Auditor Credibility Impairment Spillovers: Evidence from Three Andersen Engagements

Rajib Doogar*
Theodore Sougiannis

University of Illinois

Hong Xie

Syracuse University

October, 2007

* Corresponding author. Address all correspondence to: 1206 S. Sixth Street, Champaign, IL 61820, U.S.A. e-mail: doogar@uiuc.edu, Phone: 217.244.8083, Fax: 217.244.0902.

An earlier version of this study was titled "The Impairment of Auditor Credibility: Stock Market Evidence from the Enron-Andersen Saga." We acknowledge research assistance from John Grenier, our many conversations with Jim McKeown and helpful comments and suggestions by two anonymous reviewers and seminar participants at The University of Cincinnati, Hong Kong University of Science and Technology, The University of Illinois at Chicago, The University of Illinois at Urbana-Champaign, The University of Kentucky, McMaster University, The University of Notre Dame, New York University, Texas A&M University and Yale University and attendees at the 2003 ISAR Conference, the 2004 Auditing Section Mid-Year and the 2004 AAA Annual Meeting.

Auditor Credibility Impairment Spillovers:
Evidence from Three Andersen Engagements

Abstract

Does bad news about one auditor's conduct affect the credibility of *all* auditors? We examine auditee abnormal stock returns around twenty-five bad news events involving Arthur Andersen LLP's (Andersen) questionable audits of Waste Management, Sunbeam, and Enron. In twelve of these twenty-five windows, market-wide mean abnormal returns are negative and significant. During these twelve windows, we also find, consistent with expectations, that portfolios of low quality (high abnormal accruals, low cash flow) auditees are, on average, penalized significantly more than portfolios of high quality (low abnormal accruals, high cash flow) auditees. Eight of these twelve windows predate October 2001 when news of Enron's financial problems first became public. The strongest reaction occurs in June 2001 when the SEC penalizes Andersen over the Waste Management engagement. Overall, we present strong and systematic evidence that bad news about the conduct of one auditor generates significant negative externalities (spillovers) to *all* auditors.

Auditor Credibility Impairment Spillovers:
Evidence from Three Andersen Engagements

I. Introduction

We investigate whether bad news about one auditor impairs the credibility of *all* auditors (*inter-firm credibility impairment spillovers*). Several studies have investigated such spillovers around events related to the 2001-2002 implosion of Enron and to the demise of its auditor, Arthur Andersen, LLP (hereafter, Andersen). Overall these studies find little evidence of inter-firm spillovers. After testing for spillovers around the time of the infamous January 10, 2002, "shredding admission date" Chaney and Philipich (2002, p. 1244) conclude that "Whether the decline in reputation observed for Andersen may spill over to other audit firms is yet to be determined, but clearly others are already worried." In subsequent studies, Callen and Morel (2002) find no evidence of spillovers to other auditors while Asthana, Balsam and Krishnan (2003) report that the market reaction appears to be only weakly related to proxies for investors' assurance and insurance needs.[1] Our study contributes to this stream of research by investigating, in addition to the Enron-related events investigated in prior research, a number of pre-Enron events that set the stage for Andersen's eventual demise and may have impaired auditor credibility. We also employ a novel, portfolio-based approach to impairment detection.

Enron started to implode in October 2001, shortly after the SEC had (in May and June 2001) imposed unprecedented sanctions on Andersen for inadequate audits of Waste Management and Sunbeam. The press coverage of the SEC sanctions had questioned not only the credibility of

[1] Callen and Morel (2003) summarize their findings thus "Univariate and regression results indicate that event day abnormal returns and two-day cumulative abnormal returns are generally not significantly different from zero for both Andersen and big five non-Andersen audit clients during the months of October 2001 through January 2002. Cumulative abnormal returns over all event days are marginally significantly negative for the Andersen sample and insignificantly negative for the non-Andersen control sample. The Andersen sample lost 4% more than the non-Andersen sample in risk-adjusted returns over the events in the four-month period." We discuss differences between each of these studies and our work in more detail in Section II.

Andersen's audits but also that of audits in general. If Andersen's conduct in the Waste Management and Sunbeam audits had already caused substantial inter-firm spillovers, studies focusing on later Enron-related events *alone* would detect only the *incremental* inter-firm spillover effects of those (later) events. In this study we therefore examine, in addition to Enron-related events examined in prior research, key events related to Andersen's audits of Waste Management and Sunbeam since these earlier events could also potentially have triggered inter-firm spillovers.

