pi_t \in \mathbb{R}$, an aggregate shock to all firm returns reflected by a value-weighted stock index, plus $x_t \in \mathbb{R}$, an excess return that is firm specific and independently distributed across t. Then shareholders compensate the CEO for period $t-1$ according to previously agreed terms with $w_t \in \mathbb{R}$. Candidates for CEO in period t enter the period with accumulated wealth e_t (including their compensation from employment in $t-1$) and make their consumption choice c_t. The shareholders, represented by their board, make an offer to a candidate: it is defined by a compensation plan, endogenously determined in the model as an optimal contract. The candidate receiving the offer then chooses whether to accept the employment terms or not, setting $l_{t0} = 1$ if he rejects it and seeks employment elsewhere, and setting $l_{t0} = 0$ if the candidate becomes CEO. An unknown state denoted by $s_t \in \{1, 2\}$ is drawn that affects the probability distribution of the firm's excess returns in period $t+1$. The probability that $s_t = s$ is denoted by $\varphi_s \in (0, 1)$, and is independent across t.

If the CEO accepts employment with the firm ($l_{t0} = 0$), the state of the firm, s_t, is now fully revealed to the CEO but remains partially hidden from the shareholders. Upon privately observing s_t, the CEO reports $r_t \in \{1, 2\}$ to the board. Shareholders know the CEO is lying if $r_t = 2$ and $s_t = 1$, but cannot directly infer whether the CEO is being truthful or not if $r_t = 1$ and $s_t = 2$. Then given (t, s, r) the CEO makes an unobserved labor effort choice, denoted by $l_{tsrj} \in \{0, 1\}$ for $j \in \{1, 2\}$, either by working in the interests of shareholders, and setting $l_{tsr2} = 1$, or by following personal interests on the job, called shirking, in which case $l_{tsr1} = 1$. As rejecting the firm's offer, working for shareholders, and shirking in the firm are mutually exclusive activities, $l_{t0} + l_{tsr1} + l_{tsr2} = 1$ for each (t, s, r). We also let:

$$l_{tj} \equiv \sum\nolimits_{s=1}^{2} \sum\nolimits_{r=1}^{2} 1\{s_t = s\}1\{r_t = r\} l_{tsrj}$$

denote whether the CEO chooses activity $j \in \{0, 1, 2\}$ in period t or not; for example, $l_{t1} = 1$ means the CEO shirks in period t.

Preferences. We assume shareholders can neutralize their exposure to risk from firm specific excess returns through portfolio diversification, and are therefore risk neutral on this dimension of uncertainty. The CEO is an expected utility maximizer, and we parameterize his preferences with a utility function that is additively separable over periods and multiplicatively separable between consumption and work activity within periods, expressing lifetime utility as:

$$-\sum\nolimits_{t=0}^{\infty} \sum\nolimits_{j=0}^{2} \beta^t \alpha_j \exp(-\gamma c_t) l_{tj}, \tag{1}$$

where β is the constant subjective discount factor, γ is the constant absolute level of risk aversion (CARA), and α_j is a utility parameter that measures the distaste from three mutually exclusive options $j \in \{0, 1, 2\}$. Because optimal choices are invariant to linear transformations of utility, we normalize $\alpha_0 \equiv 1$. As working is more distasteful than shirking, we assume $\alpha_2 > \alpha_1$, an inequality lying at the heart of the agency problem in this model. The CARA assumption simplifies the analysis, because it implies the preferences of a CEO toward his risk exposure are independent of his wealth. This implies the optimal contract, derived below, is independent of outside wealth, namely, personal assets with returns that are independent of his own firm's idiosyncratic excess return. Relaxing CARA would be straightforward if the data on each CEO included holdings of all his assets (financial and real), not just those related to his firm.

Financial markets and the distribution of excess returns. Neither the actions of the CEO nor the state of the firm are contractible in this model, because neither are fully observed by shareholders. We assume there are no other impediments to trade. The shareholders are well-diversified, and the CEO has access to well-functioning market to smooth consumption streams with wealth they have accumulated. Formally, we assume a complete set of markets for all publicly disclosed events exists, and attribute all deviations from the law of one price to the information asymmetries of PMH and private information.[10] No assumptions are imposed on the distributional properties of π_t, the aggregate shock to the economy. However, mainly because the time-series length of our panel data is short, but also to simplify the exposition of the theory, we model the interest rate process parsimoniously.

We denote the probability density function for excess returns when the CEO works and the state is s by $f_s(x)$. Also let $f_s(x)g_s(x)$ denote the probability density function for excess returns in period t when the CEO shirks. Thus, $g_s(x)$ is the likelihood ratio for the shirking density relative to the working density.[11] We assume the likelihood of shirking declines to zero as excess returns increase without bound:

$$\lim_{x\to\infty}\left[g_s(x)\right] = 0 \tag{2}$$

for both states $s_t \in \{1, 2\}$. Loosely speaking, this assumption states that very high excess returns are most unlikely to have been achieved from shirking.

[10] The assumption of complete markets, and tests of the assumption, have been used when applying structural econometric models to panel data of consumption and labor supply (Altug and Miller [1990, 1998]), housing size and labor supply (Sieg and Miller [1995]), first home purchase, fertility, and labor supply (Khorunzhina and Miller [2021]), as well as managerial compensation (Margiotta and Miller [2000], Gayle and Miller [2009, 2015], Gayle, Golan, and Miller [2015]). A common alternative assumption in structural labor econometrics is that all current income is immediately consumed. The alternative is inappropriate for this context, because CEOs of these large companies are at this point in their life primarily saving for their retirement, and have access to many financial vehicles for saving.

[11] In our model, $g(x)$ also measures the nature and quality of information shareholders receive. Loosely speaking, if excess returns x are such that $g(x) = 0$, then the CEO must have worked; no information is received from x if $g(x) = 1$, while if $g(x)$ is extraordinarily high, the CEO most probably shirked.

We assume shareholders maximize expected excess returns. Consequently, working (in the interests of shareholders) implies:

$$\int x f_s(x) g_s(x) \mathrm{d}x \equiv E_s\big[x g_s(x)\big] < E_s[x] \equiv \int x f_s(x) \mathrm{d}x. \tag{3}$$

Finally, we assume the weighted likelihood ratio of the second state occurring relative to the first, converges to an upper finite limit as x increases:

$$\lim_{x \to \infty} \big[\varphi_2 f_2(x)/\varphi_1 f_1(x)\big] \equiv \lim_{x \to \infty} [h(x)] = \sup_{x \in \mathbb{R}} [h(x)] \equiv \overline{h} < \infty. \tag{4}$$

Intuitively, the assumption gives the CEO license to falsely report with impunity that the bad state occurred, putting the onus on shareholders to design a contract that provides the CEO with sufficient incentive to disclose the good state when it occurs.

4. *Evaluating the Importance of Agency*

A conflict of interest arises between shareholders and their CEO, because from (3) shareholders prefer the CEO to work, but the inequality $\alpha_2 > \alpha_1$ implies the CEO prefers shirking. To quantify this conflict of interest, we compare how much the firm loses from the CEO shirking with how much the CEO gains from shirking when the CEO is paid a fixed wage. We also decompose expected CEO compensation into two additive pieces, how much the CEO would be paid if his private information about the firm was common knowledge and shareholders observed his actions, which we call administrative pay, and the expected extra pay he receives because of the agency problem. The formulas defining these measures, displayed in figure 2, are the objects of our structural estimation, and this section explains how they arise in the context of our model.

Conflict of interest. Absent incentives, the interests of a firm's shareholders and their CEO are not aligned because the shareholders prefer the CEO to work, but the CEO prefers to shirk. Shareholders incur a gross loss in returns of:

$$\rho_1 \equiv \sum\nolimits_{s=1}^{2} \int \varphi_s [x - x g_s(x)] f_s(x) \mathrm{d}x \tag{5}$$

if the CEO shirks rather than works.

When comparing shirking and working, the CEO weighs the nonpecuniary benefits of the activity against the probability distribution that defines its pecuniary benefits. The compensating differential of the nonpecuniary benefit from being employed in the current period and exerting effort $j \in \{1, 2\}$ is $\gamma^{-1} \ln \alpha_j$. Therefore, the compensating differential in nonpecuniary benefits between working and shirking denominated in current consumption units is $\gamma^{-1} \ln(\alpha_2/\alpha_1)$.

The timing of the nonpecuniary benefits to the CEO and his pay is nonsynchronous: nonpecuniary utility accrues in the current period but pay is

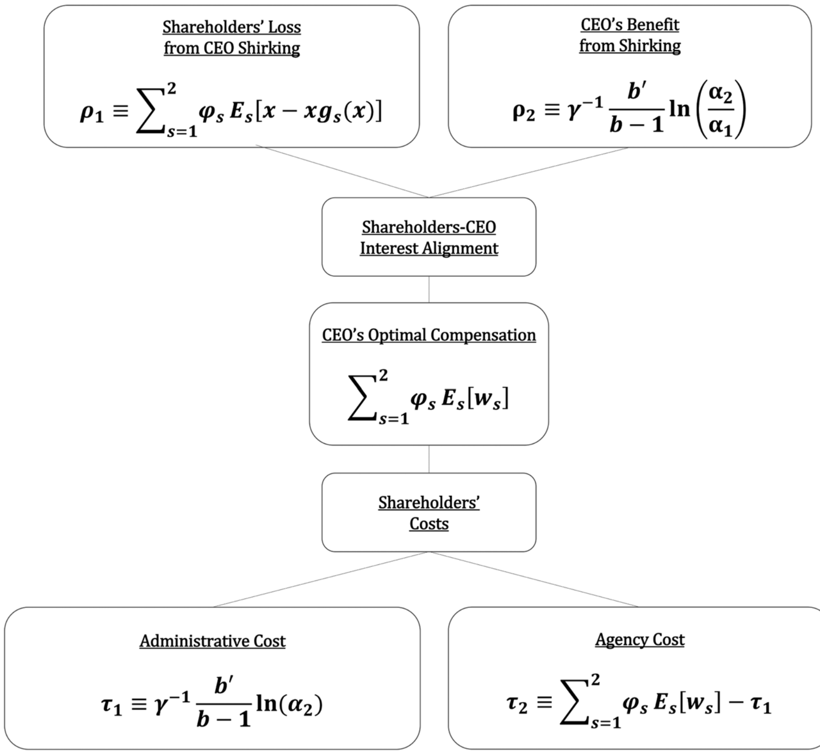

FIG 2.—Welfare costs summary.

received next period. To make these two sources of benefit comparable, we compute the annuitized utility gain that would occur from next period onward. Let b denote the current bond price, b' the bond price next period and ι the current one period interest rate. Noting that $b - 1 = (1 + \iota)^{-1} b'$, we define the certainty equivalent differential gain, denominated in next period's consumption units, to a CEO from shirking instead of working in the current period as:

$$\rho_2 \equiv (1 + \iota)\gamma^{-1} \ln(\alpha_2/\alpha_1) = b'\left[(b - 1)\gamma\right]^{-1} \ln(\alpha_2/\alpha_1). \quad (6)$$

Compensation costs. Let $w_r(x)$ denote the compensation paid to the CEO at the beginning of next period as a function of firm excess returns, if the CEO reports the current state $r \in \{1, 2\}$, the CEO is truthful, and works. Denominated in next period's consumption units, total expected compensation, averaged over both states, can be expressed as:

$$\tau_0 \equiv \sum_{r=1}^{2} \int \varphi_r w_r(x) f_r(x) \mathrm{d}x. \quad (7)$$

The null hypotheses that CEO compensation does not vary with the firm's state or with the firm returns is rejected by our data. In the equilibrium of our model derived below, the only way to induce the CEO to work is to offer a contract that depends on x. If the actions of the CEO were monitored by shareholders, they would pay him a fixed wage, denoted by τ_1. We interpret τ_1 as payment to the CEO for undertaking administrative services within the firm instead of pursuing an outside option; in other words τ_1 is the certainty equivalent, or compensating differential, for being employed as CEO and working. If the CEO could be monitored perfectly, then his tasks would be essentially administrative, directing and coordinating human resources without taking responsibility beyond the requirements of the job statement: hence the name. It is easy to prove that in terms of next period's consumption units:

$$\tau_1 \equiv (1+\iota)\gamma^{-1} \ln(\alpha_2) = b'\left[(b-1)\gamma\right]^{-1} \ln \alpha_2. \tag{8}$$

We now define the agency cost as $\tau_2 \equiv \tau_0 - \tau_1$, the difference between expected total compensation and the administrative wage. It is the expected amount the firm pays the CEO because shareholders cannot monitor his activities. In the optimal contract derived below, τ_2 is also the risk premium paid to the CEO with certainty equivalent τ_1.

