Market and Political/Regulatory Perspectives on the Recent Accounting Scandals

Ray Ball

The University of Chicago Booth School of Business
5807 S. Woodlawn Ave
Chicago, IL 60637
Tel. (773) 834 5941
ray.ball@gsb.uchicago.edu

November 23, 2008


Acknowledgments
Helpful comments were gratefully received from Sudipta Basu, John Coates, Ron Dye, Bob Jensen, Clive Lennox, Sam Peltzman, Sarah Zechman, participants at the 2008 Journal of Accounting Research Conference, and especially Christian Leuz (the editor). Financial support from the University of Chicago, Graduate School of Business also is gratefully acknowledged.
Abstract

Not surprisingly, the recent accounting scandals look different when viewed from the perspectives of the political/regulatory process and of the market for corporate governance and financial reporting. We do not have the opportunity to observe a world in which either market or political/regulatory processes operate independently, and the events are recent and not well-researched, so untangling their separate effects is somewhat conjectural. This paper offers conjectures on issues such as: What caused the scandalous behavior? Why was there such a rash of accounting scandals at one time? Who killed Arthur Andersen – the SEC, or the market? Did fraudulent accounting kill Enron, or just keep it alive for too long? What is the social cost of financial reporting fraud? Does the US in fact operate a “principles-based” or a “rules-based” accounting system? Was there market failure? Or was there regulatory failure? Or both? Was the Sarbanes-Oxley Act a political and regulatory over-reaction? Does the U.S. follow an ineffective regulatory model?
1. Introduction

The tsunami of accounting scandals at the beginning of the millennium is well-known. A partial list of companies involved includes AOL, Bristol-Myers Squibb, Cendant, Computer Associates (CA), Conseco, Dynegy, Enron, Federal Home Loan Mortgage Corporation (“Freddie Mac”), HealthSouth, Peregrine Systems, Qwest, Rite Aid, Sunbeam, Tyco, Waste Management, WorldCom, and Xerox, with Enron and WorldCom being the most familiar due to the scope and audacity of their deficient reporting.¹ In Europe, ComROAD AG (Germany), Lernout & Hauspie Speech Products (Belgium), Parmalat (Italy), and Royal Ahold (Netherlands) achieved similar notoriety during the period, and the United Kingdom earlier had been rocked by scandals in the 1960s and again in the late 1980s and early 1990s (notably, Polly Peck). Nevertheless, the clear epicenter of the 2001-02 scandals was in the United States.

The scandal waves spread widely and quickly. Resulting damage included: a decline in the worldwide reputation of a wide variety of United States institutions, including U.S. GAAP, auditors, security analysts, regulators and financial markets generally; the conviction of over one thousand executives; extensive private litigation and damages awards against companies, managers, auditors and complicit banks; the most substantial increase in the regulation of U.S. public financial reporting in seventy-five years, under the Sarbanes-Oxley Act of 2002; creation of a Public Company Accounting Oversight Board (PCAOB) with almost unfettered powers to adopt and enforce rules governing the audit industry and to discipline audit firms and employees;² the demise of a one-proud accounting firm; and a forest of largely unhelpful literature.

A variety of press and academic commentators have tried to make sense of the scandals and the events that ensued. What caused them? Why did they occur then? What were the

¹ The precise extent of accounting fraud over the period is unclear. Lists tend to include non-accounting illegalities (such as insider trading, obstruction of justice and money laundering), bankruptcies, and unproven allegations. See http://www.trinity.edu/rjensen/Fraud.htm and http://en.wikipedia.org/wiki/Accounting_scandals.
consequences? Was more regulation the optimal response? Will they happen again? Not surprisingly, these questions have been addressed from many different perspectives. For example, there has been discussion of the roles of social norms, morals, ethics, corporate culture, corporate governance, incentive-based compensation, emphasis on short term profits, audit committees, audit standards, audit firm conflicts of interest due to non-audit revenues, reporting rules, securities laws, and regulatory oversight. Differences in perspective can enrich a debate but – if not made clear – they also can add confusion to it. The objective of this essay therefore is to consider the scandals and several of the important ensuing events from two important and contrasting perspectives: the perspective of the political/regulatory process and the perspective of the market for corporate governance and financial reporting. The general intent is to clarify our understanding of the events and their consequences.

The difference between the market and political/regulatory perspectives on the scandals has more than historical significance: it substantially affects how one interprets not just the scandals, but also the substantial regulation of US corporate governance and financial reporting that ensued. A specific objective of this essay therefore is to assist in the debate on whether the Sarbanes-Oxley Act’s historical increase in regulation was a good idea, in terms of both its scope and its form.

We do not have the opportunity to observe a world in which either market or political/regulatory processes operate independently, so untangling their separate effects is difficult at best. Further, the events still are recent and not well-understood. Insufficient time has passed to allow a distanced perspective or to conduct definitive research, the implication being that opinions tend to be based on only anecdotal or small-sample evidence. Thus, while the following analysis is based where possible on the evidence and on hopefully valid reasoning, much of it is conjectural.

A further caveat is in order. When comparing the political/regulatory and market views of the accounting scandals, I am much influenced by the Stigler (1964, 1971) - Peltzman (1976)
skepticism for regulation. From this perspective, the political process that creates and monitors regulation tends to be captured by the regulated industry and organized special-interest groups, so it typically does not promote general social welfare. Further, even if it attempts to, regulation does not always succeed in promoting social welfare, in part because it is enforced by poorly informed and incented government employees. Hopefully, this skeptical perspective will provide some counterpoint to the post-scandal rush to regulate governance and reporting.

The following section describes and interprets the scandalous events of 2001-2002. Sections 3 and 4 analyze the political/regulatory and market responses to the events, in an attempt to assess their relative roles. Section 5 argues that regulation encourages a rules-based approach to U.S. financial reporting, and the concurrent belief that technical compliance with codified Generally Accepted Accounting Principles (GAAP) is sufficient, whereas markets demand compliance with the wider concept of the accounting principles that are generally accepted. Section 6 examines the principal political/regulatory response to the scandals – the Sarbanes-Oxley Act of 2002 – from a market perspective, and Section 7 speculates on the long run consequences of the Act. Section 8 offers some conclusions.

2. The 2001-02 Accounting Scandals

The analysis of market and political/regulatory perspectives that follows is unavoidably influenced by my interpretation of these scandalous events, including my views on what the scandals were, why they were scandalous, what caused them and why they occurred when they did. This section attempts to make those views more transparent.

2.1 Reporting Negligence and Fraud versus “Earnings Management”

The term “earnings management” is used to describe managers intervening in the reporting of their own financial performance. It encompasses a range of practices, including:

---

Practices that are legal, violate no accounting rules or principles, and are generally viewed as ethical – such as structuring transactions with regard to their effect on the financial statements (leasing being a prominent example);

Practices that are legal, violate no accounting rules or principles, and are viewed by many as ethical – such as timing asset sales to book gains in years with lower profits, and to book losses in years with higher profits (Bartov, 1993);

Practices that are legal, violate no accounting rules or principles, but might violate accepted standards of disclosure – such as giving year-end quantity discounts to major customers, generating sales “pull forwards,” but failing to disclose that they inflate current earnings and borrow against future earnings;

Negligent or grossly negligent financial reporting – such as unwittingly failing to comply with GAAP;

Fraudulent financial reporting– such as knowingly failing to comply with GAAP.

“Earnings management” is a generic term for all of these practices. While they differ in economic, legal and ethical severity, they all undermine the quality of financial reporting to some degree. Earnings management therefore is analogous to a suite of financial system viruses: some are more severe than others, but all weaken the efficiency of the system.

This paper addresses financial reporting that is negligent or fraudulent. Negligence involves managers or auditors making unintentional errors due to factors such as inadequate experience, training, knowledge, supervision or effort. Negligent behavior involves failing to do what is considered reasonable and prudent for a person in their position under the circumstances: in the context of financial reporting, failing to meet accepted accounting, disclosure, or auditing standards. Gross negligence involves reckless disregard for accepted standards. Negligence that

---

4 The term “earnings management” appears to be popular because most financial statement manipulation involves earnings. Nevertheless, some manipulation affects balance sheets or operating cash flow, without affecting earnings.
can be proven to harm others, such as investors, gives rise in law to an actionable tort. Fraud is a more serious transgression, because proving fraud requires establishing \textit{scienter}, defined loosely as intentional wrongdoing.\textsuperscript{5}

In contrast, much academic research on earnings management establishes such a weak burden of proof that earnings management appears almost universal, perhaps reflecting the Hobbesian outlook that has overtaken accounting academe, its limited knowledge of accrual accounting, and the ease with which correlated omitted variables can produce the result one is looking for.\textsuperscript{6} The advantages of focusing on negligent and fraudulent reporting include: a proven case of negligent or fraudulent financial reporting is an institutional “fact,” as distinct from an error-prone academic estimate; reporting negligence and fraud have been shown to have substantial adverse effects on firm values (Palmrose, Richardson, and Scholz, 2004); and the sight of executives being led away in handcuffs under indictment for reporting fraud created more scandal than a whole literature of Jones-model discretionary accrual estimates.\textsuperscript{7}

\textbf{2.2 Why Were The Events Scandalous?}

James Q. Wilson opens \textit{The Moral Sense} with the insightful observation that two conditions must hold for a public scandal to occur: the events must be unusual (i.e., relatively infrequent in occurrence); and they must be shocking (i.e., counter to our norms).\textsuperscript{8} The public pays little attention to frequent events; and there is no scandal arising from infrequent but morally acceptable events.

\textsuperscript{5} The legal standard of proof is subjective. For securities cases, the Private Securities Litigation Reform Act of 1995 set the standard as “strong inference” of \textit{scienter}. In \textit{Tellabs, Inc. v. Makor Issues & Rights, LTD} (21 June 2007), the Supreme Court interpreted this as a requirement to prove "cogent and compelling evidence" of \textit{scienter}.

\textsuperscript{6} “Discretionary” accruals are estimated as the residual component of accruals that cannot be explained by accruals models. These models typically are mechanical, perfect-certainty descriptions of “non-discretionary” accruals, and are not built on credible economic models of why firms hold working capital (e.g., to absorb shocks in production, demand, and cash receipts). They thus are so poorly specified that they explain only a minor proportion of accruals. They seem likely to substantially over-estimate the amount of even the mildest (and presumably most frequent) type of earnings management: practices that are legal, violate no accounting rules or principles, and viewed as ethical.

\textsuperscript{7} On the other hand, an unobserved quantity of negligence and fraud goes undetected or unproven, even though it is argued below that material financial misreporting is difficult to conceal for an extended period.

\textsuperscript{8} Wilson (1993, p.2).
The events of 2001-02 certainly were unusual. The relative frequency of negligent or fraudulent reporting in the U.S. is difficult to estimate precisely, mainly because it is unclear how to define the exact population of reports from it is drawn, in terms of both the total number of firms in existence and the length of the time period used for reference. Nevertheless, the reference population undoubtedly is large relative to the number of convictions for malfeasance, and the relative frequency is correspondingly low, despite the extraordinarily large number that came to light over two years. For example, if 4,000 public companies on average file quarterly reports over 50 years, the population is eight million total filings.9 As noted above, an unknown amount of negligence and fraud goes undetected, so the true incidence undoubtedly exceeds the observed. Nevertheless, even a generous allowance for undetected malfeasance in past years seems unlikely to alter the conclusion that the events of 2001-02 were scandalous in large part because they occur with such a low relative frequency.10

The events of 2001-02 also were scandalous because they clashed so strongly with the norm. Common law countries, chief among them the U.S., have built large debt and equity markets, and correspondingly large public corporate sectors, by following a model founded on high quality financial reporting and disclosure. This allows lenders, shareholders, suppliers and customers to transact at arm’s length with corporations, across a public market, without private (insider) access to information about them. Consequently, high-quality financial reporting has long been viewed as a foundation of the U.S. financial system.11 In contrast, public financial reporting

---

9 In the US, a public company is defined for regulatory purposes as one that is required to register its securities (including stocks and bonds) for sale to the public on a major stock exchange such as NYSE or NASDAQ, or “over the counter” through quotation services such as the OTC Bulletin Board and Pink Sheets. In addition, small companies are publicly traded without being required to register.
10 A similar point was made in a speech on 14 July 2003 by Thomas J. Donohue, the President of the U.S. Chamber of Commerce: “There are more than 17,000 publicly traded companies in the U.S., and only a couple (of) dozen of them have been accused of wrongdoing.” [Link](http://www.uschamber.com/press/speeches/2003/030714tdj_hitachi.htm).
11 In one of several untimely remarks, then Deputy Secretary of the Treasury Lawrence Summers wrote (Financial Times, London, 11 March, 1998): “If one were writing a history of the American capital market, it is a fair bet that the single most important innovation shaping that market was the idea of generally accepted accounting principles.”
has not played such an important economic role in code law countries, including most of Continental Europe. Under the code law model, access to information has been more on a relationship basis, public financial reporting has been lower in quality by observable measures, large corporations have been more likely to be private, and public capital markets have been a comparatively smaller part of the economy.\textsuperscript{12} The reporting transgressions that came to light in 2001-02 therefore were scandalous both because negligent or fraudulent reporting is infrequent and because it so strongly violates expected standards of behavior.