We employ a four-step process to document spillovers. First, using Lexis-Nexis we identify twenty-five auditor-related bad news windows pertaining to Andersen's Waste Management, Sunbeam and Enron audits. Our approach to window selection favors inclusiveness, primarily because prior research provides little guidance in identifying specific types of events that can trigger spillovers. Second, from this collection of bad-news windows, we select as *potential spillover windows* those windows during which mean cumulative abnormal returns for the sample auditees are negative.[2] There are a couple of reasons for this approach. Miller (2006) finds not all press coverage is necessarily news to investors. Neither does any theory predict which news stories can be expected to be treated by investors as bad news (and therefore trigger spillovers). So we follow prior research and pick events with negative market-wide (or sample-wide) returns to look for spillovers.[3]

[2] Prior studies of spillovers generally stop at this point and interpret negative mean abnormal returns during bad news windows as evidence of auditor credibility impairment. While this is a valid test in short windows, it is worthy of note that when only one or two events are studied—as is the case in most prior studies of inter-firm spillovers — such returns also can reflect the effects of other contemporaneous, and potentially confounding, events.

[3] If our goal were to investigate whether auditor bad news stories on average result in spillovers to other auditors, partitioning windows in this fashion might be construed as *data snooping*. Our goal, however, is to document the existence and extent of spillovers. In this case, lumping all dates into one analysis, regardless of whether investors reacted to the news or not is tantamount to reducing the power of our tests—because spillovers may happen in only a few event windows—and may result in the phenomenon of "throwing the baby out with the bathwater".

Third, we investigate whether the cross sectional variation in abnormal returns during the set of potential spillover windows selected in step two is, on average, positively related to the quality of auditee accounting estimates and to auditee financial health. In Section III we show formally that the value of auditing is higher for lower quality auditees (auditees about which investors' *ex-ante* uncertainty is higher). Consequently when auditor credibility is impaired, security prices of low quality auditees suffer more than prices of high quality auditees. Empirically, when auditor credibility is impaired, investor uncertainty about the quality of accounting estimates (e.g. accruals) can be expected to increase: auditees with lower financial reporting quality (larger *abnormal* accounting estimates) will, then, be viewed as less attractive than auditees with higher financial reporting quality (smaller *abnormal* accounting estimates).[4] Prior research also suggests that investors view managers of financially weaker auditees as having greater incentives to misreport (e.g. Palmrose 1987). Consequently, both theory and intuition suggest that when auditor credibility is impaired, one would expect investors to "fly to quality" so that abnormal stock returns to *low quality* auditees (auditees with lower financial reporting quality *and* poor financial health) would be expected to be more negative than those to *high quality* auditees (auditees with higher financial reporting quality *and* good financial health).[5] Fourth, prior research argues that news that affects only financially distressed (*distressed*) auditees may not speak unambiguously to the impairment of auditor credibility (Baber et al. 1995). We, therefore, repeat step two separately for distressed and non-distressed (*solvent*) auditees.

[4] We measure abnormal accounting estimates using both absolute and signed values of abnormal accruals estimated using the modified Jones (1991) model and, alternatively, the Dechow and Dichev (2002) measure of accrual quality. Absolute abnormal accruals are widely used in the accounting literature to measure earnings management, earnings quality or audit quality (e.g., Klein 2002; Johnson, Khurana and Reynolds 2002; Myers, Myers and Omer 2003).

[5] Auditee financial health is measured using either cash flow from operations or free cash flow.

Steps three and four extend the standard negative mean abnormal return test for spillovers (step two). By linking the magnitude of investor reactions to accruals and cash flows, step three increases our confidence that the news heightened investors' uncertainty about audited financial statements, i.e., impaired auditor credibility. By showing that solvent auditees, who are less likely to need auditor-provided insurance, also are penalized, step four further bolsters the conclusion that the news impaired auditor credibility.[6] We also include in our tests all non-Big-Five auditees in the Compustat-CRSP universe. While non-Big-Five auditors are thought to provide a lower level of assurance than Big Five auditors (Simunic 2003, Francis 2004), the possibility that auditor-related bad news triggers differential investor reactions for Non-Big Five auditees relative to Big Five auditees has not been addressed in prior research. Our study provides empirical evidence on the extent to which actions of a Big Five auditor affect the credibility of non-Big-Five auditors.