5. *Optimal Contracting*

The values of ρ_1, ρ_2, τ_1, and τ_2 depend on the parameters defining the model. Some parameters, namely, $f_s(x)$, can be estimated directly if there are data on (s, x) and the CEO truthfully reports the state and works; similarly $f_s(x)g_s(x)$ can be estimated directly if the CEO truthfully reports and shirks. Others, such as α_j and γ, cannot be inferred without making assumptions about the behavior of shareholders and their CEO.

The premise of our econometric analysis is that shareholders design contracts reflecting their interests, and that CEOs respond rationally. Thus, shareholders calculate the expected gross benefits from employing a CEO to shirk and also to work, offset those benefits with the expected CEO compensation from the respective cost minimizing contracts, and select the maximum if it is positive. It is straightforward to prove the cost minimizing contract for employing a CEO to shirk is $(1+\iota)\gamma^{-1}\ln\alpha_1$. Because our data on CEO compensation show that in both states their pay depends on the firm's excess returns, we focus on the problem of deriving the cost minimization problem for inducing the CEO to work in both states.

There are no gains from a long-term arrangements between shareholders and the CEO in this framework, because the distribution of the firm's financial returns is independent of his actions taken more than one period ago, and his private information is only useful for forecasting returns one period ahead. The benefits from a long term contract arise if, for example, the hidden actions of the CEO were only revealed some years after they were taken. Consequently, the optimal long-term contract between shareholders and the CEO in this model decentralizes to a sequence of short-term one-period contracts (Malcomson and Spinnewyn [1988], Fudenberg et al [1990], Rey and Salanie [1990], Gayle and Miller [2015]).

When comparing two employment activities, such as shirking and working, the CEO weighs the nonpecuniary benefits of the activity in the current period against the probability distribution defining its pecuniary benefits next period. To render these two sources of benefit comparable, we compute the utility gain that would occur from next period onward. Similar to the calculation of ρ_2 in (6), the total value of undertaking activity $j \in \{1, 2\}$ in the current period (relative to the outside option) and being paid w in the next period, denominated in units of next period's consumption, is therefore:

$$b'\left[(b-1)\gamma\right]^{-1}\ln(\alpha_2) + w. \tag{9}$$

In our model's equilibrium, optimally smoothing CEO consumption over the remaining periods of his life amounts to valuing additional wealth by a factor that scales up his current utility function by an annuitized amount, a property that derives directly from the exponential utility (or CARA) assumption (Margiotta and Miller [2000]). Compensation depends on contractible events, including the bond price, the state the CEO reports, and subsequent excess returns. Then the annuity value of $w_r(x)$ starting with (the payment of) a consumption unit next period is $w_r(x)/b'$, which has a utility equivalent of:

$$v_r(x) \equiv \exp\left[-\gamma w_r(x)/b'\right]. \tag{10}$$

This representation of the CEO's lifetime indirect utility function considerably simplifies the shareholders' contracting problem: appealing to (9) and (10) when considering the CEO's responses to the contract terms, we only need to compare the expected value of expressions like $\alpha_j^{1/(b-1)} v_r(x)$ for different (j, r, s).

Appealing to the revelation principle (Myerson [1982]), the optimal contract is solved by a direct mechanism: in the current period t shareholders choose $w_r(x)$ for each (r, x) to minimize the expected cost of managerial compensation subject to the constraints that the CEO prefers to truthfully report rather than lie, and that he prefers working to shirking in both states.[12]

Participation. To induce an honest candidate for CEO to accept employment with the firm and work, his annuitized expected utility from employment must exceed the utility obtained from taking the outside option, which is unity. Setting $(j, r, s) = (2, s, s)$ in $\alpha_j^{1/(b-1)} v_r(x)$, taking the

[12] We assume legal concerns deter CEOs from reporting $r = 2$ when $s = 1$. An extended analysis could incorporate this additional choice, with the consequence that law enforcers catch the liar with some probability less than one.

expectations over x and averaging over states, the candidate would only accept the position if

$$\sum_{s=1}^{2} \int \varphi_s \alpha_2^{1/(b-1)} v_s(x) f_s(x) \mathrm{d}x \leq 1. \tag{11}$$

Incentive compatibility. Given his decision to be employed and truthfully reveal the state, the incentive compatibility constraint induces the CEO to prefer working to shirking in both states $s \in \{1, 2\}$. Shirking yields an immediate nonpecuniary benefit, because $\alpha_1^{1/(b-1)} < \alpha_2^{1/(b-1)}$, yet the expectation of the (annuitized consumption) value of compensation $v_s(x)$ is taken with respect to the $g_s(x) f_s(x)$ density rather than $f_s(x)$. Comparing $\alpha_j^{1/(b-1)} v_r(x)$ evaluated at $(j, r, s) = (2, s, s)$ and $(1, s, s)$, the incentive compatibility constraint is

$$\int \alpha_2^{1/(b-1)} v_s(x) f_s(x) \mathrm{d}x \leq \int \alpha_1^{1/(b-1)} v_s(x) g_s(x) f_s(x) \mathrm{d}x \tag{12}$$

Truth telling. 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 with the expected utility from reporting honestly in the second state and working, we obtain the truth-telling constraint. Noting the nonpecuniary utility component $\alpha_2^{1/(b-1)}$ is the same for both activities (and therefore cancels in the inequality):

$$\int v_2(x) f_2(x) \mathrm{d}x \leq \int v_1(x) f_2(x) \mathrm{d}x. \tag{13}$$

Sincerity. Finally, 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

$$\int \alpha_2^{1/(b-1)} v_2(x) f_2(x) \mathrm{d}x \leq \int \alpha_1^{1/(b-1)} v_1(x) g_2(x) f_2(x) \mathrm{d}x. \tag{14}$$

Optimization. Minimizing expected compensation is equivalent to choosing $v_s(x)$ that maximizes:

$$\sum_{s=1}^{2} \int \varphi_s \ln\left[v_s(x)\right] f_s(x) \mathrm{d}x. \tag{15}$$

Appealing to the Kuhn–Tucker theorem, there is a unique positive solution to the equation system formed from the first-order conditions augmented by the complementary slackness conditions. The Lagrangian and the first-order conditions are laid out in the appendix.

It is well-known that the hidden action component to the agency problem explains why the slope of the compensation schedule qualitatively follows the likelihood ratio function for the shirking to working densities. A positive slope indicates the density of excess returns from working is increasing relative to the corresponding density from shirking; the flattening at the upper end of compensation is consistent with (2). If shareholders directly observed $s \in \{1, 2\}$, they would maximize (15), the same objective function, with two fewer constraints, dropping (13) and (14). The appendix shows that in this closely related model of PMH, the expected utility to the manager is equalized across states. In the HMH, the CEO never falsely announces the good state, to avoid the risk of criminal prosecution. Thus, the optimal compensation schedule for the good state in the HMH is almost identical to its PMH counterpart. The only difference is that in the HMH expected utility of the CEO is higher in the good state than the bad.

In the HMH, the CEO can fail to disclose the good state (e.g., a windfall gain to the firm that might occur before any action is taken) and instead misreport by declaring the bad state, subsequently attributing the increased financial return to an exceptionally lucky draw from the distribution of the bad state. To deter him from lying about the good state shareholders compensate the CEO: (1) for revealing the good state with $v_2(x)$, by increasing the annuitized value of utility in the PMH model for each x by exactly the same amount, and (2) with a lower amount than in the PMH if the CEO reports the bad state and firm returns are high. Summarizing, the appendix shows the (annuitized) expected utility is equalized in both states of the PMH, more than the expected utility conditional on the bad state in the HMH, but less than the expected utility the CEO receives conditional on the good state in the HMH.

These features are on display in both figure 1, which plots the reduced form of compensation on excess returns, and the left side of figure 3, which plots the optimal contract for the HMH model given parameters estimated for low-leveraged small firms in the primary sector. The illustrated schedules vary with excess returns, increase throughout most of the domain, and flatten at very high rates of excess returns. In the HMH, they are steeper for the good state than the bad, and expected compensation in the good state noticeably exceeds expected compensation in the bad state, a prominent feature of table 1. The PMH model, illustrated on the right side of figure 3 for the same parameter values, is less convincing.[13] In contrast to figure 1, the slopes across states are comparable over most of the domain of x, as are expected payoffs by state, in seeming contradiction to table 1.

6. *Identification*

The parameters defining the model characterize the excess returns to the firm and CEO preferences. The probability density functions for excess return in each state, conditional on working or shirking, are represented by $f_s(x)$ and $g_s(x)$ for $s \in \{1, 2\}$, while φ_2 is the probability of the second state occurring. CEO preferences are defined, relative to the normalized

[13] Online appendix D outlines how we numerically computed the solution to the PMH model.

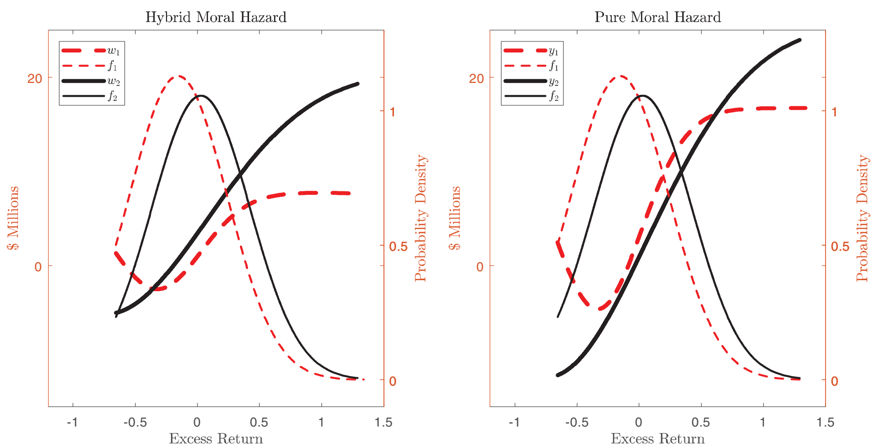

FIG 3.—Optimal compensation schedules.
The plots apply to firms with small size and low leverage in the primary sector in the pre-SOX era. The risk aversion parameter γ equals 0.08. The shirking parameter $\alpha_1$equals 0.96 and the working parameter α_2 equals 1.20. Bond prices are set to $b_t = 16.5$ and $b_{t+1} = 16.4$. The excess return is approximated by a one-side truncated normal distribution $\mathcal{TN}(a,\mu,\sigma)$ with left truncation point a, and a parent normal distribution with mean μ and standard deviation σ as follows. When the CEO: works in the bad state $\mathcal{TN}(-0.66, -0.16, 0.39)$; works in the good state $\mathcal{TN}(-0.66, 0.03, 0.39)$; shirks in the bad state $\mathcal{TN}(-0.66, -0.25, 0.27)$; shirks in the good state $\mathcal{TN}(-0.66, -0.11, 0.36)$. The probability of $s = 1$, the bad state, is 0.54, implying the probability of the good state is 0.46.

utility from taking the outside option, by their distaste for working, α_2, and shirking, α_1, as well as their risk aversion parameter, γ. To summarize, the model is parameterized by:

$$\theta \equiv \left(\gamma, \alpha_1, \alpha_2, \varphi_2, \left\{ f_s(x), g_s(x) \right\}_{s=1}^{2} \right) \in \Theta, \tag{16}$$

where the definition of Θ follows from the preceding discussion of its elements. Denote by $\theta^* \in \Theta$ the (true) model generating the data.