2.3 What Caused Such a Rash of Scandals Come to Light in 2001-02?

How can the notion that scandals are infrequent be reconciled with such a large number of cases coming to light in such a short period? There is limited evidence on what was different about this period, so the following is uncomfortably conjectural. The closest supporting theory and evidence is in Harris and Bromiley (2006) and Kedia and Philippon (2007), and is consistent with the following hypothesis that the economic cycle was a major factor.\textsuperscript{13}

The longest boom in U.S. history ended in March 2001. Bearing this in mind, the following sequence of cycle-related events seems plausible. First, in an extended boom high growth becomes built into performance expectations: into earnings and revenue forecasts, budgets, share prices, option values, investment decisions, and debt commitments. Managers therefore come under peer and financial pressures to deliver strong earnings growth and share market performance. Second, lax practices likely develop in a long boom: corporate monitors (boards, internal and external auditors, analysts, rating agencies, the press, and regulators) come to accept high growth as normal, and there is a risk of “falling asleep at the switch.” Third, the boom “busts,” growth suddenly falters, and around the same time many managers are unable to meet expectations. Fourth, some managers resort to either faking transactions or adopting unaccepted accounting methods to

\textsuperscript{12} Ball, Kothari and Robin (2000).
\textsuperscript{13} See also Alexander and Cohen (1996), Strobl (2008) and Davidson (2008).
disguise their faltering performance. Fifth, with some probability the reporting malfeasance is detected. Sixth, after a series of malpractices are uncovered, they reach widespread public attention, and become scandalous.

Other factors surely were at play. A “rule-checking” mentality appears to have crept over the accounting profession, including the Financial Accounting Standards Board (FASB, a fittingly named successor to the Accounting Principles Board), audit firms, researchers and educators. Under this narrow world view, financial reporting requirements are embodied in rules, not principles, and compliance with rules-based GAAP is sufficient. In several of the scandalous cases (Enron in particular), investors were alleged to have been mislead by financial statements that were in technical compliance with GAAP but that did not reflect the economic substance of the transactions they were reporting. Regulation downplays the basic common law principle that market participants are expected to behave according to the standards that are generally accepted and practiced by informed, competent market participants (here including financial statement users as well as preparers). In contrast, regulators prefer to administer rules. I conjecture that regulation played a substantial role in fostering this rules-based perspective, and hence played an important role in setting the stage for the accounting scandals. This conjecture lies in contrast with the more orthodox view that the scandals occurred in spite of regulation – a view that is difficult to

---

14 The likelihood of detection decreases with post-malfeasance performance. Many schemes to overstate revenues (e.g., “channel stuffing” and “sales creep”) involve borrowing revenue from future periods, thereby increasing present earnings and correspondingly reducing future earnings. They are less likely to be detected if future revenues recover. Further, companies such as Enron and WorldCom, whose false reporting temporarily disguised financial insolvency, are forced into bankruptcy if real performance does not improve, at which point managers acting on behalf of creditors can examine the books.

15 The above is uncomfortably close to Galbraith’s (1961, p. 153) account of the embezzlement scandals that came to light after the 1929 market crash, and contributed to the creation of the SEC: “In good times people are relaxed, trusting, and money is plentiful. But even though money is plentiful, there are always many people who need more. Under these circumstances the rate of embezzlement grows, the rate of discovery falls off, and the bezzle increases rapidly. In depression all this is reversed. Money is watched with a narrow, suspicious eye. The man who handles it is assumed to be dishonest until he proves himself otherwise. Audits are penetrating and meticulous.”

16 This point is developed further in Section 5.7 below.
reconcile with the fact that the U.S. financial reporting market is highly regulated by international standards (Coffee, 2007).

A widely viewed culprit in the scandals was conflicts of interest arising from audit firms performing non-audit work for clients. To many observers, auditors had incentives to appease their clients by giving them favorable audit treatment, in order to retain lucrative consulting engagements. I return to this issue in section 4.3, and conclude that conflicts of interest most likely were not a major factor in the scandals. Nevertheless, user trust in financial reporting requires perceived as well as actual audit independence, so the issue is unlikely to go away.

A less obvious but possibly more important culprit in the scandals was a subtle drift in the auditor-client relationship over time. The relationship always has been unusual in a legal sense. Auditors supply a certification that is used by investors and the public at large, and it is to them that they owe a standard of care. However, they are engaged and compensated by client firm managers, who are interested parties (i.e., prefer a clean audit certificate). The argument here is not that non-audit revenues create a potential conflict of interest, but that audit fees themselves do. Perhaps related to this issue, anecdotal evidence suggests that over time auditors have drifted from a skeptical, adversarial interaction with clients and toward a co-operative approach. The drift possibly accelerated as a consequence of the lax practices that develop in the prolonged economic expansion of the 1990s. It needs emphasizing that this argument is based only on anecdotal evidence; salient elements of the auditor-client are not readily observable, and research on the issue tends to use proxies such as length of auditor tenure that are correlated with important omitted variables. Further, the argument ignores reputational and other mechanisms that constrain such conflicts of interest, as discussed in section 4.3 below.

Coffee (2005) argues that dispersed share ownership makes the U.S. more prone to financial reporting problems than the concentrated ownership systems of Continental Europe,
where scandals are more likely to involve appropriating private benefits of control. While this argument has some legitimacy, it is not consistent with much evidence. First, the argument implies that, if anything, financial reporting should be of higher quality in private companies than public companies, because they have more concentrated ownership. The evidence supports the opposite view (Ball and Shivakumar, 2005; Burgstahler, Hail and Leuz, 2006). Second, the dispersed ownership argument makes no allowance for the evidence that common-law countries such as the U.S. generally set a higher standard for financial reporting. Many of the scandalous U.S. cases involved practices that historically have fallen within legitimate managerial discretion in many code-law countries, such as drawing on “cookie jar” reserves to inflate profits in bad years. In that regard, attributing the scandals to dispersed ownership also seems inconsistent with the evidence on international differences in reporting quality (Ball, Kothari and Robin, 2000; Leuz, Nanda and Wysocki, 2003).

Increased use of stock options in executive compensation in the U.S. was blamed by several commentators, including Coffee (2005) and, notably, Alan Greenspan (2002). While the argument has merit, in that performance-based compensation provides an additional motive to misrepresent performance and was found to be a factor in at least one case (Computer Associates), there is evidence stock options were not a major factor.17

While the number of cases that came to light over 2001-02 is astonishing, and while additional cases presumably went undetected, there are reasons to suspect the press exaggerated the scandal. The press generally over-reports risks. Bailis and MacCoun (1996) document the press over-represents the frequency of tort litigation generally, the frequency of disputes going to court without prior settlement, the frequency of trials won by plaintiffs, and the average size of jury awards. During the scandal period, accounting-related issues were unusually newsworthy, and the

---

17 Compare Erickson et al. (2006) with Burns and Kedia (2006) and Efendi et al. (2007). Option back-dating became a separate scandal, several years later.
press reported even minor issues that previously would not have made the news. This bias was fed by panicking companies that, in a post-Enron “witch hunt,” combed their books for the slightest hint of an accounting issue, which the press duly reported as scandalous. The press thereby exaggerated the extent of an already substantial scandal.

2.4 Why Do Managers Commit Financial Reporting Fraud?

Managers face a variety of well-known financial incentives to meet performance expectations. These include: gaining earnings-based bonuses; increasing their promotion prospects; avoiding termination; avoiding a decline in the value of their stocks, stock appreciation rights, and options; avoiding a downgrade of the company’s debt, which could result in higher borrowing costs and further reductions in earnings; avoiding debt covenant violations that could lead to restrictions on dividend payments, new investment and further borrowing, or accelerated debt repayment; avoiding corporate bankruptcy; and hiding some other fraud (e.g., stolen assets, including cash).

My view, based on mainly anecdotal experience, is that non-financial motives are more powerful than is commonly believed, and sometimes are the dominant reason for committing accounting fraud. An important motivator seems to be maintaining the esteem of one’s peers, ranging from co-workers to the public at large. Enron executives reportedly were celebrities in Houston, and in important places like the White House. Much of this celebrity derived from the firm’s reputation as an extremely successful innovator in the post-deregulation energy market. Much apparently was purchased using shareholders’ seemingly abundant money to support a wide

---

18 Similar reporting behavior occurs for small earthquakes and plane crashes, which are more likely to be covered immediately after a large earthquake or crash brings the issue to public and press attention.
19 Wall Street Journal Online (December 31, 2002) ran a headline “Tyco's Internal Report Finds Extensive Accounting Tricks” over a story that actually reported “the problems weren't large enough to be ‘material’ to the company's overall profits” and “many of the maneuvers fell within accounting rules and didn't require any correction” and also “the report said these were relatively minor and didn't amount to a ‘systemic fraud.’”
20 In addition to providing substantial employment and stock gains to Houston residents, Enron’s name was on Enron Field, it generously supported the Alley Theater, the Houston Ballet, the Houston Symphony, and Museum of Fine Arts, among others. See: Kinzer (2001) and Yardley (2002).
range of charitable causes. Either way, the celebrity would have evaporated if the company’s poor profitability and weak financial condition had been truthfully revealed.\textsuperscript{21} In general, there appear to be substantial private non-monetary benefits of controlling a company for managers committing fraud, even though they presumably are more difficult to prove in court than monetary benefits from bonuses, options, etc.

Financial frauds appear to share three properties:

1. Inability to meet performance expectations;
2. Personal costs – pecuniary or non-pecuniary – of failing to meet expectations; and
3. Being able to convincing oneself that real performance will improve soon.

The last property arises for several reasons. First, if real performance does not subsequently improve, the deception must be repeated to disguise the continuing performance shortfall. Second, fraud is very difficult to conceal for an extended period. Fraud typically requires co-operation among several employees, or involvement by parties on the other side of a fraudulent transaction. Financial reporting fraud usually leaves a document trail that is subject to internal and external audit. Third, if real performance does not increase and the deception must continue, its effects cumulate in the accounts and the probability of being detected rises.\textsuperscript{22} Fourth, many schemes that overstate current earnings involve accelerating revenue recognition (incorporating revenues in earnings too early), and hence only “borrow” revenue and earnings from future periods. This increases reported performance in the current period, but reduces it in future periods, apparently in the hope or belief that future real performance will increase and will allow the deception to be covered up. Convincing oneself that real performance will improve before the fraud is discovered

\textsuperscript{21} Painter (2002) and Brudney and Ferrell (2002) discuss legal issues. After heavy lobbying by charities, the final version of Sarbanes-Oxley dropped requirements in early drafts that corporations disclose their charity recipients.

\textsuperscript{22} Parmalat reported fake sales for decades, the cumulative effect being that by 2002 it reported €3 billion in fake bank accounts. The alternative – faking uncollected Accounts Receivable -- would more likely have been detected by auditors (verifying individual receivables) or analysts (noticing an unusually high receivables/sales ratio).
would be assisted by an excessively strong belief in oneself and in one’s company.\textsuperscript{23} Thus, committing accounting fraud frequently involves a “gambler’s mentality.”\textsuperscript{24} This is an important point in assessing whether the increased penalties under the Sarbanes-Oxley Act are likely to deter future financial reporting fraud, discussed below.

3. Political/regulatory Reaction to the Scandals

The political and regulatory response to the scandals had two branches: criminal proceedings against companies and individuals accused of acting negligently or fraudulently in the past, and legislative proceedings aimed at reducing the incidence of future negligence and fraud.