Our principal findings are as follows. During twelve of the twenty five event windows we examine, for both Andersen and non-Andersen auditees, mean cumulative abnormal returns are significantly negative, low-quality auditees are penalized significantly more than high-quality auditees and mean abnormal returns are significantly negative for both solvent and distressed auditees. Overall, during these twelve windows we find compelling evidence of market-wide auditor credibility impairment spillovers. Interestingly, eight of these windows precede the Enron-related windows examined in prior studies.

[6] Intuitively, the magnitude of the event window returns provides another criterion for evaluating whether auditor credibility was impaired. The average market capitalization of our sample is about $9.5 trillion per window. Even one-tenth of one percent of this amount, $9.5 billion would have been beyond the collective ability to pay of the Big Five auditors. Consequently, negative abnormal returns in a single window in excess of 1% (a loss of about $95 billion) and total value losses of about 5% ($475 billion) over all windows are difficult to reconcile with any reasonable estimate of auditor-provided insurance.

SEC sanctions and criticisms of Andersen's audits affect all auditors' reputations to about the same extent. On the other hand, disclosures about the specific nature of the accounting irregularities overlooked by or mistakes made by Andersen appear to affect the reputations of other Big Five auditors more than those of non Big Five auditors. The most compelling evidence of market-wide spillovers from a single event occurs around June 19, 2001, when the SEC censures Andersen for substandard audits of *Waste Management* and imposes an unprecedented $7 million fine on the auditor.

Overall, our results should be of interest to regulators entrusted with the well functioning of financial markets, to auditors, and to financial statement producers and users. Our study makes three contributions to the literature. First, it provides systematic evidence consistent with the existence of substantial auditor credibility impairment spillovers. This evidence can be helpful in evaluating the social desirability of actions aimed at stemming or mitigating such spillovers.[7] Second, we find that the magnitude of the spillover to Other Big Five and to non-Big Five auditors varies with the type of event. Whether different types of events have such differential effects in other settings as well offers an interesting avenue for future research. Third, our approach to detecting spillovers by decomposing stock returns based on accounting variables can be helpful in future research that analyzes the market response to accounting or auditing events. The rest of the paper is organized as follows. Section II reviews key related studies. Section III presents a model of security price revisions when investor perceptions of audit quality are impaired. Section IV describes our research methodology and section V discusses sample

[7] The magnitude of investor losses we document can serve as a lower bound for the costs to society of such spillovers (there may be other costs that stock returns do not capture) and thus, provide a useful referent for policy makers. Policies expected to impose costs vastly in excess of the losses suffered by investors could be readily ruled out of consideration while those expected to impose costs small in relation to the magnitudes of investor losses may find greater acceptance.

selection and variable measurement issues. Section VI presents the main findings, Section VII the results of several robustness checks, and Section VIII a summary and conclusions.

II. Related Research

Why might bad news about one auditor impair the credibility of all auditors? U.S. auditors share substantial similarities in the inputs and audit processes they use and the outputs they produce. On the input side, the audit labor used by all U.S. auditors consists largely of individuals who have demonstrated sufficient knowledge of U.S. accounting and auditing standards to acquire certification as a Certified Public Accountant (CPA). With respect to the audit process, all U.S. audits must comply with the requirements of U.S. Generally Accepted Auditing Standards (GAAS) (Palmrose 1987; Buckless and Peace 1993; Jennings, Kneer and Reckers 1993; Latham and Linville 1998). Finally, on the output side, all financial statements accompanied by a clean audit opinion must conform to U.S. Generally Accepted Accounting Principles (GAAP). In addition to these structural elements, labor mobility in the audit industry contributes to the diffusion of audit know-how as well. Consequently, it would not be surprising if inter-firm spillovers in audit markets were to occur because perceptions of a quality problem with one auditor's work lead investors to suspect "perhaps a similar, yet undetected, quality problem" for other auditors as well (Chaney and Philipich 2002, pp. 1239-40).