Suppose the data come from a cross-section of firms: bond prices (b, b') are observed in the two consecutive periods, as is each firm type $z \in Z$ (defined in section 2), an accounting earnings report $r \in \{1, 2\}$ for the current period and financial excess returns $x \in \mathbb{R}$ for the following period. Given θ^* optimal CEO compensation $w_r^*(x)$ is defined conditional on (b, b', z), but its dependence on (b, b', z) is subsumed for notational convenience: we assume $w_r^*(x)$ is observed with random error by w.

Several parameters in the model are *point* identified and estimated without recourse to the optimality conditions. Thus, φ_2^* can be estimated by its relative frequency in the sample. Data on excess returns are generated by $f_s^*(x)$ and can be estimated nonparametrically. Both are identified and estimated conditional on z. We assume the measurement error $w - w_r^*(x)$ is identically and independently distributed about $w_s^*(x)$, implying:$w_r^*(x) = E[w|r, x]$. This proves the reduced form of the optimal compensation equation $w_r^*(x)$ is point identified too. Our estimates show that $w_r^*(x)$ varies with accounting earnings reports $r \in \{1, 2\}$ and financial excess returns $x \in \mathbb{R}$, proving that if the model is correctly specified, shareholders induce the CEO to work. In that case CEOs truthfully reveal the state, implying $r = s$ in the equilibrium optimal contract.

The remaining parameters in θ^*, that is the likelihood ratio of excess returns in each state for shirking versus working $g_s^*(x)$, plus the preference parameters of the CEO $(\gamma^*, \alpha_1^*, \alpha_2^*)$, are only *set* identified. We say an element $\theta \in \Theta$ is *observationally equivalent* to θ^* if it has the same data-generating process as θ^*, so cannot be distinguished from θ^* regardless of how many observations are in the data set. Denote by $\Gamma \subseteq \Theta$ the subset of observationally equivalent elements to θ^*. Then θ^* is point identified if Γ is a singleton, and set identified otherwise.

We exploit restrictions that arise from the first-order conditions for the shareholders' cost minimization problem, the complementary slackness conditions associated with the constraints, plus an inequality that reflects shareholders preferences for the CEO to work rather than shirk. Identifying $g_s^*(x)$ and $(\gamma^*, \alpha_1^*, \alpha_2^*)$ decomposes into two steps. First, given $f_s^*(x)$ and φ_s^* for $s \in \{1, 2\}$, we write (α_1, α_2) and $g_s(x)$ as mappings of γ by exploiting some of the conditions in the model implied by optimal contracting. The second step forms a criterion function that consolidates the remaining equalities and inequalities not used in the first step to set identify γ^*, the only remaining parameter.

Concentrating the parameter space. Let $v_s(x, \gamma) \equiv \exp[-\gamma w_s^*(x)/b']$. Also let $\overline{v}_s(\gamma) \equiv \lim v_s(x, \gamma)$ as $x \to \infty$. The first step concentrates the parameter space to a single dimension, given by γ:

$$
\begin{aligned}
\alpha_1(\gamma) &\equiv \left\{ E^*\left[v_s(x, \gamma)\right] \frac{\overline{v}_2(\gamma)^{-1} - E_2^*\left[v_2(x,\gamma)\right]^{-1}}{\overline{v}_2(\gamma)^{-1} - E_2^*\left[v_2(x,\gamma)^{-1}\right]} \right\}^{1-b} \\
\alpha_2(\gamma) &\equiv E^*\left[v_s(x, \gamma)\right]^{1-b} \\
g_1(x, \gamma) &\equiv \frac{\overline{v}_1^{-1}(\gamma) - v_1(x,\gamma)^{-1} + \eta_3(x,\gamma)\left[\overline{h} - h(x)\right] - \eta_4(x,\gamma) g_2(x,\gamma) h(x)}{\overline{v}_1^{-1}(\gamma) - E^*\left[v_s(x,\gamma)\right]^{-1} + \eta_3(x,\gamma)\overline{h}} \left(\frac{\alpha_2(\gamma)}{\alpha_1(\gamma)}\right)^{1-b} \\
g_2(x, \gamma) &\equiv \frac{\overline{v}_2(\gamma)^{-1} - v_2(x,\gamma)^{-1}}{\overline{v}_2(\gamma)^{-1} - E_2^*\left[v_2(x,\gamma)^{-1}\right]},
\end{aligned}
\tag{17}
$$

where the expectation operator $E^*[\cdot]$ is defined by integrating over (x, s) using $f_s^*(x)$ and φ_s^*, similarly $E_s^*[\cdot]$ is defined by integrating over x using $f_s^*(x)$ conditional on s, and:[14]

$$
\begin{aligned}
\eta_3(x, \gamma) &= E_2^*\left[v_2(x, \gamma)\right]^{-1} - \eta_4(x, \gamma) - E^*\left[v_s(x, \gamma)\right]^{-1} \\
\eta_4(x, \gamma) &= \frac{\frac{E_1^*\left[v_1(x,\gamma)\right]}{E^*\left[v_s(x,\gamma)\right]} - E_1^*\left[v_1(x,\gamma)h(x)\right]\left\{E_2^*\left[v_2(x,\gamma)\right]^{-1} - E^*\left[v_s(x,\gamma)\right]^{-1}\right\} - 1}{\left(\frac{\alpha_2(\gamma)}{\alpha_1(\gamma)}\right)^{1-b} E_1^*\left[v_1(x,\gamma) g_2(x,\gamma) h(x)\right] - E_1^*\left[v_1(x,\gamma)h(x)\right]}.
\end{aligned}
\tag{18}
$$

[14] η_3 corresponds to the truth-telling constraint and η_4 corresponds to the sincerity constraint.

Gayle and Miller [2015] prove $\alpha_s^* = \alpha_s(\gamma^*)$ and $g_s^*(x) = g_s(x, \gamma^*)$ for $s \in \{1, 2\}$ any period t. Therefore, if γ^* is known, the remaining components in $\theta \in \Theta$ are identified. More generally, tracing out the regions implied by the transformations $\alpha_s(\gamma^*)$ and $g_s(x, \gamma^*)$ from the identified set for γ^* set identifies α_s^* and $g_s^*(x)$.

To convey a flavor of their proof, consider $\alpha_2(\gamma)$, expressed more fully as

$$\alpha_2(\gamma) = \left[\sum\nolimits_{s=1}^{2} \int \varphi_s^* v_s(x, \gamma) f_s^*(x) \mathrm{d}x\right]^{1-b}. \tag{19}$$

It is easy to establish by a contradiction argument that the participation constraint binds in the optimal contract, and hence (11) holds with equality (Margiotta and Miller [2000]). Equation (19) now follows by making $\alpha_2(\gamma)$ its subject. As $f_s^*(x)$ and φ_s^* for $s \in \{1, 2\}$ are identified, along with $w_s^*(x)$, it follows from (10) that if γ^* is known, then α_2 is point identified if there is a unique solution to (19). The other expressions are derived from the cost minimization optimality conditions in a similar manner.

Set identifying the risk aversion parameter. The index representation of $\alpha_j(\gamma)$ and $g_s(x, \gamma)$ given in (17) does not draw upon all the restrictions imposed by the cost minimizing contract that ensures participation and work. In particular, the truth-telling (13) and sincerity (14) inequalities are not used in this derivation. For example, appealing to (13), the truth-telling constraint implies:

$$\int v_2(x, \gamma^*) f_2^*(x) \mathrm{d}x \leq \int v_1(x, \gamma^*) f_2^*(x) \mathrm{d}x. \tag{20}$$

A similar inequality follows from (14). Both inequalities are necessary conditions circumscribing observationally equivalent values of γ to γ^*.

Further inequalities helping to pin down γ^* come from the model's prediction that if shareholders design contracts, or alternatively are presented with contracts maximize firm value shareholders must prefer to the CEO to work rather than shirk when compensation varies with the firm's excess returns in both states: implies that paying the CEO to work in both states is more profitable than paying him to shirk in at least one of them. Both types of conditions restrict the set of *admissible* risk parameters, values of γ that are observationally equivalent to γ^*. A second source of restrictions on γ stems from shareholders objective to maximize profits by inducing the CEO to work in both states, rather than shirk in at least one of them and only pay a fixed, relatively low wage. For example, the expected net present value of paying the CEO to work in both states exceeds the value of paying him (less) to shirk in both states, or

$$V_s \int x\big[1 - g_s\big(x, \gamma^*\big)\big] f_s^*(x) \mathrm{d}x \geq E[w|s, x, b] - \big[\ln \alpha_1\big(\gamma^*\big)\big]/\gamma^*, \tag{21}$$

where V_s is firm value in state s. and $[\ln \alpha_1(\gamma^*)]/\gamma^*$is the (fixed) wage of employing him to shirk.

The parameter representation given in (17) is based on data covering one full compensation cycle, that is from setting the CEO contract at the beginning of period t to payment at the beginning of period $t+1$. Data on multiple periods with varying bond prices further restrict the set of admissible risk aversion parameters. For example, suppose there exists $(\underline{b}, \underline{b}') \neq (b, b')$ for two pairs of sampled periods. From (19) and (10) it follows from () and the definition of $\alpha_2(\gamma)$ that:

$$(1-b)\ln\left\{E_s^*\left[\exp\left(-\gamma^* E[w|s, b, b', x]\Big/b'\right)\right]\right\} \quad (22)$$
$$=\left(1-\underline{b}\right)\ln\left\{E_s^*\left[\exp\left(-\gamma^* E[w|s, \underline{b}, \underline{b}', x]\Big/\underline{b}'\right)\right]\right\}.$$

Let $Q(\gamma)$ denote a quadratic form of the minus norm composed of those equalities and inequalities implied by the model that are not used in the derivation of (17).[15] For example, one additive component of $Q(\gamma)$ is formed by squaring the difference between the left and right side in (22); another from first substituting the expression for g from (17) into (21), and then squaring the minimum of 0 and the difference between the left and right side of (21). Gayle and Miller [2015] provide a *sharp* and *tight* list of equalities and inequalities captured by $Q(\gamma)$, sharp because every element is admissible, and tight because it contains every admissible element.[16] Appealing to (17) proves Γ can be represented by values of γ satisfying those conditions:

$$\Gamma=\left\{\gamma : Q(\gamma)=0\right\}. \quad (23)$$

In estimation we follow their list, which is reproduced in online appendix C1. As Γ is sharp and tight, the model is misspecified if and only if Γ is empty. To estimate a confidence region for Γ and conduct a misspecification test, we exploit the fact that our estimates of $Q(\gamma)$ are strictly positive only because expectations and population proportions differ from their sample analogs. Thus, the null hypothesis that the data were generated by the model for any $\theta \in \Theta$ is rejected if the sample approximation to $Q(\gamma)$ falls outside the confidence region for all $\gamma \in \mathbb{R}$.