3.1 Criminal Proceedings

Criminal law is created, administered, and prosecuted by governments. In criminal law, defendants found guilty in securities cases can be punished with prison sentences, fines, asset seizures, or disbarment from practice. Additional costs borne by defendants include legal fees, loss of reputation, loss of employment prospects, time, and stress on themselves and their families.

Criminal proceedings were initiated by the United States against a range of individuals and corporations involved in the accounting scandals.

The audit firm Arthur Andersen was indicted and convicted with astonishing speed. The SEC opened its investigation of Andersen on 29 November 2001. On 9 January 2002, the Department of Justice upgraded it to a criminal investigation. Andersen first attracted widespread public attention the following day, 10 January 2002, when it announced it had destroyed certain Enron audit documents the previous year, some after learning of the SEC investigation. On 6 May 2002, in the United States District Court for the Southern District of Texas, the United States charged Andersen with obstruction of justice by failing to retain audit documents. After a widely-

\textsuperscript{23} Kenneth Lay and Jeffrey Skilling (Enron), Richard Scrushy (HealthSouth) and Bernard Ebbers (WorldCom) claimed in their trial defense that their companies were fundamentally sound, despite compelling evidence to the contrary. The link between financial fraud and excessive self belief is explored by Schrand and Zechman (2008).

\textsuperscript{24} The diagnostic criteria for pathological gambling of the American Psychiatric Association (2000) include: “after losing money gambling, often returns another day to get even (“chasing losses”).”
covered trial, Andersen was convicted by a jury on 15 June, barely five months after the allegation of destroying documents initially surfaced. Andersen then lost an appeal against the conviction, in the United States Court of Appeals for the Fifth Circuit, and because federal regulations disbar convicted felons from auditing public companies, it surrendered its license to practice on 31 August. Three years later, on 31 May 2005, the conviction was unanimously overturned by the U.S. Supreme Court – a Pyrrhic victory for Andersen, which had long since closed its doors.25

I hold no brief for Arthur Andersen. In fact, I argue below that, in the absence of government involvement, the market most likely would have closed Andersen by itself, because its reputation for independent professional auditing was in tatters due to a series of poor audits, and its financial viability was threatened by pending damage claims. It also is possible that Andersen would have been convicted in the criminal case even if the District Court had shown greater objectivity. But the fact that Andersen’s conviction was unanimously overturned by the U.S. Supreme Court, in an opinion written in such strong language by the Chief Justice, does raise substantial concerns about Federal prosecution in high-profile cases. The District Court judge in Andersen’s Houston trial was placed under substantial pressure by a press baying for conviction, by politicians denouncing Andersen, by a local Houston public deeply affected by the Enron bankruptcy, and by aggressive Federal prosecutors. Prosecutors enjoy considerable power in such cases, because a high-profile prosecution played out in the world press – however unwarranted it may be – is a crippling penalty per se for the company and its managers.

Compared with the Andersen case, the criminal prosecution of individuals involved in the scandals has taken longer and been more successful. Acting under intense political pressure to “fix the problem,” President Bush created the President’s Corporate Fraud Task Force on 9 July 2002,  

---

25 The court ruled that Andersen’s original conviction was marred by the judge instructing the jury to the effect that criminal intent need not be proven (Andersen claimed it had followed normal policy of retaining important audit documents but discarding volumes of unimportant documents, and was unaware it was taking any illegal action). Chief Justice Rehnquist wrote the opinion, which included: "Indeed, it is striking how little culpability [on the part of Andersen] the instructions [to the jury] required." [http://www.oyez.org/cases/2000-2009/2004/2004_04_368/](http://www.oyez.org/cases/2000-2009/2004/2004_04_368/).
“to restore public and investor confidence in America’s corporations following a wave of major corporate scandals.” Department of Justice statistics for prosecutions between July 2002 and June 2007 record 1,236 convictions attributed to the Task Force, including 214 Chief Executive Officers and Presidents, 129 Vice Presidents, 53 Chief Financial Officers, and 23 Corporate Counsels. Penalties included over $2 billion in restitution and fines. These figures include convictions for securities fraud, insider trading, market manipulation, wire fraud, obstruction of justice, false statements, money laundering, Foreign Corrupt Practices Act violations, stock option backdating and conspiracy, and are not limited to financial reporting, but they give an indication of the aggressive and effective criminal prosecution by the U.S. Government in response to the scandals.

3.2 Legislative Proceedings: The Sarbanes-Oxley Act 2002

The legislative response to the accounting scandals was straight out of the Stigler (1964, 1971) - Peltzman (1976) playbook. Applying Peltzman’s (1976) theory, Watts and Zimmerman (1986, pp. 229-231) argue that the political process has an incentive to avoid perceived responsibility for investor losses, and that legislative action is a political attempt to escape blame. Spurred on by the White House and the press, by declining US prestige, by a declining dollar, by declining share prices, by the events of September 11, and by an economic downturn, Congress rushed to pass the Sarbanes-Oxley Act. The House vote was 423-3 and the Senate vote was 99-0. President George W. Bush immediately signed it into law, on 30 July 2002.

The extensive provisions of Sarbanes-Oxley include:

- CEOs and CFOs must certify their financial statements and internal controls, with substantial criminal penalties (notably, 20-year prison terms) for reckless certification;
- An independent audit committee of the board must be responsible for hiring and overseeing auditors;

26 Figures are from United States Government (2007).
• Fraudulently influencing or misleading auditors is prohibited;

• Non-audit work by the company’s auditor that would compromise their independence is prohibited (such as bookkeeping, internal audit, and legal work);

• Rotation of audit engagement partners is mandated;

• Increased disclosure of a variety of matters is required, including companies’ internal controls, codes of ethics, and off-balance-sheet transactions;

• Security analysts are prohibited from publishing reports on companies where conflicts of interest exist (notably, due to investment banking work);

• Corporate loans to officers are prohibited; and

• The Public Company Accounting Oversight Board (PCAOB) is created to oversee the audit industry.

The PCAOB is a regulator with daunting powers. It is empowered to promulgate rules that affect almost all aspects of the audit industry, including auditing standards and procedures, audit firm practices, and client firms’ internal controls. It monitors compliance with its own rules. Further, it enforces its own rules: it is empowered to take disciplinary action against audit firms and their employees. In many ways, the PCAOB is legislator, policeman, judge and jury. It is given wide-ranging discretion and is subject only to supervision by the SEC. In the long run, it has the capacity to regulate the audit industry with an iron fist.

Overall, Sarbanes-Oxley provides the most extensive regulation of the securities markets since the Securities Act of 1933 and the Securities Exchange Act of 1934, which among other things created the Securities and Exchange Commission (SEC). That legislation also was a response to a political crisis: the Great Crash of 1929 and the Great Depression. Sarbanes-Oxley substantially expands government and regulatory involvement in the securities markets, and financial reporting in particular. It represents a major move toward a Continental European
securities market model and a substantial abandonment of a model that served the United States well for a century. Yet the Act was drafted, debated, voted upon and signed into law within seven months of the SEC opening its investigation into Arthur Andersen. And this by a Congress that takes decades to address Social Security reform.

4. Market Reaction to the Scandals

We do not have the opportunity to observe a world in which either market or political/regulatory processes operate independently, so it is only possible to conjecture what the response of a completely unregulated marketplace would have been. Even if we could observe market forces separately, Hayek (1945, 1988) continually reminds us that they are complex, dispersed, difficult to fully identify, and easy to underestimate.

Nevertheless, it is possible to identify some market mechanisms that operated, and to attempt an assessment of their effectiveness. In parallel to the political and regulatory response to the scandals, the market response had two branches: penalties assessed against those accused of past negligence or fraud, and adaptive changes in institutional arrangements aimed at reducing its future incidence.

4.1 Civil Proceedings and Damages Awards

While the criminal prosecution cases attracted most worldwide public attention, a blizzard of private litigation was launched against firms, managers, board members, audit firms, insurance companies, and any parties alleged to have been complicit in financial reporting malpractices. Civil litigation is prosecuted by private litigants who allege they were harmed by the actions of others. Litigants included stockholders, creditors, bondholders, employees, labor unions and pension plans. Defendants found guilty in securities cases are punished by the courts awarding monetary compensation to the litigants for the damages they have incurred as a consequence of the harmful actions. Unlike criminal proceedings, civil litigation is a private, market process of enforcing
explicit and implicit contracts between firms, managers, auditors, creditors, shareholders and other contracting parties.27

The press attention given to the criminal trials makes it easy to overlook the scale of the market-based penalties that have been meted out. Total damages settlements in securities litigation over 2000-07 exceeded $50 billion (Simmons and Ryan, 2008). In the Enron case alone, extensive private litigation against banks alleged to have been complicit in just some aspects of Enron’s financial misreporting has been settled for a total of approximately $9 billion (for comparison, Enron’s total assets peaked at approximately $70 billion). WorldCom settlements exceeded $6 billion, and Cendant settlements exceeded $3 billion. Some litigation arising from the 2001-02 scandals is still ongoing.

4.2 Reputation, Bonding and Insurance Effects: Who Killed Arthur Andersen?

Reputation effects have long been viewed as a powerful market mechanism, imposing penalties on parties found to have acted inappropriately.28 Karpoff and Lott (1993) document substantial reputational costs to firms committing fraud generally. In the audit industry, the audit firm’s reputation for independent, professional work is central to performing its economic role of verifying financial statements for use by uninformed outsiders.29 DeAngelo (1981) argues that large audit firms, like Arthur Andersen before its demise, earn substantial quasi-rents which they stand to lose if they perform poor-quality work. Research has shown that audit firm reputation is associated with audit fee premiums (Simunic, 1980; Francis, 1984; Francis and Simon, 1987; Palmrose, 1986; Craswell, Francis and Taylor, 1995) and with the market valuation of their clients (Kellogg, 1984; Beatty, 1989). Palmrose (1986, 1987) links litigation risk to audit quality. Events that could reduce auditors’ reputations, such as regulatory action or private litigation against them,

27 While it is a fundamentally private (market) process, civil litigation is subject to regulation (notably, via the Private Securities Litigation Reform Act of 1995) and is affected by the existence of securities laws (notably, private lawsuits citing the violation of Section 10(b) and Rule 10b-5 of the Securities Exchange act of 1934).
28 For example, Telser (1980) and Kreps and Wilson (1982).
are associated with stock price reductions for clients (Loebbecke et al., 1989; Firth, 1990; Moreland, 1995; Franz et al., 1998), with loss of clients (Firth, 1990; Wilson and Grimlund, 1990), and with reductions in audit fees (Davis and Simon, 1992).

By the time it ceased business as an auditor, Andersen’s reputation for quality, independent auditing was in tatters due to a series of deficient audits. These included its audits of Enron, Sunbeam, Waste Management and WorldCom, which garnered international attention. But there were lower-profile transgressions, including its audit of Baptist Foundation of Arizona, which became the largest bankruptcy of a religious nonprofit in U.S. history. Further, there was evidence (discussed below) that Andersen’s responsibility for deficient audits was not due to isolated rogue engagement partners, but was systemic and reached the top of the company. Based on prior research on the economic role of auditor reputation, one could expect that by the time Andersen’s clients digested the implications of its deficient audits and had been able to negotiate a replacement auditor, reputation effects alone would have caused it to shed much of its client base.

A related market mechanism for enforcing discipline on audit firms is bonding. Audit firm partners are jointly and severally liable for partnership debts, and hence in principle the entire pool of capital in the audit firm is available to satisfy damages arising from the actions of individual partners. 30 This acts as a bonding mechanism, giving the audit firm incentives to create high standards of behavior for individual partners, monitor their performance, and enforce compliance. Andersen’s capacity to bond itself to provide high audit quality was impaired by another market mechanism, civil litigation. As discussed in the previous subsection, civil litigation is a market process for recovering damages from parties who do not follow accepted standards of behavior. Investors and lenders are owed a duty of care by audit firms which, if not performed, can trigger lawsuits to recover damages. Because litigation is on behalf of damaged parties, it usually occurs in

30 The Private Securities Litigation Reform Act of 1995 restricted audit firm joint and several liability. In addition, audit firms are organized with limited liability, so individual partners’ liabilities are limited to their invested capital.
relation to client firms whose stock prices have fallen substantially (Lys and Watts, 1994) and whose capacity to pay therefore is limited. Audit firms then are named as defendants.