A number of studies have investigated own-firm effects (*intra-firm* spillovers) resulting from auditor credibility impairment (Firth 1990; Menon and Williams 1994; Baber, Kumar and Verghese 1995; Chaney and Philipich 2002; Callen and Morel 2003; Asthana, Balsam and Krishnan 2003; Krishnamoorthy, Zhou and Zhou 2006; Cahan, Emmanuel and Sun 2005). These studies find that bad news about an auditor's work lowers its client retention rates, slows

its practice growth in immediately subsequent periods and triggers negative mean abnormal returns to its own auditees.

Three of these studies have also investigated spillovers to other auditors (*inter-firm* spillovers) from bad news about Andersen's conduct of the Enron audit. Chaney and Philipich (2002) examine whether Andersen's January 10, 2002, "shredding admission" resulted in negative mean abnormal returns to Andersen and other Big Five auditees. Using the S&P1500 sample, they find reliable evidence of intra-firm spillovers to other Andersen auditees, but not of inter-firm spillovers to other Big Five auditees (*ibid* p. 1244, quoted earlier).[8] Callen and Morel (2003) examine mean abnormal stock returns to Andersen and non-Andersen auditees between October 2001 and January 2002.[9] They find no evidence of inter-firm spillovers: daily and cumulative abnormal returns to non-Andersen Big Five auditees in their study are not significantly different from zero. Finally, Asthana et al. (2003) find significantly negative mean abnormal returns to both Andersen and Other Big Five auditees around the January 10, 2002 *shredding window* and the February 4, 2002 *Powers report window.* However, they find that "variables proxying for insurance and assurance needs of the clients" explain little of the cross-sectional variation in stock returns for non-Andersen auditees.[10]

---

[8] Discussing the results reported in Table 7 of their study, Chaney and Philipich (2002) report that "in general we find a negative BHAR when using the Russell 3000 index for all auditors and mostly positive BHARs when using the equally weighted index. … The BHAR for Andersen's audit clients is more negative and, in most cases, statistically different from the BHAR for the other Big 5 auditors regardless of the market index used." Overall, they conclude that the market reaction on January 10, 2002, was "attributable to a decline in Andersen's reputation."

[9] Callen and Morel (2003) do not employ any cross-sectional tests. [Chaney and Philipich contain some cross-sectional tests. See their table 9].

[10] Asthana et al. (2003) use absolute abnormal accruals as their primary proxy for assurance needs and Return-on-Assets (ROA), the annual stock return and Altman's z-score as proxies for insurance needs. They conclude that "assurance effects, if they exist, appear to be much weaker for clients of the other Big 5 firms than they were for Andersen." With respect to the insurance effects, their evidence is similarly mixed: financial distress is never significant, ROA is significant twice in eight regressions (once with the wrong sign) while annual stock returns are significant in one window (January 10, 2002) but not in the other (February 4, 2002). They interpret this evidence as "providing some support for the insurance hypothesis."

Our study differs from these prior studies in several important ways. First, and most critically, we theorize and document that during spillover windows (identified as in prior research by a negative mean market-wide abnormal return for the event window) low quality auditees are more heavily penalized than high quality auditees. By contrast, the approach used in prior studies relies on documenting the assurance and insurance impairment effects resulting from auditor-related bad news (Menon and Williams 1995; Baber et al. 1995; Asthana et al. (2003). One concern with the prior approach is that when auditor credibility is impaired, investors will perceive the audits to be of lower quality (assurance impairment) and, as a consequence, will also expect more future audit-related litigation against that auditor. This higher expected litigation rate will in turn, lead to a decline in the insurance value of the audit as well. In other words, it may be practically impossible to separate insurance impairment and assurance impairment effects when one suspects market-wide auditor credibility impairment to have occurred. We therefore focus on examining whether, during potential spillover windows, there is a discernible flight-to-quality effect in which low quality auditees are penalized more than high quality auditees.