7. *Estimation*

Our estimation strategy adapts the identification analysis described above to the modest sample size of some firm types, especially in the post-SOX era. Aside from the risk aversion parameter γ, which we assume is the same for

---

[15] The minus norm of q, denoted $\|q\|_-$, is the norm of the maximum of $-q$ and 0, or $\|q\|_- = \|\max(-q, 0)\|$.

[16] This approach does not preclude imposing additional restrictions on the primitives, but to prove point identification one must show Γ is single valued. For example, suppose $g_s^*(x)$, the true likelihood in state s, is declining in x. Assuming $g(x, \gamma^*)$ is differentiable in x then $\partial g(x, \gamma^*)/\partial x < 0$. This inequality can be added to the set of conditions admissible γ must satisfy.

all CEOs, we treat each firm category separately (using indicator variables). We make two additional assumptions further specializing the model that are not necessary in the theory or in identification, but are practical ways of confronting some data limitations. They stem from the small number of bond prices observed, and the small number of observations in some firm-type cells. We assume that: (1) both $f_s(x)$ and $g_s(x)$ are independent of bond prices, and (2) b_t follows a first-order difference equation taking the form $b_t = \delta(b_{t+1})$.

Broadly speaking there are three pieces: nonparametrically estimating CEO compensation, using subsampling procedures to estimate the critical value that defines a confidence region for the risk aversion parameter, and computing the four welfare measures to compare the pre-SOX and post-SOX eras with the mappings given by (17). We also undertake four analyses to assess how robust our conclusions are: (1) we vary the sample years to check how sensitive our estimates might be to perceptions about anticipating the implementation of SOX; (2) we compute the key measures of agency indices for two pairs of different bond prices to check the sensitivity of our estimates to aggregate conditions; (3) we estimate the risk aversion parameter regions for the pre-SOX and post-SOX eras separately, checking for evidence that risk preferences changed when SOX was implemented; and (4) we conduct a DID analysis to control for other contemporaneous industry and aggregate factors, by comparing a treated group of firms compelled by legislation to change the structure of their board with a control group that already met some key legislative requirements before SOX was implemented. The implementation details are in online appendices C2 and C3. This section summarizes the three pieces, explains how we deal with the data limitations, and describes the four sensitivity analyses.

Overview. The first piece, estimating the compensation equation, exploits the assumption that $\delta(b) : \mathbb{R}^+ \to \mathbb{R}^+$ is a well-defined but unknown function. We estimate the compensation schedule nonparametrically and separately for each of the firm categories defined in section 2, plus another indicator variable for the pre-SOX versus post-SOX eras, by forming a bivariate kernel estimator that regresses measured compensation $\widetilde{w}_t$ on excess returns x_t and bond price b_t (see online appendix C2 for more details.) Reducing the dimension of the nonparametric regression from $(r_{t-1}, b_{t-1}, b_t, x_t)$ to (r_{t-1}, b_t, x_t) for each firm type is useful because there are only 13 different values corresponding to the sample years. Moreover, observed variation in bond prices hardly affects the results, as our sensitivity analysis demonstrates.

Let $\widehat{Q}(\gamma)$ denote a sample analog to $Q(\gamma)$ and define $\widehat{\Gamma} \equiv \{\gamma : \widehat{Q}(\gamma) \leq \widehat{c}_{0.95}\}$, where $\widehat{c}_{0.95}$ is a consistent estimator of $c_{0.95}$, the 95% critical value. As number of observations diverges, $\widehat{\Gamma}$ converges to those elements in Γ that are sampled 95% of the time. The second piece applies the approach of Chernozukov, Hong, and Tamer [2007], drawing 100 subsamples from the original full sample, following the joint distribution of the public states and the private states. Each subsample contains 80% of the observations in the original sample. For each subsample, we calculate the value of the objective function and use these values to compute $\widehat{c}_{0.95}$. The model is rejected at the 0.05 level if $\widehat{\Gamma}$ is empty.

Having determined $\widehat{\Gamma}$ numerically, we infer the estimated confidence region for the other parameters used to compute the welfare measures. For example the estimated confidence region for $\tau_1(\gamma^*)$ is:

$$\left\{\tau_1(\gamma) \equiv b'\left[(b-1)\gamma\right]^{-1} \ln \alpha_2(\gamma) : \gamma \in \widehat{\Gamma}\right\}.$$

Comparing the welfare measures in the pre-SOX and post-SOX eras, is essentially a counterfactual experiment: while the bond price in the pre-SOX must equal the bond price in the post-SOX to make the measures comparable, table S2 in the online appendix shows that none of the bond prices in the post-SOX era exactly match any of those in the pre-SOX era.

Adapting the procedure to the sample. The simplest way to match bond prices between pre-SOX and post-SOX eras is to fix interest rates and bond prices throughout the entire sample at their sample means. As bond prices are the only form of aggregate time varying heterogeneity affecting the optimal contract, this approximation not only resolves the matching problem, but also justifies combining the observations used in forming the equalities and inequalities defining $Q(\gamma)$ for each firm type over several years. The chief benefit of aggregating this way stems from the fact that it alleviates a shortcoming in the data that several firm type cells only contain a relatively small numbers of observations (see table S3 in the online appendix.) Pursuing this approach entails two costs. The optimal compensation contract varies with bond prices, so lumping together contacts made for different bond prices misspecifies the model.[17] Using overidentifying restrictions of the form given by (22) also tightens the region of observationally equivalent risk parameters used in (23). Therefore, how much variation in bond prices we exploit in estimation is necessarily a matter of judgment.

We divide the first ten annually reported bond prices corresponding to the pre-SOX era into a set of five $B' \equiv \{15.6, 16.1, 16.4, 16.8, 17.8\}$ and the remaining three bond prices into two of those five $\underline{B}' \equiv \{16.4, 16.8\}$, following the rule described in online appendix A5 that roughly approximates their dispersion over the 13 years.[18] When forming $\widehat{\Gamma}$, we condition on pre- versus post-SOX, substitute the wage regression estimates of $E[w|s, \delta(b'), b', x] = E[w|s, b', x]$, obtained from the whole sample at the observed bond prices, into $v_s(x, \gamma) = \exp\{-\gamma E[w|s, b', x]/b'\}$, evaluated at

[17] For example, bond prices help determine the value of consumption smoothed over the life of the CEO, and help shape the participation and incentive compatibility constraints.

[18] Recall that a bond price of 15.6 means that the current price receiving a dollar now, plus a dollar each year in perpetuity, is $15.60.

the bond prices $b' \in B'$, and apply the restrictions laid out in online appendix C1, which include (20), (21) and (22).

When computing the counterfactuals in the third piece of estimation, we make smoothing assumptions. Following Margiotta and Miller [2000] and Gayle and Miller [2009], we impose a truncated normal distribution, by applying a minimum distance estimator (MDE) to our nonparametric estimates of f and g evaluated at different quantiles. Finally, the counterfactuals also require a set $\underline{B} \equiv \{\delta(16.4), \delta(16.8)\}$ because both b and b' explicitly appear in three of the four measures, namely, ρ_1, τ_1 and τ_2. To pair $\underline{B}'$ with a counterfactual set $\underline{B}$, we estimate $\delta(b')$ with an autoregressive process containing linear and quadratic terms.

To summarize the steps: (1) we estimate the compensation regression for B' in the pre-SOX era and $\underline{B}'$ in the post-SOX era, (2) these estimates are used to form restrictions in the pre-SOX era and in the post-SOX era, that are used in constructing $\widehat{Q}(\gamma)$, (3) the subsampling procedure described above yields estimates of the critical value for $\widehat{Q}(\gamma)$ and hence the confidence region for γ, (4) the primitives for f and g are smoothed with an MDE to obtain four truncated normal distributions, (5) we regress the bond price sample on its forward operator and its square, imputing $\underline{B}$, and (6) two sets of estimates of the changes in welfare are computed, corresponding to different bond price pairs and sample sizes.

Sample years and bond price pairs. The text below reports estimates for the pairs of bond prices $(b, b') = (16.5, 16.4)$ obtained for the main sample, which omits the two years bordering on the SOX legislation, namely, 1993–2001 (16,894 observations) and 2004–2005 (3,781 observations). As a robustness check, we repeated the analysis for the pair bond price $(b, b') = (16.5, 16.8)$ on the extended sample, which covers 1993 through 2002 for the pre-SOX era (18,855 observations) and 2003–2005 for the post-SOX era (5,670 observations). The differences are inconsequential, suggesting that a precise determination of the cutoff dates for the two regimes is empirically unimportant, and that the welfare measures are insensitive to relatively small changes in bond prices. For this reason, we relegate the results on the extended sample to the online appendix, where they are displayed in tables S10– S13.

Structural DID analysis. To separate changes due to the implementation of SOX from other aggregate factors we conduct a DID analysis. We define *noncompliant* (treated) firms as those that missed at least one of the three following criteria: (1) an entirely independent compensation committee before July 25, 2002, when SOX was approved, (2) an independent majority board before February 13, 2002, when the SEC asked NYSE and NASDAQ to review their corporate governance requirements, and (3) an entirely independent audit committee before December 31, 2000.[19] The

[19] This last criterion follows Duchin et al. [2010], who suggest that the exchanges adopted the recommendations of the independence of audit committee as early as the end of 1999.

remaining firms for which we have data on these criteria are called *compliant*, and the resulting sample is called *restricted*.[20] Because the control group complied with SOX before it was implemented, at least on these three criteria, we conjecture the response to SOX would be more pronounced within the treatment group.

The rationale for this DID analysis is that if the reasons for belonging to the treated versus control groups are independent of the other features of governance and firm excess returns, then the different behavior between the two groups after SOX was implemented could be attributed to its implementation. A key feature of our structural DID analysis is that we embed within the econometric framework the equilibrium behavior of firms before and after the implementation, helping to inoculate our results against endogenous factors that might otherwise jeopardize this interpretation.

This approach has two limitations. First, the three criteria mentioned above only constitute a subset of the regulations SOX imposed, implying that some firms classified here as controls really belong to the treated group. Second, typical of almost all DID analyses, the presumption is that the factors leading firms to sort this way before SOX was implemented are exogenous. Nevertheless, we believe this approach is informative; the three criteria are among the most onerous of the regulations, and firms that had already met those criteria subsequently faced lower compliance costs. For this reason we report the DID results at the same level of detail as those results not differentiating between these two groups.