Andersen was deluged with lawsuits. Its settlements included: Sunbeam for $110 million, Baptist Foundation of Arizona for $217 million, Waste Management for $27 million including fines, WorldCom for $65 million, and Enron for $40 million. These amounts do not include legal expenses. Andersen currently is named as a defendant in more than one hundred suits still pending in the courts. Settlements and legal expenses depleted Andersen’s resources, to the point where its ability to credibly bond itself to investors and lenders against future financial reporting negligence and fraud would have been severely impaired.

A market mechanism closely related to bonding is a form of insurance provided by auditors. Kellogg (1984), Wallace (1987), Chow, Kramer and Wallace (1988), and Kothari, Lys, Smith, and Watts (1988) develop the hypothesis that, because audit firms are named as co-defendants in corporate financial reporting cases, their resources provide insurance to investors and lenders against financial reporting negligence and fraud. Whether they self-insure against this potential liability or can purchase insurance, a competitive audit market should allow audit firms to pass this insurance cost on to their client firms as a component of their audit fees (Dye, 1993). The depletion of Andersen’s resources from litigating and settling past cases would have impaired its ability to insure investors and lenders against future financial reporting negligence and fraud.

Given its loss of both reputational and monetary capital, it is not surprising that Andersen started shedding clients. It lost a net 21 clients in 2000 and 73 clients in the first three quarters of 2001, while during the same intervals the other Big 4 audit firms had net gains of 44 and 14 clients. These defections preceded the Enron and WorldCom scandals, suggesting that the losses were due primarily to the Waste Management case. Barton (2005) studies Andersen’s subsequent

---

32 Numbers supplied in private communication by Clive Lennox.
post-Enron client losses. He reports the earliest defections were by clients that were more visible in the capital markets, in the sense that they attracted more analysts and more press coverage, had larger institutional ownership, had larger turnover, and issued more securities. By the time Andersen surrendered its license, it had lost its client base.33 Barton (2005, p. 549) concludes: “Overall, my study suggests that firms more visible in the capital markets tend to be more concerned about engaging highly reputable auditors, consistent with such firms trying to build and preserve their own reputations for credible financial reporting.”

There is some debate about the relative contributions of reputation, insurance and bonding effects in the demise of audit firms. Menon and Williams (1994) study the fraud-induced 1990 bankruptcy of Laventhol and Horwath, then the seventh-largest U.S. audit firm. They conclude the insurance hypothesis explains the negative stock price reaction of its clients to its bankruptcy. However, Baber, Kumar, and Verghese (1995) conclude the Laventhol and Horwath clients’ price movements are consistent with both insurance and reputation effects. Sinason and Pacini (2000) and Khurana and Raman (2004) reach a similar conclusion from wider samples. Chaney and Philipich (2002) attribute the stock price behavior of Arthur Andersen clients around the time of the Enron scandal to reputation effects, but Nelson, Price and Rountree (2008) document confounding effects. Weber, Willenborg and Zhang (2008) study the interesting German accounting scandal involving ComROAD AG, which is informative because damages awards are both rare and minor in Germany, so bonding and insurance effects are not substantial. They report that clients of KPMG, ComROAD’s audit firm, experienced negative abnormal returns and tended to change auditors. They conclude that reputation effects are important.

33 As Barton (2005, p.566) observes, we know only the dates when a change in auditor was announced, which would have substantially lagged the decision to change. Most clients had December 31 fiscal year-ends, and the earliest they could have terminated their Andersen engagements would have been after the Annual Report and 10-K were completed and filed, several months after the year-end. In addition, it takes time for managers to locate alternative auditors (e.g., to explore potential conflicts of interest is a detailed, complex task), and to request and review the audit programs proposed by replacement candidates, and for the audit committee to meet and vote on the replacement.
The most likely explanation is that reputation, insurance and bonding effects cannot be separated, either logically (because the concepts overlap) or empirically (because their effects are correlated). This seems particularly likely in common law countries, where loss of reputation is associated with litigation and consequential loss of monetary capital. Nevertheless, all are market mechanisms.

While it is impossible to completely untangle the effects of market and political/regulatory processes in the demise of Arthur Andersen, there appears to be ample evidence the audit market would have closed Andersen on its own accord, because the firm’s greatest asset (a reputation for quality, independent auditing) and its financial viability (hence its capacity to bond the quality of the accounts it audited and to insure against harm to users) were in ruins. In addition, even if it had won its criminal trial, the flurry of private civil lawsuits ensuing from the Enron collapse surely would have bankrupted it. It seems reasonable to conclude that market forces, left to their own devices, would have closed Andersen.

Finally, it is worth reflecting on what the closure of an audit firm implies, and why an audit firm is not closed whenever a single negligent or fraudulent audit comes to light. It is important to distinguish between the actions of an individual partner, those of a regional office, and those of an audit firm as a whole. Individual partners acting alone in negligence or fraud typically lose future employment prospects, lose personal assets in damages settlements, and can be disbarred, fined or jailed. Some financial and reputational losses also are suffered by those who are not directly implicated. Financially, under joint and several liability the partnership as a whole is responsible for fines and damages awards, which normally exceed individual-partners’ wealths. Firm-wide financial and reputational losses due to the actions of a rogue partner provide an incentive for partners to monitor each other (though this mechanism apparently was not sufficient in Andersen’s case). When a scandal is due to poor oversight or a corrupt audit culture in an entire regional
office, but is clearly contained within that office, the penalties fall more heavily on partners in that office, but here too they typically do not threaten the entire firm.\textsuperscript{34}

So why was Andersen forced to close its doors when this normally does not happen? Presumably because in Andersen’s case there was clear evidence that culpability did not simply lie with individual audit engagement partners, or even with a single regional office, but involved managers at the very top. The firm itself was tainted. In the Waste Management scandal that preceded Enron, the following top-level Andersen people were directly implicated in a blatant attempt to disguise improper accounting, while at the same time Andersen was issuing clean audit opinions: Andersen’s Managing Partner, the Practice Director for its Central Region, its Director of Global Risk Management, and the Audit Division Head of its Chicago head office.\textsuperscript{35} When this firm-level culpability was discovered, Andersen was fined a then-record $7 million, and consented to a permanent injunction against further violations. The Enron and WorldCom cases, which broke soon thereafter, flew in the face of this injunction.\textsuperscript{36} It is reasonable to conclude that Arthur Andersen’s firm-wide capacity to credibly perform independent audits was severely damaged by its

\textsuperscript{34} In the May 2007 ChuoAoyama PricewaterhouseCoopers (PwC) scandal in Japan, the regional office was closed, many auditors were dismissed, but those not associated with the scandal were retained in a newly-created firm under the PwC global network, with new management and 30% fewer clients.

Parmalat S.p.A. is a similar case. In December 2003, it was revealed that an Italian audit firm associated with the U.S. auditor Grant Thornton had accepted a fax -- allegedly from Bank of America, but faked by Parmalat managers to conceal large losses – as confirming the existence of accounts totaling approximately €4 billion held on behalf of a Parmalat subsidiary. Dianthus S.p.A., an Italian audit firm associated with the U.S. firm Deloitte & Touche, audited Parmalat’s accounts, and relied on the Grant Thornton associate’s audit of the subsidiary. Websites and other public statements made it clear that both U.S. audit firms accepted no responsibility for their Italian associates’ work (possibly in recognition of Italy’s reputation for lax enforcement of audit procedures). Nevertheless, there were spillover effects on the U.S. firms. Deloitte & Touche paid $149 million to settle a €13.2 billion lawsuit by Parmalat, Grant Thornton is still defending a similar suit for €10 billion, and Grant Thornton experienced U.S. client losses almost immediately (“Grant Thornton hit by client losses after Parmalat scandal,” \emph{Financial Times} March 18, 2004).

\textsuperscript{35} SEC, Accounting and Auditing Enforcement Release Nos. 1405 and 1410 (June 19, 2001).

\textsuperscript{36} Two Andersen partners banned from auditing public companies for three years were not fired, but reassigned to non-audit jobs. Robert V. Kutsenda, the Director of Global Risk Management, was banned for a year but reassigned to revising Andersen's document retention policy. Ironically, this revised policy received widespread adverse publicity when the Enron scandal broke. Asked why they had not been fired, Andersen's Managing Partner Joseph F. Berardino said: "You need to be fair to people who are trying to do a good job." Floyd Norris, “Enron's Collapse: News Analysis; For Andersen and Enron, The Questions Just Keep Coming,” \emph{The New York Times} January 16, 2002.
firm-level culpability in these events, and that this explains why the firm – and not just individual rogue partners or branch offices – went out of business.\textsuperscript{37}

4.3 Audit Firm Conflicts of Interest

Part of the adverse public and political reaction to the accounting scandals was the revelation that audit firms conduct substantial non-audit work for their clients, thereby creating at least the appearance of conflicts of interest. Conflicts of interest were given at least partial blame for the scandals by the press (e.g., Herrick and Barrionuevo, 2002) and by some researchers (e.g., Coffee, 2002). In response to the adverse reaction, Sarbanes-Oxley prohibits the type of non-audit work by the company’s auditor that would compromise its independence in performing the audit. Examples include the provision of bookkeeping and internal audit services, where as an external auditor it would in effect be auditing its own work. The Act sensibly does not prohibit other services, such as tax advice. There are several reasons to doubt whether this statutory prohibition was advisable or necessary.

First, the hypothesis that audit firms allow their audit judgment to be compromised by non-audit revenues does not make as much sense as one might initially believe. Why would audit firms attach such a low value to their reputations as independent auditors? Why would they willingly place the entire capital of the partnership at risk by cutting audit quality? The hypothesis might seem to make sense if one accepted the premise that they were earning quasi-rents on non-audit business, but none on audits. They then might seem to have little to lose by reducing audit quality to attract lucrative non-audit engagements. But even if one accepted that premise, the argument still would make no sense. Why would they willingly risk losing quasi-rents on their non-audit work by gaining a reputation for poor audit quality? Would not the existence of quasi-rents earned from

\footnote{Eisenberg and Macey (2004) show that the frequency of financial restatements by Andersen clients – a measure of financial reporting quality – was not significantly different from that of clients of other large audit firms, controlling for important variables such as size. This is consistent with the hypothesis that Andersen closed its doors because its culpability in several important cases reached the top of the company, not because of audit failures \textit{per se}.}
non-audit business imply that firms with substantial non-audit revenues put up a larger bond to guarantee their audit quality and hence are *less* likely to compromise it? All things considered, the motives of audit firms are not as clear-cut as many commentators have portrayed them.

Second, if client firms view their reputations are valuable assets, one would expect them to voluntarily avoid contracting for the audit firm to provide any non-audit services that could compromise audit independence. Kinney, Palmrose and Scholz (2004) test this hypothesis, using the need to subsequently restate previously issued financial statements as an indicator of low-quality financial reporting and auditing. They do not find a pervasive relation between non-audit services fees and restatements.

Third, one should not lose sight of the benefits non-audit work brings to clients. The accounting firms have built substantial businesses in areas such as consulting, systems, taxation, and litigation support. They have done so in a competitive market. Their comparative advantage appears to lie in a combination of training, cost and specific client and industry knowledge. Excessive restrictions on using their comparative advantage can impair economic efficiency.

Fourth, there is substantial evidence that non-audit business in fact does not lead to audit firms compromising their audit judgments. If anything, the evidence in these studies points to non-audit revenues being associated with *less* favorable audit treatment, most likely because financially weaker firms exhibit greater use of consultants and also receive harsher audit opinions.

Fifth, before the passage of Sarbanes-Oxley and in response to the scandal-induced perception that conflicts of interest influenced audit judgment, all but one of the major accounting

---


39 Coffee (2006, pp. 147-148 and 322-325) counters that non-audit revenues created industry-wide, not firm-specific, incentives and hence audit firms catered to all clients, independently of client-specific non-audit fees. While this argument has some validity, it does not seem plausible that industry-wide catering would co-exist with audit firms catering equally to clients, independent of the amount of their non-audit revenues.
firms decided to cease providing internal audit and audit-related technology consulting services to clients where they are the external auditor. This decision was made by Arthur Andersen (it was still operating at the time), Ernst & Young, KPMG, and PricewaterhouseCoopers. Deloitte & Touche alone among the then Big Five audit firms resisted the change, arguing with some justification that the issue was one of perception arising from “the level of hyperbole in the debate.” Nevertheless, perceptions do matter, and it is not clear that Deloitte would have been able to hold out on this position for long. In any event, Sarbanes-Oxley codified the standard that the market had largely moved to in response to the scandals.