Second, our dependent variable, abnormal stock returns, is computed based on a four-factor model as described in Carhart (1997) which controls for the effects of beta, book-to-market, size and momentum—a more comprehensive set of known risk factors than used by either Chaney and Philipich (2002) or Callen and Morel (2003). We, thus, treat all four factors as control variables in contrast to some studies (e.g., Menon and Williams 1994, Baber et al. 1995, Asthana et al. 2003) that treat factors such as beta and momentum as test variables. Third, we examine *more* and *different* events (i.e., pre-Enron events) that could lead to spillovers than prior studies. Fourth, unlike prior studies which focus only on Big Five auditees, we examine all auditees in

the Compustat-CRSP universe.[11] Fifth, we present evidence that the spillover affected both financially distressed as well as non-distressed (*solvent*) auditees, suggesting that the effects of the news cannot be attributed *purely* to impairment in the insurance value of the audit (Menon and Williams 1994, Baber Kumar and Verghese 1995). Sixth, we use both portfolio and regression approaches to conduct our tests since each approach has notable strengths and some weaknesses (Campbell et al., 1997, MacKinlay 1997). Prior studies, in contrast, rely mainly on regression analysis to investigate determinants of cross-sectional return variation. Finally, our results are much stronger and more unambiguously link the cross-sectional variation of stock returns during spillover windows to cross-sectional variation in audit quality *and* we document that this pattern is *not* present during non-spillover windows.

Collectively, the larger sample of auditees and the larger number of windows we examine, coupled with the flight-to-quality test, permit us to present clearer and more compelling evidence that auditor-related bad news leads to significant, market-wide, auditor credibility impairment spillovers in U.S. capital markets.

III. Security Prices and Auditor Credibility Impairment

Consider a setting in which two parties, *insiders* and *outsiders* are asymmetrically informed about an investment opportunity which can either be very profitable (*Good*) or moderately profitable (*Bad*).[12] Akerlof (1970) shows that in such asymmetric information settings, markets may fail. A demand for *auditing* emerges naturally in such settings: auditors facilitate credible communication by certifying insider assertions as to project type (DeAngelo 1981, Dopuch and

[11] During the period of our study, the audit firms of Arthur Andersen LLP, Deloitte and Touche LLP, Ernst and Young LLP, KPMG LLP and PricewaterhouseCoopers LLP were collectively known as the Big Five auditors. All other auditors are classified, for the purposes of this study as Non-Big-Five auditors.

[12] Insiders can be thought of, depending on the context, as owner-entrepreneurs (Leland and Pyle 1977), as managers (Datar, Feltham and Hughes 1990) or as current generation of stockholders (Dye 1993).

Simunic 1980, Gjesdal 1981, Antle 1982 & 1984, Datar, Feltham and Hughes 1991, Dye 1993, Schwartz 1997).

Our analysis focuses on the impact of a change in audit quality on investors' equilibrium security valuations. To develop the arguments formally, we construct a model of security prices when investors are fully rational and care about both project quality and about auditor credibility. More specifically, in our model economy, project types are not perfectly known to investors and auditing is costly. Investors hold beliefs about the level of audit quality (the auditor's ability to detect true project type) and value securities in keeping with expected audit quality. We then explore the cross-sectional variation in the impact of a shock to audit quality by performing comparative statics on the security price as both project quality and investors' perceptions of auditor credibility change.[13] Formally, we assume that

(A1) There are two types of projects in the economy, Good and Bad. Good projects always return 1, Bad projects 0.

(A2) The proportion of Good projects is γ ($0 < \gamma < 1$).

(A3) Bad projects are, with probability θ ($0 < \theta < 1$), *ex-ante* indistinguishable from Good projects.

(A4) Managers have no personal wealth and cannot signal by retaining ownership.

(A5) The cost of capital is r, [$0 < r < 1$].

A1 is an innocuous stylization that simplifies computations. A2 posits that there are some Bad projects to screen and A3 posits that some Bad projects are not self-evident before they are undertaken (were this not the case, no Bad projects would ever be undertaken and there would be

[13] For similar approaches to modeling the demand for auditing, see Balachandran and Nagarajan 1987, Dye 1993, Schwartz 1997, Chan and Pae 1998, Hillegeist 1999, Pae and Yoo 2001). The value of auditing arises in these models from the auditor's differential ability to (imperfectly) screen out unviable projects. This class of models is most relevant to our discussion since screening out bad projects alters the distribution of returns investors can expect and thus provides an intuitive formalization of an economy with varying risk across projects.

no need for auditing). Jointly, A2 and A3 create a demand for auditing in this model. A4 eliminates managerial signaling via ownership retention.[14] This assumption is widespread in the literature as a simplifying device, and it does not change qualitatively the nature of the conclusions (see, e.g., the references in note 13).