8. *Structural Estimates*

The set of model parameters considered in this paper are defined by Θ, and we assume the data are generated by some $\theta^* \in \Theta$. The estimates for $f_s^*(x)$, discussed in section 2, and φ_s^*, reported in table S3, are obtained directly without appealing to the optimal contracting features of the structural model. Given the data-generating process, the remaining parameters of θ^* are fully determined by γ^*, because $\alpha_s^* = \alpha_s(\gamma^*)$ and $g_s^* = g_s(x, \gamma^*)$ for $s \in \{1, 2\}$, defined by (17). The four welfare measures ρ_i and τ_i for $i \in \{1, 2\}$ are also determined this way. For example, rather than solving for the cost of agency τ_2 as a real-valued function of Θ, and then substituting observationally equivalent estimates of θ^* into that function, we exploit the equilibrium condition that $\tau_2 = E[w_s] - (1 + \iota)\gamma^{-1} \ln \alpha_2$, along with the identification condition that $\alpha_2^* = \alpha_2(\gamma^*)$.

This section reports our structural estimates evaluating the effects of SOX on the agency problem between shareholders and their CEO, as character-

[20] The main sample analyses consider three dimensions of heterogeneity including industrial sector (primary, consumer goods, and service), firm size (binary), and financial leverage (binary), which together generate 12 firm categories. To maintain a sufficient amount of observations in each firm category, the restricted sample only differentiates by industrial sector and firm size, reducing the number of firm categories to six.

ized by our principal–agent model. First we report our estimates of the risk aversion parameter. Then we quantify the two measures of conflict of interest, the two components of compensation, and offer some remarks on differences that emerged across firms types.

8.1 PARAMETER ESTIMATES

At the outset, we note that as the confidence region of a risk aversion parameter is nonempty in both eras, the data do not reject the model. As mentioned in the Introduction, some public figures thought risk attitudes might shift in response to the SOX legislation, so we tested the null hypothesis that γ is constant across the two eras. The 95% confidence region of the risk aversion parameter for our sample is almost identical in both eras; up to four significant figures, every observationally equivalent risk aversion parameter for one era appears in the other's confidence region.[21] Therefore, we do not reject the null hypothesis of no change in risk attitude, and attribute all change in the welfare measures we present to differences in nonpecuniary returns to CEOs and/or differences in the distribution of excess returns, from both working and shirking.[22] In light of the test results, we impose the restriction that the risk aversion parameter is constant over the entire sample period for the remainder of the study. The common 95% region is an interval (0.0695, 0.6158), which in economic terms amounts to the CEO paying between \$34,722 and \$290,206 to avoid an equiprobable gamble of losing or winning \$1,000,000.[23]

8.2 CONFLICT OF INTEREST

Shareholders incur a gross loss of ρ_1, defined by (5), if the CEO shirks rather than works. The certainty equivalent differential gain to a CEO from shirking instead of working in period t is ρ_2, defined by (6). Let Δ denote the change from the pre-SOX to the post-SOX era. Thus $\Delta\rho_1$ measures the effects of SOX on the sensitivity of the firm's performance with respect to the agent's effort, and $\Delta\rho_2$ measures the effect of SOX on the cost of effort to the agent. To the extent SOX provided more protection to shareholders

---

[21] The confidence region for the full sample of the post-SOX period covers a broader range (0.0616, 0.2335) than that of the pre-SOX period (0.0784, 0.2335), and 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. See table S7 in the online appendix for a more detailed report of these findings.

[22] 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 are that Nekipelov assumes linear contracting, approximates compensation with salary and bonus, and estimates his model off a different sample population.

[23] These estimates are in line with previously published work (see Gayle and Miller [2009a, 2009b, 2015], Gayle, Golan, and Miller [2015]).

TABLE 3

Gross Loss to Shareholders from CEO Shirking (in %): $\rho_1 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}\{x[1 - g_{s,pre}(x)]\}$

Panel A: Entire sample

	(Size, D/E)	ρ_1	$\Delta\rho_1$
Primary	(S, S)	(11.09, 11.31)	(−2.69, −1.96)
	(S, L)	(9.20, 11.70)	(−6.92, −4.75)
	(L, S)	(7.70, 9.67)	(−2.82, −2.10)
	(L, L)	(4.97, 5.70)	(−1.96, −1.95)
Consumer	(S, S)	(15.65, 16.28)	(−9.16, −8.72)
	(S, L)	(9.13, 13.15)	(2.12, 12.21)
	(L, S)	(6.60, 9.13)	(−0.40, 1.54)
	(L, L)	(5.46, 7.58)	(−2.68, −2.11)
Service	(S, S)	(19.64, 20.25)	(−8.93, −6.34)
	(S, L)	(10.48, 13.94)	(−3.02, −1.03)
	(L, S)	(17.25, 19.76)	(−16.59, −15.37)
	(L, L)	(7.63, 10.11)	(−5.97, −5.07)

Panel B: Subsample

		Compliant (C)		Noncompliant (NC)		NC versus C
	Size	(1) ρ_1	(2) $\Delta\rho_1$	(3) ρ_1	(4) $\Delta\rho_1$	(5) DID
Primary	S	(9.62, 9.65)	(−3.18, −3.17)	(14.14, 14.19)	(−4.45, −4.04)	(−1.27, −0.86)
	L	(4.84, 4.89)	(−0.60, −0.60)	(7.45, 7.51)	(−1.40, −1.38)	(−0.80, −0.78)
Consumer	S	(14.48, 14.56)	(−8.18, −8.11)	(16.44, 16.45)	(−8.46, −8.38)	(−0.35, −0.21)
	L	(5.91, 6.01)	(1.73, 1.89)	(9.61, 10.20)	(−4.86, −4.39)	(−6.74, −6.12)
Service	S	(20.94, 20.98)	(−9.13, −9.07)	(18.07, 18.24)	(−6.08, −6.01)	(3.05, 3.06)
	L	(10.96, 11.05)	(−5.93, −5.83)	(12.60, 12.75)	(−9.83, −9.78)	(−4.01, −3.85)

Note that ρ_1 is the reduction of excess returns due to CEOs shirking. The confidence region is estimated for the single common bond price (16.4). We use the region of the risk aversion parameter commonly shared by the two periods to calculate the values in each column. Panel A compares the pre-SOX and post-SOX values in the entire sample. Panel B reports the estimates for the treatment group and the control group in the DID analysis. The column "ρ_1" reports the estimates of the welfare measure in the pre-SOX period for each firm category. The column "$\Delta\rho_1$" reports the changes of the welfare measure across SOX. The noncompliant firms are the treatment group, including firms who missed at least one of the following criteria before SOX: (1) a majority independent board, (2) an entirely independent audit committee, and (3) an entirely independent compensation committee. The rest of our sample is denoted as compliant firms and used as the control group.

from shirking managers, we would predict that $\Delta\rho_1 < 0$.[24] Similarly, if the penalties imposed by SOX suppressed the private benefits to the CEO from shirking ($\Delta\alpha_1 > 0$) without imposing administrative burdens on working managers ($\Delta\alpha_2 = 0$), then $\Delta\rho_2 < 0$.

Gross loss to shareholders from CEO shirking (ρ_1). Panel A of table 3 reports, for the entire sample, the expected gross loss shareholder incur from the CEO shirking instead of working. Denoted by ρ_1, it is reported as a%age of market value for the pre-SOX era, and how that changes with the im-

[24] Chhaochharia and Grinstein [2007] conducted an empirical analysis of the effect of SOX on firm value and found it increased the value of firms that were less compliant with SOX relative to those that were more compliant.

plementation of SOX. Similar to estimates found in previous studies, they range from 5.0% to 20.2% per year.[25] As depicted, the variation explained by the firm category far outweighs the indeterminacy from observational equivalence that arises from set identification. In particular, large firms tend to lose proportionately less than small firms from a shirking CEO. Given size and leverage, shareholders in firms belonging to the service sector have the most to lose if the financial incentives to their CEOs are not aligned.

The most striking result of panel A of table 3 is that $\triangle\rho_1 < 0$ in all but two categories, evidence that implementing SOX reduced the gross loss shareholders would bear when managers shirked. Again the heterogeneity across firm types far exceeds heterogeneity across observationally equivalent primitives within firm type.

The DID exercise reported in panel B shows that, compared with compliant firms, noncompliant firms not only faced a larger ρ_1, but for the most part also experienced a more significant decrease in ρ_1 after SOX. That is, the intervals in five out of six firm categories, displayed in the column on the far right, are negative. Overall, SOX reduced the expected loss a CEO would impose on his firm by not pursuing a goal of expected value maximization, more so for noncompliant firms.

Benefit to CEO from shirking (ρ_2). From the perspective of the CEO, the conflict of interest is measured by ρ_2, the compensating differential equalizing the CEO lifetime annuitized utility of working rather than shirking. Panel A of table 4 shows that our estimates are a tiny fraction of the losses shareholders would incur from shirking, ranging between \$1.1 million (U.S. 2006) and \$10.8 million annually. In other words, the benefits to a CEO from shirking are far outweighed by the costs to shareholders, and there are huge gains from trade by resolving this conflict through the appropriate use of financial incentives, a finding that echoes previous results (Margiotta and Miller [2000], Gayle and Miller [2009a, 2009b, 2015], Gayle, Golan, and Miller [2015]).

The far right column in panel A of table 4 shows that introducing SOX had an uneven effect on the incentives of the CEO to shirk; the differential mostly declined in the firm categories where it was relatively high. A clearer picture emerges from the DID analysis, displayed in panel B of table 4. Two patterns are evident, and they reinforce each other. First, within each sector, and conditional on whether a firm is compliant or not, SOX dampened the differences in the conflict of interest from the CEOs perspective between small and large firms. Second, in the primary and consumer goods sectors, where compliant firms had less conflict of interest on this measure than noncompliant firms in the pre-SOX era, the

---

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

TABLE 4

Benefit to CEO from Shirking (in \$ thousands): $\rho_2 \equiv b_{t+1}[(b_t - 1)\gamma]^{-1}\ln(\alpha_{2,pre}/\alpha_{1,pre})$

Panel A: Entire sample

	(Size, D/E)	ρ_2	$\triangle\rho_2$
Primary	(S, S)	(2262, 2879)	(122, 221)
	(S, L)	(1108, 1299)	(−57, −24)
	(L, S)	(1459, 1904)	(1716, 2125)
	(L, L)	(1395, 1665)	(100, 380)
Consumer	(S, S)	(5325, 7854)	(−3213, −2091)
	(S, L)	(1947, 2596)	(287, 476)
	(L, S)	(3314, 5727)	(18, 792)
	(L, L)	(2976, 5384)	(−1078, −654)
Service	(S, S)	(4024, 5728)	(−780, −487)
	(S, L)	(1549, 2455)	(67, 446)
	(L, S)	(6492, 10841)	(−7697, −5721)
	(L, L)	(4286, 6472)	(−2041, −1985)

Panel B: Subsample

		Compliant (C)		Noncompliant (NC)		NC versus C
	Size	(1) ρ_2	(2) $\triangle\rho_2$	(3) ρ_2	(4) $\triangle\rho_2$	(5) DID
Primary	S	(1610, 1699)	(668, 691)	(3281, 3542)	(−496, −382)	(−1187, −1049)
	L	(780, 830)	(1335, 1456)	(2541, 2719)	(1024, 1069)	(−387, −311)
Consumer	S	(4403, 4795)	(−668, −586)	(6556, 7224)	(−3090, −2644)	(−2504, −1977)
	L	(2473, 2843)	(911, 1501)	(5831, 6715)	(−937, −745)	(−2246, −1848)
Service	S	(5013, 5522)	(−1102, −918)	(2824, 3188)	(2213, 2602)	(3131, 3703)
	L	(6988, 7673)	(−4640, −4465)	(5887, 6857)	(−3104, −3070)	(1370, 1570)

The confidence region is estimated for the single common bond price, 16.4. We use the region of the risk aversion parameter commonly shared by the two periods to calculate the values in each column. Panel A compares the pre-SOX and post-SOX values in the entire sample. Panel B reports the estimates for the treatment group and the control group in the DID analysis. The column "ρ_2" reports the estimates of the welfare measure in the pre-SOX period for each firm category. The column "$\triangle\rho_2$" reports the changes of the welfare measure across SOX. The noncompliant firms are the treatment group, including firms who missed at least one of the following criteria before SOX: (1) a majority independent board, (2) an entirely independent audit committee, and (3) an entirely independent compensation committee. The rest of our sample is denoted as compliant firms and used as the control group.

difference in $\triangle\rho_2$ between noncompliant and compliant firms fell; in the service sector, where noncompliant firms had less conflict of interest than compliant firms in the pre-SOX era, it rose.