Sarbanes-Oxley also restricts public companies from hiring directly from their audit firms. More specifically, Section 206 prohibits a company from employing a former auditor of its accounts in a senior accounting or financial position until at least a year has elapsed since that person left the audit firm. The principal concern with “revolving door” appointments is that they induce various conflicts of interest, including those that could occur when audit firm employees have to audit a former colleague or supervisor. This Sarbanes-Oxley prohibition was fueled by public outcry upon news that former Andersen auditors had been employed by Enron, Waste Management and other companies with deficient financial reporting. But here too, the need for such prohibition is far from clear. Lost in the argument is the potential benefit a former auditor can bring to a company, stemming from specialist skills and specific knowledge of the company. Consistent with such benefits, Geiger, Lennox and North (2008) report a positive share market reaction to pre-Sarbanes-Oxley revolving door appointments, in excess of the price reaction to other appointments. They also report that revolving door appointments are not associated with lower scores on measures of financial reporting quality.

In sum, there is reason to believe, and some supporting evidence, that market forces can resolve potential auditor conflicts, in the absence of regulatory prohibitions. This does not imply that there should be no regulation, but it does help place in perspective the Sarbanes-Oxley rush to heavily regulate the industry.

4.4 Who Killed Enron and Enron Jobs?

Was Enron forced into bankruptcy – and did employees lose their jobs – due to the accounting scandals? Or due to bad business decisions? Here too, the political and market perspectives seem to differ.

The notion that accounting transgressions led to the Enron bankruptcy, and the associated job losses, is correct in a sequential sense: revelation of the company’s financial misrepresentations was swiftly followed by its demise. That does not mean that accounting transgressions caused the Enron bankruptcy or the job losses, because revelation of the financial misrepresentations was accompanied by revelation of the motive behind them, which was to conceal the company’s true financial position. I will argue that, if anything, the accounting transgressions actually deferred bankruptcy and job losses for a short period.

The immediate press coverage of Enron dwelt extensively on job losses.41 This focus returned at the sentencing of Skilling and Lay. For example, BBC News stated on 23 October 2006:42 “The scandal at the one-time energy giant left 21,000 people out of work.” When signing the Sarbanes-Oxley Act into law on July 30, 2002, President Bush stated:43 “This law says to workers: we will not tolerate reckless practices that artificially drive up stock prices and eventually destroy the companies, and the pensions, and your jobs.”

---

41 In a retrospective review of media coverage, Myatt (2004) concluded “the TV networks and mainstream dailies focused on business, the job losses and human interest.”
At the sentencing hearing for Skilling and Lay, U.S. District Judge Sim Lake admitted testimony from several former Enron employees, including Charles Prestwood (who testified he and the other former employees were the real victims in this case) and Diana Peters (who testified her husband was diagnosed with cancer shortly before the company went bankrupt, leaving her without health insurance, and to pay for his treatment she had taken different jobs and sold her furniture, home and assets). The prosecution team attributed Enron job losses to the financial misreporting under Skilling and Lay. After their sentences were handed down, Alice Fisher, an assistant U.S. attorney general, stated:44 “Today's sentence is a measure of justice for the thousands of people who lost their jobs and millions of dollars in investments when Enron collapsed under the weight of the fraud perpetrated by the company's top executives.”

It is not a crime per se to have bad business judgment, or make bad business decisions, even if they lead to bankruptcy or unemployment. However, deliberate financial misreporting is a crime, and as Judge Sim Lake indicated the extent of job losses is a relevant consideration in sentencing under federal guidelines. But one would have thought job losses would be relevant only if the misreporting caused them. I also hold no brief for Skilling or Lay, and the following observation has no implications for their guilt or otherwise, which is a separate issue from sentencing. But I am unaware of any reliable evidence that the accounting scandals caused (as distinct from deferred) job losses. The most plausible hypothesis is provided by Kedia and Philippon (2007), who conclude that overstating earnings initially causes over-investment and over-employment, distorting the allocation of real resources. When the overstatement is discovered, the firm sheds labor and capital.45

45 Sadka (2007) similarly analyzes the product market effects of financial misrepresentation.
Some indication can be gleaned from Enron’s financials. Between 1995 and 2000, reported revenues grew from $9.2 billion to $100.8 billion: a tenfold increase in five years. However, reported profits did not even double over the period, and consequently the net profit margin on sales fell by 2000 (using reported numbers) to less than 1%. We now know that even that slim margin actually was not earned, because reported profits that year were overstated. Based on the Bankruptcy Examiner’s estimate, Enron’s true 2000 profit was $42.3 million. The restated numbers imply a paltry 0.04% margin on sales (and a mere 0.06% of restated total assets of $69.6 billion). This estimate might be biased downward because it was made by an interested litigant, but it suggests that Enron was an unprofitable company the year before it entered bankruptcy, and that its accounting transgressions were designed to hide that fact.

It is difficult to escape the conclusion that market forces caused Enron’s bankruptcy, for the simple reason that by 2000 it was not generating profits on an enormous amount of invested capital. Conditional on the company having mis-allocated capital and labor resources to unprofitable projects such as its energy trading and broadband businesses, its accounting transgressions most likely kept the company alive for some period (perhaps one or two years) longer than would have occurred if it had reported the true profitability. Absent misrepresentation, Enron would have been at least restructured and possibly liquidated, and many Enron employees would have lost their jobs, at an earlier point in time. If Enron managers anticipated the possibility of financial misrepresentation before committing to these businesses, they might even have been encouraged to over-invest and over-employ in them in the first instance. Whatever counter-factual is assumed, Enron’s levels of investment and employment would have peaked earlier in the

---

46 United States Bankruptcy Court, Southern District of New York (2003). Enron reported Operating Cash Flow of +$3.0 billion, which the Bankruptcy examiner claimed should have been -$0.2 billion.
absence of financial misrepresentation. The welfare loss arose from employing excess capital and labor that were better used elsewhere.\footnote{Tragically, employees also lost approximately $1 billion on their 401(k) portfolios, which on average were 63 percent invested in Enron stock. A welfare loss would arise from any over-investment during the period in which the company’s poor financial condition was not revealed, but the entire loss is difficult to attribute to the scandal because it primarily postponed the date of reckoning and hence the loss. Employees generally overinvest their 401(k) plans in their employer's stock, so it is difficult to attribute poor diversification entirely to the misrepresentation. The aggregate welfare loss from accounting scandals due to undiversified 401(k) plans is unclear.}

\subsection*{4.5 Summary: Market Consequences of Misreporting}

As is the case with the political/regulatory system, markets have their own mechanisms for penalizing and hence deterring harmful behavior that does not meet accepted standards. While the political/regulatory response to the scandals of 2001-02 garnered considerably more press attention, the market response was substantial. It included enormous damages awards under civil litigation, reputation effects, and bankruptcy. Furthermore, markets are adaptive mechanisms, and learn from past events, even though many of the likely market changes in response to the scandals were pre-empted by legislation. I return to this issue in Section 6 below.

\section*{5. GAAP versus “accounting principles that are generally accepted”}

Codified standards arise naturally in a variety of market settings, as an efficient way of satisfying demand. High quality individual producers, acting atomistically, have incentives to incur costs of signaling to uninformed buyers the standards met by their products (Spence, 1973). Minimum education requirements are one such signaling mechanism. In many markets there also are efficiency gains from developing, codifying and enforcing standards across multiple producers. These include networking gains (e.g., Besen and Farrell, 1994) and gains from economies of scale (designing standards once centrally, and applying them multiple times locally).

Professional standards typically exist whenever there is a profession, absent regulation. To be accepted and compensated by the marketplace as a professional, one must meet accepted market standards of practice. Standards might simply take the form of generally accepted best practice.
Alternatively, a professional organization with voluntary membership might formally codify a body of rules that it enforces through its own disciplinary procedures. There might be competition among such organizations, with different standards.

The question then arises as to why countries generally have gravitated to a single professional accounting body (in the U.S., the AICPA) and a single set of codified standards nationally (U.S. GAAP) and perhaps even worldwide (IFRS). Is this an efficient solution in the financial reporting market, or is it due to regulation? A closely related question is the nature of the interplay between market and political forces in determining reporting standards. What is the ultimate criterion against which financial reporting in the U.S. is evaluated? Generally Accepted Accounting Principles (GAAP), defined as the codified set of rules endorsed by FASB? Or “generally accepted accounting principles,” defined in the common-law sense as the accounting principles that are generally accepted by market participants as applying in the United States, a criterion that encompasses all factors that determine appropriate financial accounting?

These are important questions. The relative merits of “principles-based” and “rules-based” accounting have been argued extensively in the international arena, due largely to the belief that these different approaches underlie International Financial reporting Standards (IFRS) and U.S. GAAP, respectively. The SEC in particular has been wary of the lack of specificity in some IFRS standards. The debate took on fresh impetus as a consequence of the recent accounting scandals, which were attributed in part to companies following the letter, but not the spirit, of the rules. A “rule-checking” mentality among accountants and managers is widely believed to have contributed to the scandals that emerged in 2001-02.

Codified GAAP certainly gives the impression that U.S. financial reporting is rules-based. Accounting for leases is the paradigmatic example when examined from the viewpoint of codified

---

48 The distinction between rules-based and principles-based accounting is not totally clear because, as Schipper (2003) points out, rules generally are based on principles.
rules alone. Shortridge and Myring (2004) note that there are only seven relevant IFRS documents affecting leases, but the relevant U.S. GAAP is contained in at total of 78 FASB Statements, Interpretations, Bulletins, and EITF Abstracts. Further, the primary IFRS lease standard (IAS 17) simply defines a transaction as a capital lease “if it transfers substantially all of the risks and rewards incident to ownership” to the lessee, making it difficult for firms to write contracts that barely avoid minimum requirements and hence keep them off their balance sheets. The equivalent U.S. standard (SFAS 13) specifies four precise numerical criteria for capital leases, so companies routinely structure their lease contracts to avoid triggering the criteria by the barest of margins, thereby facilitating substantial off–balance-sheet financing. The more “principles-based” approach of IFRS therefore leads to balance sheets that better reflect the economic substance of the lease transaction. Considering only specific GAAP standards such as those applying to leasing, U.S. accounting appears unequivocally rules-based, and not principles-based.

In this section, I propose that the standard against which financial reporting ultimately is evaluated in the U.S. has substantial elements of the common law version (i.e., the standard of reporting that is generally accepted as applying in the United States). A corollary is that the United States de facto operates a “principles based” accounting system, contrary to popular belief. Further, I propose that regulation is largely responsible for fostering the incorrect view that codified GAAP alone determines acceptable financial reporting standards.

5.1 Substance Takes Precedence over Form

FASB’s Statements of Financial Accounting Concepts (SFACs) are “intended to serve the public interest by setting the objectives, qualitative characteristics, and other concepts that guide selection of economic events to be recognized and measured for financial reporting and their display in financial statements or related means of communicating information.”

---

states “The quality of reliability, and, in particular, representational faithfulness leaves no room for accounting representations that subordinate substance to form.”\(^{50}\) This statement reaffirms a long-standing and fundamental principle. Earlier, APB 4 had stated “Financial accounting emphasizes the economic substance of events even though the legal form may differ from the economic substance and suggest different treatment.”\(^{51}\)

The standard unqualified United States audit report states that the firm’s financial statements “present fairly, in all material respects, the financial position of the company at year end, the results of operations, and its cash flows, in conformity with generally accepted accounting principles.” A similar requirement has been in effect for three quarters of a century, its origins being in the first codification of generally accepted standards by the then American Institute of Accountants. The precise interaction between the “present fairly, in all material respects” and “in conformity with generally accepted accounting principles” components of the audit certificate is debatable, but they clearly are not intended to be equivalent statements.