Many rules and regulations applying to all walks of life compel the subjects to discard heterogeneous or idiosyncratically individualistic behavior. Here, in this model, even though shirking is not on the equilibrium path, a similar phenomenon appears. By increasing the penalties of illegal behavior, SOX may have discouraged questionably legal behavior, channeling the type of shirking that would occur if CEOs lack proper incentives.

8.3 COMPENSATION COSTS

In equilibrium, the CEO works and is truthful, implying total expected compensation can be expressed as τ_0, defined in (7), and decomposed into administrative costs, denoted by τ_1, defined in (8) and agency costs, $\tau_2 \equiv \tau_0 - \tau_1$.

Denote the change in the administrative payments from the pre-SOX to post-SOX eras by $\Delta\tau_1$. Thus, $\Delta\tau_1$ measures the effect of SOX on the opportunity cost of the outside option. This definition of administrative costs should be broadly interpreted within our model: for example each new law opens the possibility of unjust prosecution, bringing legal jeopardy to executives by becoming exposed following legislative enactment, and hence raises this compensating differential. Thus, increased regulations and penalties suggest that $\Delta\tau_1 > 0$. For example a senior officer of the U.S. Chamber of Commerce, Michael Ryan, expressed concerns that "the time and energy required by SOX can be a distraction. . . . The amount of time management is spending on the process to comply with SOX takes them away from running the business, increasing sales and developing new products" (Farrell [2007]). On the other hand, outsourcing some responsibilities for preventing fraudulent behavior from the internal controls of the firm to the government, funded by taxpayers rather than shareholders, might also reduce the administrative burden borne by the CEO, who in the post-SOX era faces stiffer penalties for malfeasance instead.

Denote the change in the risk premium from the pre-SOX to post-SOX eras by $\Delta\tau_2$. It measures the change in expected costs the firm paid to resolve its agency issues when SOX was implemented. 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. Alternatively, the provisions might complicate communication between shareholders and management, increasing the cost of revealing the good state.

Administrative costs (τ_1). Administrative costs, denoted by τ_1, are a compensating differential a CEO is paid, relative to his outside option, when he has no private information and his actions are observed, that accounts for the nonpecuniary costs and benefits of his position. Panel A of table 5 shows these costs vary significantly by sector and firm type. For example, in the pre-SOX regime, the 95% confidence region for the administrative cost of (S, L) firms in the primary sector is covered by the interval ranging from \$0.9 to \$1.0 million. In (L, S) firms belonging to the service sector, the corresponding region is covered by the interval ranging from \$7.9 to \$11.0 million. Every firm category within the primary sector experienced increased administrative costs of between \$2.3 and \$4.6 million and every category within the service sector experienced declines between \$0.5 and \$4.1 million, while zero change is a consistent estimate for three out of the four consumer sector groups.

These results broadly reflect our findings in table 1, which show that mean CEO compensation significantly increased in every subcategory within the primary sector following the passage of SOX, but did not increase substantially in any subcategory of the service sector. They also

TABLE 5

Administrative Costs (in \$ thousands): $\tau_1 \equiv \gamma^{-1}\frac{b_{t+1}}{b_t - 1}\ln \alpha_{2,pre}$

Panel A: Entire sample

	(Size, D/E)	τ_1	$\triangle\tau_1$
Primary	(S, S)	(1440, 1860)	(2285, 2455)
	(S, L)	(872, 1043)	(3182, 3209)
	(L, S)	(3699, 4079)	(4113, 4648)
	(L, L)	(3727, 3994)	(2829, 3165)
Consumer	(S, S)	(−279, 1282)	(−437, 31)
	(S, L)	(931, 1407)	(−25, 110)
	(L, S)	(2467, 4560)	(−1041, 590)
	(L, L)	(4734, 6766)	(−767, −389)
Service	(S, S)	(2348, 3701)	(−1473, −1153)
	(S, L)	(1877, 2642)	(−462, −112)
	(L, S)	(7942, 10951)	(−4129, −3888)
	(L, L)	(7684, 9374)	(−1738, −1262)

Panel B: Subsample

		Compliant (C)		Noncompliant (NC)		NC versus C
	Size	(1) τ_1	(2) $\triangle\tau_1$	(3) τ_1	(4) $\triangle\tau_1$	(5) DID
Primary	S	(1393, 1456)	(2988, 3060)	(1917, 2095)	(1358, 1417)	(−1702, −1572)
	L	(3614, 3674)	(3424, 3526)	(4423, 4566)	(3686, 3777)	(251, 263)
Consumer	S	(977, 1232)	(410, 480)	(668, 1117)	(−1279, −1063)	(−1760, −1473)
	L	(5160, 5607)	(277, 720)	(6285, 6999)	(−1158, −1071)	(−1792, −1435)
Service	S	(3481, 3882)	(−2458, −2305)	(2650, 2959)	(1074, 1298)	(3379, 3755)
	L	(9732, 10212)	(−1667, −1561)	(9335, 10058)	(−2862, −2753)	(−1202, −1192)

The confidence region is estimated for the single common bond price, 16.4. We use the region of the risk aversion parameter commonly shared by the two periods to calculate the values in each column. Panel A compares the pre-SOX and post-SOX values in the entire sample. Panel B reports the estimates for the treatment group and the control group in the DID analysis. The column "τ_1" reports the estimates of the welfare measure in the pre-SOX period for each firm category. The column "$\triangle\tau_1$" reports the changes of the welfare measure across SOX. The noncompliant firms are the treatment group, including firms who missed at least one of the following criteria before SOX: (1) a majority independent board, (2) an entirely independent audit committee, and (3) an entirely independent compensation committee. The rest of our sample is denoted as compliant firms and used as the control group.

echoes concerns voiced by practitioners in the primary sector that some provisions in SOX increase the workload of CEOs. An article on the Energybiz Magazine (Anand and Schwind [2006, p. 10]) discussing the SOX challenges confronting the energy industry mentioned that "Depending on the area of focus (oil, gas, electricity, nuclear, wind, hydro, etc.), the industry contends with many regulations and regulatory bodies… Add to that the Sarbanes–Oxley legislation with internal control audit and disclosure requirements. Energy companies find themselves being pulled in many directions to meet all of their regulatory obligations. Sometimes those requirements appear to conflict with each other."

Columns 2 and 4 in panel B of table 5 show that although administrative costs in the primary sector for both compliant and noncompliant firms increased, the picture is less clear in the other two sectors. For example,

TABLE 6

Agency Cost (in $ thousands): $\tau_2 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}[w_{s,pre}(x)] - \tau_1$

Panel A: Entire sample

	(Size, D/E)	τ_2	$\triangle\tau_2$
Primary	(S, S)	(56, 477)	(20, 190)
	(S, L)	(22, 194)	(3, 30)
	(L, S)	(50, 430)	(76, 611)
	(L, L)	(35, 302)	(43, 379)
Consumer	(S, S)	(222, 1783)	(−527, −59)
	(S, L)	(65, 542)	(21, 156)
	(L, S)	(302, 2395)	(182, 1812)
	(L, L)	(290, 2323)	(81, 459)
Service	(S, S)	(187, 1540)	(−360, −41)
	(S, L)	(105, 869)	(45, 395)
	(L, S)	(416, 3425)	(113, 355)
	(L, L)	(233, 1924)	(53, 529)

Panel B: Subsample

		Compliant (C)		Noncompliant (NC)		NC versus C
	Size	(1) τ_2	(2) $\triangle\tau_2$	(3) τ_2	(4) $\triangle\tau_2$	(5) DID
Primary	S	(32, 95)	(37, 108)	(93, 270)	(−90, −32)	(−199, −69)
	L	(31, 91)	(53, 155)	(73, 216)	(45, 136)	(−19, −8)
Consumer	S	(133, 388)	(45, 115)	(237, 687)	(−333, −117)	(−448, −161)
	L	(233, 681)	(240, 683)	(374, 1088)	(58, 144)	(−539, −182)
Service	S	(209, 610)	(−232, −80)	(161, 470)	(126, 349)	(206, 582)
	L	(248, 728)	(47, 153)	(380, 1104)	(75, 185)	(28, 38)

The confidence region is estimated for the single common bond price, 16.4. We use the region of the risk aversion parameter commonly shared by the two periods to calculate the values in each column. Panel A compares the pre-SOX and post-SOX values in the entire sample. Panel B reports the estimates for the treatment group and the control group in the DID analysis. The column "τ_2" reports the estimates of the welfare measure in the pre-SOX period for each firm category. The column "$\triangle\tau_2$" reports the changes of the welfare measure across SOX. The noncompliant firms are the treatment group, including firms who missed at least one of the following criteria before SOX: (1) a majority independent board, (2) an entirely independent audit committee, and (3) an entirely independent compensation committee. The rest of our sample is denoted as compliant firms and used as the control group.

compliant firms in the consumer sector experienced increased administrative costs, noncompliant firms a decrease.

Agency costs (τ_2). Agency costs, τ_2, measure the gross costs that shareholders would be willing to pay for perfect monitoring and thus avoid the penalties induced by the incentive compatibility, truth-telling and sincerity constraints. Table 6 reports the 95% confidence region for the observational equivalent values of τ_2 in the pre-SOX era and its change $\triangle\tau_2$. While agency costs are small in some firm categories, as low as $22,000 per year in (S, L) firms within the primary sector, these costs are much greater in the service sector: between $105,000 and $3.425 million.

Panel A shows that SOX increased agency costs in ten of twelve firm categories. For the most part, the absolute values of the changes are small to moderate, exceeding $1 million only in the (L, S) consumer goods category. However, as a proportion of the levels, they are quite substantial; the estimated upper bound on $\triangle\tau_2$ is at least as large as the lower bound of τ_2 in several categories. Panel B of table 6 confirms the prevalent increase for a different partition of firms. Broadly speaking, it shows that both the compliant firms and the noncompliant firms experience an increase in aggregate agency costs.

8.4 DIFFERENCES ACROSS FIRM TYPES

Finally, we compare differences that emerge between the sectors, firm sizes and the debt/equity ratio.

Sectoral differences. Tables 3–6 not only establish some overall trends that can be ascribed to SOX, such as reduced losses to shareholders from shirking CEOs, but they also display notable differential effects within and between sectors. For example, reviewing panel A of these four tables reveals that only within the primary sector did administrative and agency costs increase in every firm category. Within the other two sectors either the administrative and agency costs had offsetting effects, or both tended to fall.

Comparing column 5 in panel B of tables 4 and 6 is also illuminating. In large (small) firms belonging to the primary and consumer sectors, the agency costs increased less (decreased more) in noncompliant firms than in compliant firms, whereas in the service sector the cost increase was greater for noncompliant firms. A factor contributing to this sign reversal is found in table 4. Within the service sector, the benefit to shirking declined in compliant sectors but increased in noncompliant firms; within the other two sectors, that benefit for noncompliant firms either declined more than or increased less than for compliant firms. We deduce the change in the benefit of shirking for different firm types helps explain the shifts in agency costs.