SAS 69 clarifies the interaction between the two components, by explicitly requiring auditors to judge whether the company’s selection and use of formal GAAP pronouncements is such that transactions and events are treated in accordance with their substance. “The auditor should consider whether the substance of transactions or events differs materially from their form.”\(^{52}\) The same section also cautions that it is possible to follow the letter of a particular GAAP pronouncement but still fail to meet the objective of providing relevant and reliable information about the true economic position of the firm.

---

\(^{50}\) Statement of Financial Accounting Concept No. 2 ("SFAC 2") “Qualitative Characteristics of Accounting Information,” May 1980, paragraph 160.


SAS 90 further clarifies the interaction between “present fairly” and “in conformity with generally accepted accounting principles” by requiring public company auditors to discuss with the company’s audit committee their judgment of the quality of the accounting employed by the company, not simply its compliance with GAAP.53 The discussion should address matters such as the completeness and clarity of the financial statements, the neutrality, representational faithfulness, and verifiability of the information they contain, new accounting policies, changes in policies, estimates, judgments, uncertainties, and unusual transactions. The standard notes that these assessments require exercising judgment in the particular circumstances that cannot be subsumed by objective criteria.

Rule 203 of the AICPA Code of Conduct makes it clear that a mechanical application of formal GAAP pronouncements can lead to inappropriately misleading financial reporting:

A member shall not (1) express an opinion or state affirmatively that the financial statements or other financial data of any entity are presented in conformity with generally accepted accounting principles or (2) state that he or she is not aware of any material modification that should be made to such statement or data in order for them to be in conformity with generally accepted accounting principles, if such statements or data contain any departure from an accounting principle promulgated by bodies designed by Council to establish such principles that have a material effect on the statements or data taken as a whole. If, however, the statements or data would otherwise have been misleading, the member can comply with the rule by describing the departure, its approximate effects, if practicable, and the reasons why compliance with the principle would result in a misleading statement.

Reconciling the qualitative requirements of Rule 203 with formal GAAP pronouncements requires the exercise of professional judgment. That is, it requires principles-based rather than rules-based accounting. The Rule seems to have been widely ignored by accountants.

5.2 Simon Says: U.S. Accounting is Principles-based

In the 1969 criminal case US v. Simon, Judge Friendly held that an accountant could not escape conviction for fraud by showing that financial statements he knew were misleading had

Nevertheless been prepared in accordance with rules-based GAAP. This important Federal Court decision played a central role in the WorldCom and Adelphia trials. It adopted a principles-based view of accounting.

The auditors of Continental Vending Machine Corporation were charged with violating U.S. securities laws by certifying financial statements they knew to be false. An affiliate of Continental Vending, Valley Commercial Corporation, borrowed money from Continental and loaned it to a Continental manager who the auditors knew would be unable to repay the debt. The audited financial statements of Continental showed the Valley receivable as an asset, and an ambiguous and obscure footnote provided some details. In the trial, the defense argued that the footnote disclosure explaining the receivable from Valley complied with the letter of GAAP and that the auditors had followed Generally Accepted Auditing Standards (GAAS). The District Court judge instructed the jury that mere compliance with professional accounting standards was not a sufficient defense, and the appropriate test was whether the financial statements fairly represented Continental’s financial position. The jury found the defendants guilty, and in the appellate court Judge Friendly ruled that the District Court judge did not err in his instructions to the jury.

The generality of the Simon case (also referred to as Continental Vending) is unclear. The decision was handed down in the Second Circuit, and is not binding in other jurisdictions. Nevertheless, the decision played a central role in the recent WorldCom and Rigas (Adelphia) scandals (see below), was not overturned at the U.S. Supreme Court, and has been cited and reaffirmed in several important cases, many of them outside the Second Circuit. Perhaps the

---

55 The United States Court of Appeals for the Second Circuit is one of thirteen regional Courts of Appeals in the Federal system. Its territory comprises Connecticut, New York, and Vermont. Judge Henry Friendly was one of its most famous jurists, known for having written many important and enduring opinions.
uncertain generality of Simon helps explain why the case has attracted little commentary by the accounting profession, including auditors, standard-setters, textbook authors, and academics. Isbell (1970), Sterling (1973) and Zeff (2003b) are rare exceptions. After the Hochfelder case, Causey (1976) and Chapman, Zitek and Yellin (1976) also drew some attention to it. Not even the central role of Simon in the WorldCom and Adelphia cases seems to have drawn commentators’ attention away from rules-base GAAP.

The ruling in Simon is not contradicted by the courts placing substantial weight upon rules-based GAAP. Under Simon, compliance with rules-based GAAP is persuasive evidence the financial statements fairly represent the company’s financial position, but it is not sufficient. Further, willful failure to comply with GAAP normally is associated with an intent to mislead, and will fail the test in Simon. The implied assumption in the Simon ruling is that GAAP rules normally but do not always reflect principles (cf. Schipper, 2003).

Simon gives rules-based GAAP a role that is analogous to generally accepted precedent in common law, and the audit profession a role that is analogous to the common law judiciary. Common law is created by court or tribunal decisions resolving disagreements among private parties. It thus originates in markets, not in legislation or other state action. If a case arises with important circumstances that are covered by no authoritative precedent, common law judges must make a decision that, if it survives appeal and especially if it is cited by authoritative courts, establishes a precedent for resolving similar disagreements in future. The decisions of all courts are not equally authoritative, and are not necessarily binding in different jurisdictions. If a disagreement involving materially different circumstances needs resolution, it must be decided by imagining how it would have been resolved in common practice if the circumstances had been anticipated and provided for.
Setting aside for the moment the political/regulatory influence on GAAP, from a market perspective the SAS 69 hierarchy seems analogous to the structure of common law: FASB as a specialist accounting court, EITF as a lower court, the various AICPA Industry Audit and Accounting Guides, Practice Bulletins and Accounting Interpretations as administrative interpretations, and auditors as administrative judges.57

Like common law precedent, rules-based GAAP under Simon is persuasive, but will not alone determine whether the financial statements fairly represent the company’s financial condition in all circumstances, and audit judgment therefore needs to be applied. This suggests that accounting textbooks and instructors might balance their coverage of GAAP pronouncements with coverage of judging whether the financial statements fairly represent the company’s financial position. It also suggests that researchers might focus less on GAAP pronouncements, for example when studying international differences or international convergence on uniform rules such as IFRS, and more on what companies actually report.58

5.3 Principles-based Accounting and the Enron Case

Perhaps learning from their mistake in the HealthSouth case, the Department of Justice did not charge Enron executives Skilling and Lay with violating rules-based GAAP. They were charged with several counts of misleading investors and employees, by misrepresenting the company's earnings and financial position and by withholding materially bad news. On one count, in its closing arguments the prosecution stated: "Abracadabra, just like that. A penny to meet the consensus estimates. That's fraud. It's wrong." Another count did not contend accounting rules were violated when Enron shifted a business unit to a different segment for reporting purposes, to hide a loss, but simply argued investors were mislead. The prosecution argued: "They didn't

57 The economic role of specialized private (i.e., market) judiciaries is described in Milgrom, North and Weingast (1990). The argument that they already exist in the audit industry makes calls for an “accounting court” (Spacek, 1958; Stamp, 1970) appear misplaced.
58 See Ball, Robin and Wu (2003) and Ball (2006).
disclose these problems to the marketplace, and the investor didn't know and didn't understand what was going on. … It is a crime to omit material information about your business . . . .

Another charge involved lying about Enron’s performance on a non-GAAP financial metric called "total contract value." No violation of GAAP was alleged; the prosecution simply argued that deliberately misleading investors is a crime.

The extent to which Enron's financial statements did or did not conform to rules-based GAAP is subject to debate, and has not been resolved in court to my knowledge. Nevertheless, the Lay and Skilling convictions were for knowingly failing to present a fair picture of Enron’s profitability and financial condition, and the issue of whether Enron followed GAAP was moot.

5.4 Principles-based Accounting and the WorldCom Case

In appealing his conviction for accounting fraud, WorldCom ex-CEO Bernie Ebbers argued the prosecution had not alleged, let alone proven, that WorldCom’s accounting violated rules-based GAAP, and hence his conviction was flawed. In essence, he claimed the sole standard for U.S. financial reporting is GAAP. The Court of Appeals for the Second Circuit dismissed this argument in no uncertain terms, with words that bear repeating at length:59

“Ebbers argues that the indictment was flawed because it did not allege that the underlying accounting was improper under GAAP, and that the district court should have required the government to prove violations of GAAP at trial. He claims that where a fraud charge is based on improper accounting, the impropriety must involve a violation of GAAP, because financial statements that comply with GAAP necessarily meet SEC disclosure requirements. . . . We [The Court of Appeals for the Second Circuit] addressed a similar argument in United States v. Simon. . . . We see no reason to depart from Simon.

To be sure, GAAP may have relevance in that a defendant’s good faith attempt to comply with GAAP or reliance upon an accountant’s advice regarding GAAP may negate the government’s claim of an intent to deceive. Good faith compliance with GAAP will permit professionals who study the firm and understand GAAP to accurately assess the financial condition of the company. This can be the case even when the question of whether a particular accounting practice complies with GAAP may be subject to reasonable differences of opinion.

However, even where improper accounting is alleged, the statute requires proof only of intentionally misleading statements that are material, i.e., designed to affect the price of a security. If the government proves that a defendant was responsible for financial reports that intentionally and materially mislead investors, the statute is satisfied. The government is not required in addition to prevail in a battle of expert witnesses over the application of individual GAAP rules.

... 

In a real sense, by alleging and proving that the financial statements were misleading, the government did, in fact, allege and prove violations of GAAP according to the AICPA’s Codification of Statements on Accounting Standards, AU §312.04, "[f]inancial statements are materially misstated when they contain misstatements whose effect, individually or in the aggregate, is important enough to cause them not to be presented fairly, in all material respects, in compliance with GAAP." Thus, GAAP itself recognizes that technical compliance with particular GAAP rules may lead to misleading financial statements and imposes an overall requirement that the statements as a whole accurately reflect the financial status of the company.

To be sure - and to repeat - differences of opinion as to GAAP’s requirements may be relevant to a defendant’s intent where financial statements are prepared in a good faith attempt to comply with GAAP. The rules are no shield, however, in a case such as the present, where the evidence has showed that accounting methods known to be misleading - although perhaps at times fortuitously in compliance with particular GAAP rules - were used for the express purpose of intentionally misstating WorldCom’s financial condition in artificially inflating its stock price."

This reaffirmation of the principle in *Simon* clearly states that, at least in the Second Circuit’s view, the ultimate criterion for financial reporting is fair representation of financial condition, and while compliance with technical GAAP rules is persuasive, it is not sufficient.

5.5 Principles-based Accounting and the Adelphia Case

The principle in *Simon* also was applied in the Adelphia case involving the Rigas family. In this case, the prosecution successfully argued that Adelphia’s reclassification of certain debt in its financial statements was false and misleading, because it lacked economic substance and was designed to mislead investors. The allegation was not that the accounts did not comply with GAAP. Here too the initial conviction was appealed at the Second Circuit on the grounds that guilt requires proof GAAP was violated. The Court of Appeals disagreed, citing *Simon* to the effect that (Pet. App. 22a) “GAAP neither establishes nor shields guilt in a securities fraud case.” The appeals
court also noted (Pet. App. 23a) that good faith compliance with GAAP might indicate an absence of intent to defraud, but does not \textit{per se} shield a defendant from criminal liability.

5.6 The U.S. Audit Certificate

As noted above, the standard U.S. unqualified audit report states that the firm’s financial statements “present fairly, in all material respects, the financial position of the company at year end, the results of operations, and its cash flows, in conformity with generally accepted accounting principles.” Section 302(a) of the Sarbanes-Oxley Act requires a similar certification. The SEC view of this requirement is:

We believe that Congress intended this statement to provide assurances that the financial information disclosed in a report, viewed in its entirety, meets a standard of overall material accuracy and completeness that is broader than financial reporting requirements under generally accepted accounting principles. In our view, a “fair presentation” of an issuer’s financial condition, results of operations and cash flows encompasses the selection of appropriate accounting policies, proper application of appropriate accounting policies, disclosure of financial information that is informative and reasonably reflects the underlying transactions and events and the inclusion of any additional disclosure necessary to provide investors with a materially accurate and complete picture of an issuer’s financial condition, results of operations and cash flows.