Does firm size matter?. As a percentage of firm value, gross loss incurred from the CEO shirking is higher for small firms than for large; the two panels of table 3 show this result in every size pair of firm type category. Panel A of table 4 shows that within each sector the conflict of interest from the CEOs perspective is greater in small firms. This result is, however, sensitive to the way firms are categorized. In particular, panel B shows that, conditioning on the compliance indicator variable but averaging over the capital structure, the CEO of a large firm in the service sector would benefit more than the CEO of a small one. Table 5 illustrates yet another difference between large and small firms: the administrative costs of the latter are lower than the former.

Agency costs increased more, or did not decline as much, in large firms as in small firms. Reviewing both panels in table 6, there is only one exception to this pattern, noncompliant firms in the service sector. The differential effect of size on administrative cost changes with the implementation SOX is not one-sided. Overall our results do not support the view that SOX was disproportionately burdensome on small firms. Nevertheless, we qualify this finding with a remark that our sample comprises relatively large firms. The smallest firms in our sample exceed the threshold for compliance with SOX, and are typically larger than privately held companies also subject to the provisions of SOX, and those trading only in dark pools (that are not accessible to the investing public).

Effects of firm leverage. Across all types of firms categorized by size and sector, both measures of the conflict of interest are greater in firms with low debt/equity ratios than in firms with high debt/equity ratios. Thus, ρ_1 is greater for low leveraged firms than highly leveraged firms; ρ_2 is lower in firms with high debt/equity ratios than in firms with low debt/equity ratios. Similarly, without exception, agency costs are lower in highly leveraged firms. We attribute the size of the conflict of interest between owners and managers, and also the agency cost, mainly to the nature of the production and information technology, whereas the debt/equity ratio is primarily a choice about the means of governance. Under this interpretation, firms with less conflict of interest and lower agency costs are more attractive to bondholders. Nevertheless, we recognize the line of causality might also run the other way, bondholders demanding firms utilize technologies that reduce the conflict of interest and agency costs. Implementing SOX seemed to blur these sharp differentials somewhat, although this qualification is itself tentative, because the optimal financial leverage can take several years to reach a new steady state following the introduction of disruptive wide ranging regulations to governance.

9. *Conclusion*

SOX was a legislative response by the U.S. government to corporate governance failures in many prominent companies. This paper is an empirical analysis undertaken with panel data constructed from S&P 1500 firms to gauge the effects of SOX on internal agency issues, focusing specifically on CEO compensation plans and their incentives. As motivation, we show that after SOX was enacted, there were significant changes in the relation between a firm's excess returns and CEO compensation, and also the underlying distribution of excess returns. In the dynamic model of optimal contracting we estimate, a conflict of interest between shareholders and the CEO arises because they have different objectives, many actions taken by the CEO on behalf of shareholders are noncontractible, and in addition CEOs have valuable information about the firm that shareholders are not privy to. 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.

We estimate the model using data on CEO compensation, stock returns, and accounting disclosures, controlling for firm categories and aggregate conditions. In matching the variables in the model to the data, accounting disclosures serve as a proxy for the unverifiable discretionary messages sent by the CEO to shareholders about the future profitability of the firm. Variation in our data can be accounted for without resorting to an explanation based on changing tastes. We do not find evidence that SOX changed preferences for risk-taking by a CEO. Also, there is no systematic pattern that the impact of SOX varies by firm size, that is within the select group of S&P 1500 firms we analyze.

We adopt four measures to quantify the impact of SOX on the inherent conflict of interest between shareholders and their CEO, and the cost of resolving that conflict through incentives embedded with the CEO compensation package. The conflict of interest can be summarized by two measures, how much shareholders lose if the CEO follows his own goals instead of acting on behalf of shareholders' interests, and how much the CEO would benefit from doing so. We find, first, that SOX reduced the loss shareholders would incur if the CEO disregarded the incentives the compensation plan gives and ran the firm purely with own interests in mind (as they would if paid a fixed wage); this reduction was most noticeable in noncompliant firms. Second, the expected benefits accruing to CEOs from deviating from the firm's goals became less dependent on firm type: implementing laws that apply uniformly replace, to some extent, opportunities for malfeasance and internal penalties that vary with each firm. The conflict is resolved in equilibrium with an expected cost shareholders pay. It has two components, an administrative cost tasking the CEO in ways that reduce the scope for acting against shareholder interests, and a risk premium in exchange for exposing the CEO to uncertainty about the excess returns to the firm (inducing CEO behavior benefiting shareholders). With the enactment of SOX, administrative costs increased for almost all the firm types in the primary sector, but effects in the other sectors varied with firm type. However, agency costs increased across the board; firms became more reliant on incentive pay, and we find some evidence that the increased regulation made it more costly to motivate CEOs.

APPENDIX A

THE LAGRANGIAN OPTIMIZATION PROBLEM

This appendix lays out the Lagrangian optimization problem and the associated first-order conditions (FOC) to: (1) complete the characterization of the optimal contract, and (2) derive inequalities that compare the annuitized expected utilities from compensation under HMH with their counterparts under PMH. The Lagrangian for maximizing the expected value of $v_s(x)$ subject to the constraints on participation, incentive compatibility, truth-telling and sincerity, are

$$\mathcal{L} = \sum\nolimits_{s=1}^{2} \int \varphi_s \ln\left[v_s(x)\right] f_s(x)\mathrm{d}x \qquad \text{(A.1)}$$
$$+\eta_0\left[1 - \sum\nolimits_{s=1}^{2} \int \varphi_s \alpha_2^{1/(b-1)} v_s(x)\, f_s(x)\mathrm{d}x\right]$$

$$+\sum_{s=1}^{2}\eta_s\varphi_s\int\big(g_s(x)-(\alpha_2/\alpha_1)^{1/(b-1)}\big)v_s(x)f_s(x)\mathrm{d}x$$
$$+\eta_3\int\big[v_1(x)-v_2(x)\big]f_2(x)\mathrm{d}x$$
$$+\eta_4\int\big[(\alpha_1/\alpha_2)^{1/(b-1)}v_1(x)g_2(x)-v_2(x)\big]f_2(x)\mathrm{d}x,$$

where η_0 is the Kuhn–Tucker multiplier for (11), η_1 and η_2 are the multipliers associated with (12), η_3 is the multiplier for (13), and η_4 for (14), and for expositional convenience we scale the multipliers to φ_s. Differentiating $\mathcal{L}$ with respect to $v_s(x)$ evaluated at each point (s,x), dividing the equations through by $f_s(x)$, and recalling from (4) that $h(x)\equiv\varphi_2 f_2(x)/\varphi_1 f_1(x)$, yields the FOC:

$$\big[v_1^o(x)\big]^{-1}=\eta_0\alpha_2^{1/(b-1)}-\eta_1\big[g_1(x)-(\alpha_2/\alpha_1)^{1/(b-1)}\big]$$
$$-\eta_3 h(x)-\eta_4(\alpha_1/\alpha_2)^{1/(b-1)}g_2(x)h(x), \tag{A.2}$$

$$\big[v_2^o(x)\big]^{-1}=\eta_0\alpha_2^{1/(b-1)}-\eta_2\big[g_2(x)-(\alpha_2/\alpha_1)^{1/(b-1)}\big]+\eta_3+\eta_4. \tag{A.3}$$

As the objective function is concave and the constraints are linear, appealing to the Kuhn–Tucker theorem, there is a unique stationary point. Therefore, the first-order conditions and the complementary slackness conditions fully characterize the solution to the optimal contracting problem.

To compare compensation in the PMH and HMH models, we first prove $\eta_0=1$. To see this, multiply $(A.2)$ by $\varphi_1 v_1(x)f_1(x)$ and $(A.3)$ by $\varphi_2 v_2(x)f_1(x)$, integrate the resulting equations over x, and use the fact that (11) holds with equality (proved by a simple contradiction argument). As neither (13) nor (14) apply to PMH, the first-order conditions for PMH now simplify to

$$\big[\widetilde{v}_s(x)\big]^{-1}=\alpha_2^{1/(b-1)}-\widetilde{\eta}_1\big(g_s(x)-(\alpha_2/\alpha_1)^{1/(b-1)}\big),$$

where $\widetilde{v}_s(x)$ is optimal for PMH. Multiplying by $\widetilde{v}_s(x)f_s(x)$ and integrating over x yields

$$1=\alpha_2^{1/(b-1)}\int\widetilde{v}_s(x)f_s(x)\mathrm{d}x \tag{A.4}$$

for each $s\in\{1,2\}$. For comparison purposes, following the same steps in the HMH yield:

$$1=\alpha_2^{1/(b-1)}\int\big[v_1^o(x)f_1(x)-\eta_3 h(x)-\eta_4(\alpha_1/\alpha_2)^{1/(b-1)}g_2(x)h(x)\big]\mathrm{d}x. \tag{A.5}$$

$$1=\alpha_2^{1/(b-1)}\int v_2^o(x)f_1(x)\mathrm{d}x+\eta_3+\eta_4. \tag{A.6}$$

Denote $\int v(x) f_s(x) \mathrm{d}x$ by $E_s[v(x)]$ and note that at least one of the Kuhn-Tucker multipliers is positive. Upon comparing $(A.4)$ with $(A.5)$ and $(A.6)$:

$$E_2\left[v_2^o(x)\right] < E_2\left[\widetilde{v}_2(x)\right] = E_1\left[\widetilde{v}_1(x)\right] < E_1\left[v_1^o(x)\right].$$

Recall from the definition of $v_r(x)$ given in (10) that lower values of $v_r(x)$ are preferred (as CARA utility functions are negatively signed). In the PMH model, the annuitized expected utilities from compensation in each state is equalized (and exactly compensate the outside opportunity), whereas in the HMH they are not: conditional on truth-telling and working in both states, the annuitized expected utility is higher in the good state than the bad.

REFERENCES

Ait-Sahalia, Y.; P. J. Bickel; and T. M. Stoker. "Goodness-of-Fit Tests for Kernel Regression with an Application to Option Implied Volatilities." *Journal of Econometrics* 105 (2001): 363–412.

Altug, S., and R. A. Miller. "Household Choices in Equilibrium." *Econometrica* 58 (1990): 543–70.

Altug, S., and R. A. Miller. "The Effect of Work Experience on Female Wages and Labour Supply." *Review of Economic Studies* 65 (1998): 45–85.

Anand, S., and R. Schwind. "Energy Sector Wrestles with Sarbanes–Oxley." *ENERGYBIZ Magazine*, March/April 2006, pp. 10–11.

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

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

Ball, R. "Market and Political/Regulatory Perspectives on the Recent Accounting Scandals." *Journal of Accounting Research* 47 (2009): 277–323.

Banerjee, S.; M. Humphery-Jenner; and V. Nanda. "Restraining Overconfident CEOs Through Improved Governance: Evidence from the Sarbanes–Oxley Act." *Review of Financial Studies* 28 (2015): 2812–58.

Bargeron, L. L.; K. M. Lehn; and C. J. Zutter. "Sarbanes–Oxley and Corporate Risk-Taking." *Journal of Accounting and Economics* 49 (2010): 34–52.