In a recent public lecture, Zeff (2006) argues that the present form of the audit certification should be changed, as follows:

My premise is that principles should supplant, or at least supplement, rules in the conduct of the audit, just as they are being proposed to govern the setting of accounting standards. It should not be enough that the auditor's opinion reflects little more than a ticking off of the company's accounting methods against the rules of GAAP, even as challenging as that assignment is today. To serve the readers of financial statements and make the opinion paragraph of the auditor's report meaningful and not just a boilerplate, the auditor should be expected to treat "present fairly" as a substantive issue, and not as a "rubber stamp" of GAAP. Toward this end, I think that shareholders and the market would be better served by decoupling the auditor's opinion into whether the financial statements "present fairly" and whether they are in conformity with GAAP.

I do not understand either the need for change, or the rationale for “decoupling.” If one views financial reporting in common law terms, and not merely in terms of codified GAAP, the ultimate criterion for financial reporting in the U.S. already is fair representation of financial condition. In
other words, the system currently in place does not limit the auditor’s role to rule-ticking, even if
the recent scandals indicate that some or many managers and their auditors believed it does,
accounting textbooks routinely reinforce it and, as argued in following sub-section, bureaucratic
regulatory oversight encourages it.

5.7 Why Did Managers and Accountants View U.S. Financial Reporting as Rules-based?

In view of the many cogent reasons to view the ultimate standard for acceptable financial
reporting in the U.S. as being largely principles-based, why did a rules-based view of accounting
become so dominant? I propose the primary answer is regulation. If this hypothesis is correct, and
if a “rule-checking” mentality was in part responsible for the scandalous reporting behavior, then
regulation is one of the scandal culprits.

Some clues reside in history. Toward the end of the nineteenth century, the public joint
stock company became a dominant form of business organization in the United States. In 1895 the
first large audit firm (Haskins & Sells, now part of Deloitte & Touche) was founded. In 1904 the
first professional accounting organization was founded, and in 1916 it was reorganized into the
American Institute of Accountants. In 1917, the accounting profession and the Federal Trade
Commission together approved a uniform set of accepted audit principles and procedures. The
Institute promulgated best practice in other ways, most notably in regular commentary in its
member publication Journal of Accountancy and in 30 Special Bulletins it issued between 1920 and
1929. By 1929, when it revised the 1917 document on accepted practice, the Institute was acting
alone as an authoritative professional body, without apparent regulatory influence.

During 1932-1934, in the aftermath of the Great Crash of 1929, the American Institute of
Accountants debated and adopted six “broad principles of accounting which have won fairly
general acceptance”, and recommended a standard-form audit report certifying that the financial

---

60 This summary draws on Storey (1964), Carey (1969, Chapter 6), Moonitz (1970) and Zeff (1984, 1972, 2003a,b).
statements “fairly present, in accordance with accepted principles of accounting consistently maintained … .” The Institute then was a private-sector professional association, and a precursor to the American Institute of Certified Public Accountants (AICPA).

The SEC was established around the same time by the Securities Exchange Act of 1934, which empowered the Commission to set the rules governing financial reporting and disclosure of public companies. \(^{61}\) Initially, the SEC deferred on accounting matters to the accounting profession. In its *Accounting Series Release No. 4*, the SEC delegated authority to establish accounting standards to the private sector, and the AICPA’s Committee on Accounting Procedure (CAP) started issuing Accounting Research Bulletins, the precursor to contemporary Statements of Financial Accounting Standards. In 1953, the CAP issued ARB 43, codifying its previous pronouncements.

The Commission’s influence on financial reporting has grown substantially over the past seventy-five years. Its regulation of financial reporting is conducted primarily by its Office of the Chief Accountant, which states its mission as “establishing and enforcing accounting and auditing policy to enhance the transparency and relevancy of financial reporting, and for improving the professional performance of public company auditors in order to ensure that financial statements used for investment decisions are presented fairly and have credibility.” \(^{62}\)

From the outset the Commission strongly opposed departures from historical cost accounting (Walker, 1992; Zeff, 2007). At various points since, the Congress, the SEC and also the Treasury have intervened explicitly in standard-setting, notably in 1963 and 1967 (accounting for the investment tax credit), in 1974 (accounting for inflation), in 1978 (accounting for oil and gas exploration), in 1993-2005 (accounting for employee stock options), in 2000 (accounting for

---

\(^{61}\) Seligman (1982) provides a useful history of the SEC.

\(^{62}\) [http://www.sec.gov/about/offices/oca.htm](http://www.sec.gov/about/offices/oca.htm)
business combinations), and in 2008 (fair values for financial assets). Even when there has been no explicit intervention, the political/regulatory presence of Congress and the SEC has been a latent force in the background of FASB decision-making.

The Commission issues multiple Staff Accounting Bulletins, which “reflect the Commission staff's views regarding accounting-related disclosure practices” and “represent interpretations and policies followed … in administering the disclosure requirements of the federal securities laws.” It also issues “interpretive releases” on financial reporting and other matters. Its Accounting and Auditing Enforcement Releases outline its actions to enforce financial reporting rules via civil lawsuits and administrative proceedings.

When the SEC was created in the aftermath of the Great Depression, a proposal that the auditing of public firms be taken away from the private sector and assigned to a new government agency was debated but not enacted. In 2002, in the aftermath of the accounting scandals, the Sarbanes-Oxley Act created the Public Company Accounting Oversight Board, tightening regulatory oversight of the audit industry to a degree almost comparable to public ownership.

In sum, financial reporting in the U.S. has become increasingly regulated over time. As Mahoney (2009) concludes, “U.S. law has evolved from a system that was mostly contract-enabling to one that includes an increasing number of one-size-fits-all rules.” In parallel, codified financial accounting rules have become both more numerous and more detailed. Is this a coincidence?

Codification of law into formal rules seems a natural consequence of regulation, for several related reasons. First, a codified rulebook suits both the administrative mentality of government regulatory employees and the desire for certainty of the regulated: the regulated can construct an

---

63 Beresford (2001) provides an insider’s view of the business combinations issue.
administrative bureaucracy to interact with that of the regulator. Park (2007) expresses this explanation as follows:

Rulemaking reflects the mentality that securities regulation is a technical enterprise that should be left to experts who have created a comprehensive, efficient, administrative scheme. Principles-based enforcement actions reflect the demand that regulators punish conduct violating principles reflecting public values. For the most part, the regulated prefer a predictable regulatory regime, which rulemaking provides, while the public prefers decisive responses, which can be provided by principles-based enforcement actions.

Second, interpretation of principles is an inherently judgmental process, and involves the risk of being found wrong. Career bureaucrats are not widely viewed as not having either strong incentives or the requisite skill set to make judgments and take risks. Third, regulators receive their authority from and work under codified statutory law, not civil law, so their domain of operations naturally is heavily rules-based.

The hypothesis that regulation induces a rules-based view of the world helps explain several important financial reporting facts. Initially, it suggests that parallel increases in regulation of reporting and codification of U.S. financial accounting rules over time are related. It also suggests regulation was a contributor to the accounting scandals, many of which involved financial reporting that complied with the letter but not the principle of accounting rules (see Section 2.3 above). In addition, the regulatory preference for rules goes a long way toward explaining why the U.K. has been comparatively free of accounting scandals. Like the U.S., the U.K. is a common law country, but unlike the U.S. it historically has enjoyed very light regulation of its financial markets and financial reporting, and has more closely followed principles-based reporting standards.

By any standard, the U.S. is a litigious country (e.g., Coffee, 2007), which raises the question of whether private litigation also encourages a rules-based view of financial reporting. This is difficult to assess, because private litigation has been influenced substantially by the

---

67 Note the application of the Stigler (1964, 1971) - Peltzman (1976) theory of regulation.
Securities Acts. Most notably, the courts allow private plaintiffs to sue for damages arising from a violation of Rule 10b-5 of the Securities Exchange Act (e.g., Mahoney, 2009), a practice that has come under much criticism for increasing the prospects and hence the frequency of private litigation (e.g., Committee on Capital Markets Regulation, 2006). Consequently, we do not observe separate effects of statutory and private litigation, and one can only conjecture what the separate effect of private litigation on a rules-based view of financial reporting might be. My own view is that a pure common law system, absent regulation, would operate much as in the Simon case, described in Section 5.2 above: rules-based GAAP would be persuasive evidence the financial statements fairly represent the company’s financial position, because rules would tend to codify generally accepted best practice; willful failure to comply with GAAP would suggest an intent to mislead; but compliance with rules-based GAAP would not necessarily imply financial statements satisfy the over-arching principle of fairly representing the company’s financial position. That is, there would be rules, but they would be subservient to principles.

Ironically, in the aftermath of the accounting scandals, Section 108(d) of the Sarbanes-Oxley Act directed the SEC to conduct a study on the adoption of a principles-based accounting system and deliver it to Congress by July 30, 2003. The duly-delivered report argued in idyllic terms for the best of both worlds: for what the Commission termed “objectives-oriented standard setting,” which it explained as follows (emphasis in original):69

In our view, the optimal principles-based accounting standard involves a concise statement of substantive accounting principle where the accounting objective has been incorporated as an integral part of the standard and where few, if any, exceptions or internal inconsistencies are included in the standard. Further, such a standard should provide an appropriate amount of implementation guidance given the nature of the class of transactions or events and should be devoid of bright-line tests. Finally, such a standard should be consistent with, and derive from, a coherent conceptual framework of financial reporting.

---

SEC Chief Accountant Robert K. Herdman earlier had stated: “An ideal accounting standard is one that is principle-based and requires financial reporting to reflect the economic substance, not the form, of the transaction.” A similar stance was taken by FASB (2002).

If the U.S. did attempt a move toward principles-based accounting, for example by allowing domestic companies to report under IFRS, a skeptical prediction is that the SEC would pay only lip service to the change, and that this would be fine with the audit firms. Principles-based financial reporting would require companies and their auditors to be both able and willing to make the judgments necessary to apply broad principles in specific circumstances. It also would require SEC staff to be able and willing to evaluate those judgments. A principles-based set of accounting standards such as those embodied in IFRS soon would be supplemented by a flurry of implementation guidelines or rules issued by FASB, the SEC or individual audit firms, the net effect being similar to the current system. Stated differently, the current financial reporting system is an endogenous result of U.S. market and political/regulatory forces, and hence is unlikely to change in substance (as distinct from appearance) unless those forces themselves change. The most substantial change in response to the accounting scandals – drastically increased regulation under the Sarbanes-Oxley Act of 2002 – seems likely to create more rule-making, not less.

6. Was Sarbanes-Oxley Necessary?

Perhaps the most controversial issue is whether Sarbanes-Oxley was necessary. Seventy-five years after the only remotely comparable increase in U.S. securities regulation – the 1933-34 passage of the Securities Acts and the creation of the SEC – there still is scant evidence on whether it was beneficial. Stigler (1964) and Benston (1969, 1973) concluded the Securities Acts did not

---

benefit investors, though the issue remains controversial to this day.\textsuperscript{72} It therefore is not surprising that, only six years after its enactment, the legacy of Sarbanes-Oxley is largely conjectural.\textsuperscript{73}

This section discusses the market and regulatory perspectives on Sarbanes-Oxley in terms of first the deterrence and second the detection and correction of reporting negligence and fraud. It then discusses the issue of whether the focus on heavy deterrence penalties in the U.S. is an effective regulatory model (Coffee, 2007).

At the outset, it is worth noting that the desirability or otherwise of additional regulation is not the same question as whether there was market failure. The U.S. has operated a mixed market-and-regulatory system of financial reporting since the Securities Acts of 1933-34. It makes no more sense to argue that the market failed and hence more emphasis should be placed on regulation, than it does to argue that regulation failed and hence more emphasis should be placed on markets.

Similarly, the Sarbanes-Oxley Act is not the only change that has taken place in financial reporting in response to the accounting scandals. It is a mistake to compare Sarbanes-Oxley with the pre-Sarbanes-Oxley system, without taking into account the changes that would have occurred in market mechanisms, absent legislative action. It seems unlikely that managers, boards, auditors and other parties would have failed to change their practices. To take a small example, after paying approximately $9 billion to settle private litigation arising from their transactions with Enron, the banks would seem less likely to enter similar transactions with their clients in future.