Bertomeu, J.; E. Cheynel; E. X. Li; and Y. Liang. "How Pervasive is Earnings Management? Evidence from a Structural Model." *Management Science* (2021): 5145-5162.

Bertomeu, J.; P. Ma; and I. Marinovic. "How Often Do Managers Withhold Information?" *The Accounting Review* 95 (2020): 73–102.

Bertomeu, J.; I. Marinovic; R. A. Miller; and F. Varas. "*An Empirical Model of Long-Term Contracting with Manipulation.*" Washington, Stanford, Carnegie Mellon and Duke Universities Working paper, 2017.

Beyer, A.; I. Guttman; and I. Marinovic. "Earnings Management and Earnings Quality: Theory and Evidence." *The Accounting Review* 94 (2019): 77–101.

Bird, A.; S. A. Karolyi; and T. G. Ruchti. "Understanding the 'Numbers Game'." *Journal of Accounting and Economics* 68 (2019):1-26.

Bolton, P., and M. Dewatripont. *Contract Theory*. Cambridge, MA: MIT Press, 2005.

Breuer, M., and D. Windisch. "Investment Dynamics and Earnings-Return Properties: A Structural Approach." *Journal of Accounting Research* 57 (2019): 639–74.

Carter, M. E.; L. J. Lynch; and S. L. C. Zechman. "Changes in Bonus Contracts in the Post-Sarbanes–Oxley Era." *Review of Accounting Studies* 14 (2009): 480–506.

Center for Audit Quality (CAQ). "The CAQ's Fourth Annual Individual Investor Survey." The Glover Park Group (2010). Available at Center for Audit Quality (CAQ). https://www.thecaq.org/wp-content/uploads/2019/03/2010_caq_main_street_investor_survey.pdf

CHEN, X.; Q. CHENG; and X. WANG. "Does Increased Board Independence Reduce Earnings Management? Evidence from Recent Regulatory Reforms." *Review of Accounting Studies* 20 (2015): 899–933.

CHERNOZHUKOV, V.; H. HONG; and E. TAMER. "Estimation and Confidence Regions for Parameter Sets in Econometric Models." *Econometrica* 75 (2007): 1243–84.

CHEYNEL, E., and M. LIU-WATTS. "A Simple Structural Estimator of Disclosure Costs." *Review of Accounting Studies* 25 (2020): 201–45.

CHHAOCHHARIA, V., and Y. GRINSTEIN. "Corporate Governance and Firm Value: The Impact of the 2002 Governance Rules." *Journal of Finance* LXII (2007): 1789–825.

CHHAOCHHARIA, V., and Y. GRINSTEIN. "CEO Compensation and Board Structure." *Journal of Finance* 64 (2009): 231–61.

CHOI, J. "Accrual Accounting and Resource Allocation: A General Equilibrium Analysis." *Journal of Accounting Research* 59 (2021): 1179-1219.

COATES, J. C., and S. SRINIVASAN. "SOX after Ten Years: A Multidisciplinary Review." *Accounting Horizons* 28 (2014): 627–71.

COHEN, D. A.; A. DEY; and T. Z. LYS. "Real and Accrual-Based Earnings Management in the Pre-and Post-Sarbanes–Oxley Periods." *The Accounting Review* 83 (2008): 757–87.

COHEN, D. A.; A. DEY; and T. Z. LYS. "Corporate Governance Reform and Executive Incentives: Implications for Investments and Risk Taking." *Contemporary Accounting Research* 30 (2013): 1296–332.

COHEN, J. R.; C. HAYES; G. KRISHNAMOORTHY; G. S. MONROE; and A. M. WRIGHT. "The Effectiveness of SOX Regulation: An Interview Study of Corporate Directors." *Behavioral Research in Accounting* 25 (2013): 61–87.

DUCHIN, R.; J. G. MATSUSAKA; and O. OZBAS. "When are Outside Directors Effective?" *Journal of Financial Economics* 96 (2010): 195–214.

ENGEL, E.; R. M. HAYES; and X. WANG. "The Sarbanes–Oxley Act and Firms' Going-private Decisions." *Journal of Accounting and Economics* 44 (2007): 116–45.

FARRELL, G. "Sarbanes–Oxley Law has been a Pretty Clean Sweep." 2007. Available at https://abcnews.go.com/Business/story?id=3428252&page=1.

FILBECK, G.; R. GORMAN; and X. ZHAO. "SOX and the Regulated Firm." *Journal of Accounting and Public Policy* 30 (2011): 526–50.

FUDENBERG, D.; B. HOLMSTROM; and P. MILGROM. "Short-Term Contracts and Long-Term Agency Relationships." *Journal of Economic Theory* 51 (1990): 1–31.

GAO, F.; J. S. WU; and J. ZIMMERMAN. "Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes–Oxley Act." *Journal of Accounting Research* 47 (2009): 459–506.

GAYLE, G. L.; L. GOLAN; and R. A. MILLER. "Gender Differences in Executive Compensation and Job Mobility." *Journal of Labor Economics* 30 (2012): 829–72.

GAYLE, G. L.; L. GOLAN; and R. A. MILLER. "Promotion, Turnover and Compensation in the Executive Labor Market." *Econometrica* 83 (2015): 2293–369.

GAYLE, G. L.; L. GOLAN; and R. A. MILLER. "*Interlocked Executives and Insider Board Members: An Empirical Analysis.*" Carnegie Mellon and Washington Universities, Working paper, 2021.

GAYLE, G. L., and R. A. MILLER. "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" *American Economic Review* 99 (2009a): 1740–69.

GAYLE, G.-L., and R. A. MILLER. "Insider Information and Performance Pay." *CESifo Economic Studies* 55 (2009b): 515–41.

GAYLE, G.-L., and R. A. MILLER. "Identifying and Testing Models of Managerial Compensation." *Review of Economic Studies* 82 (2015): 1074–118.

GOW, I. D.; D. F. LARCKER; and P. C. REISS. "Causal Inference in Accounting Research." *Journal of Accounting Research* 54 (2016): 477–523.

GROSSMAN, S. J., and O. D. HART. "An Analysis of the Principal–Agent Problem." *Econometrica* 51 (1993): 7–45.

GUTHRIE, K.; J. SOKOLOWSKY; and K.-M. WAN. "CEO Compensation and Board Structure Revisited." *Journal of Finance* 67 (2012): 1149–68.

Hall, B. J., and J. B. Liebman. "Are CEOs Really Paid Like Bureaucrats?' *Quarterly Journal of Economics* 103 (1998): 653–91.

Hart, O. "Regulation and Sarbanes–Oxley." *Journal of Accounting Research* 47 (2009): 437–45.

Hochberg, Y. V.; P. Sapienza; and A. Vissing-JO Rgensen. "A Lobbying Approach to Evaluating the Sarbanes–Oxley Act of 2002." *Journal of Accounting Research* 47 (2009): 519–83.

Holmstrom, B., and S. N. Kaplan. "The State of U.S. Corporate Governance." *Journal of Applied Corporate Finance* 15 (2003): 8–20.

Iliev, P. "The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices." *The Journal of Finance* 65 (2010): 1163–96.

Jain, P. K., and Z. Rezaee "The Sarbanes–Oxley Act of 2002 and Capital-Market Behavior: Early Evidence." *Contemporary Accounting Research* 23 (2006): 629–54.

Joo, J. H., and Chamberlain, S. L. "The Effects of Governance on Classification Shifting and Compensation Shielding." *Contemporary Accounting Research* 34 (2017): 1779–811.

Kamar, E.; P. Karaca-Mandic; and E. Talley. "Going-Private Decisions and the Sarbanes–Oxley Act of 2002: A Cross-Country Analysis." *Journal of Law, Economics, and Organization* 25 (2009): 107–33.

Kang, Q.; Q. Liu; and R. Qi. "The Sarbanes–Oxley Act and Corporate Investment: A Structural Assessment." *Journal of Financial Economics* 96 (2010): 291–305.

Khorunzhina, N. and R. A. Miller. "American Dream Delayed: Shifting Determinants of Homeownership." *International Economics Review* 63 (2022): 3-35.

Labro, E., and L. Stice-Lawrence, "Updating Accounting Systems: Longitudinal Evidence from the Healthcare Sector." *Management Science* 66 (2020): 6042-6061.

Leuz, C.; A. Triantis; and T. Y Wang. "Why Do Firms Go Dark? Causes and Economic Consequences of Voluntary SEC Deregistrations." *Journal of Accounting and Economics* 45 (2008): 181–208.

Li, C. "Are Top Management Teams Compensated as Teams? A Structural Modeling Approach." *Management Science* (2021): Forthcoming.

Li, Q., and J. S. Racine. *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ: Princeton University Press, 2007.

Linck, J. S.; J. M. Netter ;and T. Yang. "The Effects and Unintended Consequences of the Sarbanes–Oxley Act on the Supply and Demand for Directors." *Review of Financial Studies* 22 (2009): 3287–328.

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

Lu, J., and W. Wang. "Board Independence and Corporate Investments." *Review of Financial Economics* 24 (2015): 52–64.

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

Malcomsen, J. M., and F. Spinnewye. "The Multiperiod Principal–Agent Problem." *Review of Economic Studies* 55 (1988): 391–407.

McClure, C. "How Costly Is Tax Avoidance? Evidence from Structural Estimation." University of Chicago, Working paper, 2019.

Merz, A. "What Have We Learned from SFAS 123r and IFRS 2? A Review of Existing Evidence and Future Research Suggestions." *Journal of Accounting Literature* 38 (2017): 14–33.

Mirrlees, J. A. "The Theory of Moral Hazard and Unobservable Behaviour." Mimeograph, Oxford, 1975.

Miller, R. A., and H. Sieg. "A Microeconometric Comparison of Household Behavior between Countries." *Journal of Business* 15 (1997): 237–53.

Myerson, R. "Optimal Coordination Mechanisms in Generalized Principal–Agent Problem." *Journal of Mathematical Economics* 10 (1982): 67–81.

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

Nekipelov, D. "Empirical Content of a Continuous-Time Principal–Agent Model: The Case of the Retail Apparel Industry." Working paper, University of California at Berkeley, 2010. Available at http://eml.berkeley.edu/ñekipelov/pdf/papers/paper12.pdf.

Protiviti Risk and Insurance Consulting Internal Audit. "2012 Sarbanes–Oxley Compliance Survey." (2012). https://www.protiviti.com/sites/default/files/united_states/insights/2012-sox-compliance-survey-protiviti.pdf

Rey, P., and B. Salanie. "Long-Term, Short-Term and Renegotiation." *Econometrica* 58 (1990): 597–620.

Romano, J. P., and A. M. Shaikh. "Inference for the Identified Set in Partially Identified Econometric Models." *Econometrica* 78 (2010): 169–211.

Stulz, R. "Managerial Discretion and Capital Structure." *Journal of Financial Economics* 26 (1990): 3–28.

Wintoki, M. B.. "Corporate Boards and Regulation: The Effect of the Sarbanes–Oxley Act and the Exchange Listing Requirements on Firm Value." *Journal of Corporate Finance* 13 (2007): 229–50.

Zakolyukina, A. A.. "How Common Are Intentional GAAP Violations? Estimates from a Dynamic Model." *Journal of Accounting Research* 56 (2018): 5–44.

Zhang, I. X.. "Economic Consequences of the Sarbanes–Oxley Act of 2002." *Journal of Accounting and Economics* 44 (2007): 74–115.

Zwiebel, J. "Dynamic Capital Structure Under Managerial Entrenchment." *American Economic Review* (1996): 1197–215.