Markets generally are not static mechanisms: their institutional structure learns from and evolves in response to past events (e.g., Hayek, 1945, 1988; Wilson, 1975; Nelson and Winter, 1982; North, 1990; Waymire and Basu, 2008). Considerable changes in financial reporting and governance norms and practices could be expected to have occurred in response to what was learned from the scandals, though many such changes would not be observable because they were

\textsuperscript{72} Coffee (1984) provides a survey at the fifty-year point.
\textsuperscript{73} Analyses are provided by Romano (2005), Coates (2007), Engel, Hayes and Wang (2007), Leuz (2007), and Zhang (2007).
pre-empted by and seemingly due to legislation. Nevertheless, some changes were evident during the brief interval between the scandals and the enactment of Sarbanes-Oxley. For example, Leuz and Schrand (2008) document increased disclosure in firms’ 2001 year-end 10-K filings, particularly in the Management Discussion and Analysis section, in describing related-party transactions (an important issue in Enron), and in the financial statements and footnotes. The increases are larger for Arthur Anderson clients. They also report an increased number of 8-K filings and conference calls.

Assessing the impact of Sarbanes-Oxley therefore is not a simple matter of comparing pre- and post-Sarbanes-Oxley practices, without careful specification of the counter-factual. Many of the observed changes would have occurred due to market forces. Furthermore, it even is possible that Sarbanes-Oxley inhibited (as distinct from pre-empted) market-based changes; for example, its increased jail penalties described below conceivably could have deterred managers and audit firms from pushing for a more principles-based system of reporting. Assessing the impact of Sarbanes-Oxley is not as easy as the flurry of pre-post regulation studies might imply: it requires careful specification of the assumed counter-factual.

6.1 Deterrence

Obviously, market and regulatory mechanisms were not sufficient to deter the spate of accounting scandals revealed in 2001-02. Leaving aside the issue of the optimal frequency of financial reporting negligence and fraud (not zero, because deterrence costs are not zero), the scandals imply a failure of both market and regulatory deterrence mechanisms.

Much attention has been paid to the Sarbanes-Oxley requirement that the financial statements be certified by the CEO and CFO, and to the Section 802(a) statutory fines and jail penalties described below conceivably could have deterred managers and audit firms from pushing for a more principles-based system of reporting. Assessing the impact of Sarbanes-Oxley is not as easy as the flurry of pre-post regulation studies might imply: it requires careful specification of the assumed counter-factual.

6.1 Deterrence

Obviously, market and regulatory mechanisms were not sufficient to deter the spate of accounting scandals revealed in 2001-02. Leaving aside the issue of the optimal frequency of financial reporting negligence and fraud (not zero, because deterrence costs are not zero), the scandals imply a failure of both market and regulatory deterrence mechanisms.

Much attention has been paid to the Sarbanes-Oxley requirement that the financial statements be certified by the CEO and CFO, and to the Section 802(a) statutory fines and jail penalties described below conceivably could have deterred managers and audit firms from pushing for a more principles-based system of reporting. Assessing the impact of Sarbanes-Oxley is not as easy as the flurry of pre-post regulation studies might imply: it requires careful specification of the assumed counter-factual.

6.1 Deterrence

Obviously, market and regulatory mechanisms were not sufficient to deter the spate of accounting scandals revealed in 2001-02. Leaving aside the issue of the optimal frequency of financial reporting negligence and fraud (not zero, because deterrence costs are not zero), the scandals imply a failure of both market and regulatory deterrence mechanisms.

Much attention has been paid to the Sarbanes-Oxley requirement that the financial statements be certified by the CEO and CFO, and to the Section 802(a) statutory fines and jail penalties described below conceivably could have deterred managers and audit firms from pushing for a more principles-based system of reporting. Assessing the impact of Sarbanes-Oxley is not as easy as the flurry of pre-post regulation studies might imply: it requires careful specification of the assumed counter-factual.

---

74 Ball and Shivakumar (2008) report a sharp increase in the proportion of total annual price-relevant information that is released at earnings announcements, starting during or after 2000 but, in comparison with the research design in Leuz and Schrand (2008), their time dating is not fine enough to distinguish pre- and post-Sarbanes-Oxley effects.

75 Mahoney (2009) notes another event that confounds the effect of Sarbanes-Oxley: prosecutorial aggressiveness increased at the same time, in response to the scandals (see also Section 3.1 above).
sentences of up to twenty years for a knowingly false certification. I doubt the effectiveness of
heavy additional statutory penalties as a deterrence mechanism. Prior to Sarbanes-Oxley, the CEO
and CFO were required to sign and attest to the financial statements, and faced a heavy
combination of code law penalties (fines and jail sentences of up to five years) and common law
penalties (loss of reputation and employment prospects, damage awards under civil lawsuits), but
this did not deter the accounting frauds. Coffee (2007, p.2) concludes from an extensive survey of
securities regulation internationally.\footnote{These data were collected earlier by Jackson (2007), who reached similar conclusions on this issue.}

The United States is unique not in its expenditures on securities regulation, but in the amount
and severity of the penalties it imposes. … . For example, in 2005/06, the financial penalties
imposed by the SEC exceeded those imposed by the U.K.'s Financial Services Agency ("FSA")
by a thirty to one ratio, which, even after adjustment for differences in market capitalization,
still translates into a ten to one ratio. … The greater use of public enforcement in the United
States is more than paralleled by corresponding disparities in private enforcement and the use
of the criminal sanction. Virtually alone, the United States recognizes the class action and the
contingent fee. The actual financial sanctions imposed by private enforcement in the United
States exceed those imposed by public enforcement … .

The pre-Sarbanes-Oxley regime in the U.S. relied more than any other on penalties and
enforcement to deter fraud, yet it was the epicenter of the 2001-02 scandals. What reason is there to
believe that even larger penalties will have a substantial deterrence effect? In this sense, the
Sarbanes-Oxley prescription looks like “just more of the same.”

A possibly relevant factor, discussed in section 2.4 above, is that managers who commit
accounting fraud frequently have a “gambler’s mentality,” loosely defined as an excessively strong
belief in oneself or one’s company, and in the likelihood that real performance will improve before
their fraud is detected. Increased penalties seem unlikely to deter people who can convince
themselves that they are doing the right thing. The fear of twenty year jail sentences seems more
likely to make innocent but prudent managers report excessively conservatively.
One lesson from the demise of Arthur Andersen (and the demise a decade earlier of Laventhol and Horwath) is that market-based effects such as reputation, bonding and insurance, are large. Reliable financial information is valued by investors, lenders, suppliers, employees, customers and the public, so markets place a high value on trust. In principle, reputational and other market mechanisms deter managers and auditors from violating that trust. However, another lesson from the scandals is that reputation effects sometimes can work the other way. Managers with reputations built around strategies that are widely believed to be successful, but in fact are not (as in the Enron, WorldCom and HealthSouth cases), might perceive an incentive to hide their poor performance. Especially if they have an excessively strong belief that real performance will improve before the fraud is detected, they might perceive their incentive as being to preserve their reputation in the short run by faking results.

Deterrence mechanisms aim to reduce financial reporting fraud by imposing costs on the manager or auditor, conditional on the fraud being detected. The U.S. system historically has imposed seemingly high deterrence costs, yet this did not deter the scandalous behavior in 2001-02. All things considered, I am not convinced that increasing deterrence costs even further is an efficient use of resources.

6.2 Detection and Correction

Attention also has been given to the extensive inspection, documentation and certification of internal control structure and procedures for financial reporting required by Sarbanes-Oxley’s Section 404, and other provisions of the Act directed at improved fraud detection. These include requiring external auditors to have direct access to the board’s audit committee, requiring audit partner (but not audit firm) rotation, prohibiting audit firms from supplying various consulting activities to clients, and protecting “whistle blowers” from retaliation.
In contrast to its record in deterrence, there is some evidence the pre-Sarbanes-Oxley system worked relatively well in detecting and correcting false reporting. One piece of evidence is the speed with which detection occurred. As Kedia and Philippon (2007) observe, the welfare loss from firms overstating their financial condition is that they over-utilize both labor and capital, distorting the allocation of real resources. Only when the overstatement is revealed do they shed excess labor and capital. The welfare loss arising from the overstatement therefore arises from the inefficient use of resources during the period between inflation and detection. While the conclusion is subjective, my impression is that most of the financial reporting fraud was detected and revealed relatively quickly, and that failing managers and their failing strategies were dumped relatively quickly, thus limiting the continuing losses. Miscreant managers, auditors, board members and counterparties were replaced relatively quickly.\footnote{See Srinivasan (2005), Arthaud-Day \textit{et al.} (2006) and Desai, Hogan and Wilkins (2006).}

Another piece of evidence is the parties who uncovered the frauds. To my knowledge, there has been no systematic, reliable compilation of accounting fraud detection during 2001-02, but I am aware of no financial reporting frauds that were detected by regulators, as distinct from internal auditors, employees, external auditors, private parties on the other side of irregular transactions, security analysts, short sellers, the plaintiffs bar, and the press. This impression is broadly consistent with the evidence of Dyck, Morse and Zingales (2007).\footnote{The Dyck, Morse and Zingales (2007) data do not directly identify the party originally detecting fraud. For example, many or most of the 14\% of all frauds attributed to the media would have originated in information detected by other parties and then revealed to the media. A similar observation can be made about the 6\% attributed to the SEC, which primarily acts on information received from other parties, including auditors.}

My conclusion is that the pre-Sarbanes-Oxley system worked much better in detecting and correcting the problem that emerged in the 2001-02 scandals than is commonly believed. Whether the extensive provisions of the Sarbanes-Oxley Act will lead to better or earlier detection is a matter for conjecture at this point.
6.3 Is Deterrence the Best Regulatory model?

Coffee (2007) raises the intriguing thesis that the U.S. employs an inefficient regulatory model, by allocating disproportionately high resources to enforcement of the law and punishment of securities fraud, after it has occurred and been detected. He describes this as “ex post” regulation, and contrasts it with a system in which regulators consult more with companies “ex ante,” advising them on appropriate behavior. He observes (2007, p. 80):

U.S. regulators are roughly comparable to those of other common law countries in their budget and staffing levels, but not in the number of enforcement actions brought or the magnitude of the sanctions imposed. Here, they inhabit a world largely of their own. This raises at least the possibility that, outside the United States, more is done through “ex ante” regulation than through “ex post” enforcement. Although the SEC also uses “ex ante” regulation, it seems more committed than other regulators to a policy of general deterrence through large penalties.

Two features of the recent accounting scandals are broadly consistent with Coffee’s thesis: the considerably lower frequency of reporting fraud in the U.K., where the Financial Services Agency regulates securities markets in a more consultative fashion and with less emphasis on enforcement after the fact; and the comparative success of the U.S. system in detecting and correcting financial reporting fraud, but not in deterring it in the first instance.

7. The Long Run Cost of Sarbanes-Oxley.

From a political/regulatory perspective, the principal downside of Sarbanes-Oxley seems to have been viewed as its short-term compliance costs, especially for small companies and for potential foreign registrants in the U.S., and the public debate has been focused on this issue.\textsuperscript{79} From a market perspective, however, the principal costs would seem to be more long run in nature. Chief among them is that the advent of Sarbanes-Oxley signals an accelerated politicization of corporate governance and financial reporting.

Common law countries, the U.S. included, historically have enjoyed a largely common law governance system, with modest code law intervention. They have operated a system of

\textsuperscript{79} Compliance costs for 2005 alone are estimated at $6.1 billion (Wall Street Journal 17 October 2005).
“shareholder value” governance, with major decision rights attached to shareholders, who are the residual claimants on the firm. Increased politicization of corporate governance and financial reporting places the historically successful “shareholder” governance system at risk of degenerating into a “stakeholder” governance system, with representation of major political blocs in writing the rules and in running corporations. This is the system that produced the ultra-conservative, low-quality financial reporting system that is associated with Continental Europe (Ball, Kothari and Robin, 2000). A major move toward such a system would involve abandoning – not strengthening – the historical strengths of the U.S. system, in a political over-reaction to one episode of scandalous financial reporting.

8. Concluding Observations

Markets need rules, and rely on trust. U.S. financial markets historically had very effective rules by world standards, the rules were broken, and there were immense consequences for the transgressors. The system worked surprisingly well in detecting but not in preventing the problem. Was Sarbanes-Oxley an over-reaction? It depends on whether one takes a market or a political/regulatory perspective.

---

80 Alchian and Demsetz (1972).
References