In the absence of credible communication, all insiders can be expected to assert that the project is Good and outsiders cannot rely on such statements (Akerlof 1970). In this setting, a fraction $1 - \theta$ of Bad projects will be so obviously unviable that if proposed, investors would never invest in them. Obviously no manager would propose such a project, or no such project, if proposed, would be funded. However a fraction θ of Bad projects cannot, *ex-ante*, be obviously distinguished from Good projects. The unconditional probability that a Bad project will appear to be Good is, therefore, $(1 - \gamma)\theta$. If investors cannot further screen between Good and Bad projects, then the expected return from, and the security price (P_U) of, an unaudited project is just the outsiders' expected return, i.e., the probability weighted payoff

$$P_U = \frac{\gamma}{\gamma + (1-\gamma)\theta}.1 + \frac{(1-\gamma)\theta}{\gamma + (1-\gamma)\theta}.0 = \frac{\gamma}{\gamma + (1-\gamma)\theta}.$$

As investors' assessment of θ increases (formally, as $\theta \rightarrow 1$), the security price decreases. If $P_U \leq r$ then investment will not occur and the market for securities will close down. This feature of the model permits us to interpret θ as investors' assessment of the *intrinsic quality* of the auditee. Intuitively, high (low) quality auditees are auditees for which investors' assessment of θ are low (high), i.e., closer to zero (closer to one).

[14] For instance, Datar, Feltham and Hughes study the role of audit quality in a signaling game where projects returns are distributed normally around some mean $m \in M = [\underline{m}, \overline{m}]$ with (common) variance σ^2. In other words, firms do not vary in risk in the cross-section. Proposition 7 of their study notes that the value of a higher quality audit increases in σ^2: As (all) projects become more risky, higher audit quality becomes more valuable. However, to obtain this result in the presence of the complications introduced by signaling, they must posit rather restrictive assumptions about the nature of the audit technology. Our model eliminates any role for managerial signaling in order to focus more closely on the fundamental *cross-sectional* tensions.

We now introduce auditing into the model. Suppose auditors can, at sufficiently low cost c, perfectly ascertain project type and some mix of *ex-post* penalties and reputation can ensure perfect screening of Bad projects.[15] In this case no Bad projects will ever be certified by the auditor and once a project has been certified, it is guaranteed to pay 1. When θ=0, prices of perfectly-audited projects (P_P) will equal outsiders' expected profits *net* of audit costs, i.e.,

$$P_P = \frac{\gamma}{\gamma}.1 - c = 1 - c .$$

If, however, the audit technology is imperfect and auditor wealth is miniscule relative to investor wealth at stake[16] then securities prices must take into account the likelihood of audit failure. Suppose that investors expect a reputable auditor to detect Bad projects with probability α. The (expected) audit failure rate is $1 - \alpha$ and the probability that a Bad project eludes the auditor's scrutiny is $(1 - \alpha)\theta$. The price (P_I) of such imperfectly audited projects will be

$$P_I = \frac{\gamma}{\gamma + (1-\gamma)(1-\alpha)\theta}.1 - c = \frac{\gamma}{\gamma + (1-\gamma)(1-\alpha)\theta} - c .$$

Since

$$\frac{\gamma}{\gamma + (1-\gamma)\theta} < \frac{\gamma}{\gamma + (1-\gamma)(1-\alpha)\theta} < 1$$

auditing increases security prices as long as

$$c < \frac{\gamma}{\gamma + (1-\gamma)(1-\alpha)\theta} - \frac{\gamma}{\gamma + (1-\gamma)\theta} .$$

[15] Auditor reputation serves as a bond that the auditor potentially forfeits (but investors may not be able to appropriate) when audit failures occur. Such bonding may reflect sunk costs of brand name development or of acquiring costly resources and developing business networks that have significantly lower value if the auditor is forced out of business. The fear of losing this bond keeps the auditor "honest" even absent any direct liability to investors (Dopuch and Simunic 1980, DeAngelo 1981).

[16] The total market capitalization of U.S. public companies is in excess of 13 trillion dollars, so this is a reasonable assumption.

Consider now the drop in security prices if investors' expectations about project quality drop (θ increases). For auditing to be more valuable on a lower quality auditee, an increase in audit quality, α, must be more valuable for a lower quality project, i.e., the increase in security prices as audit quality increases should be higher for projects with lower project quality (higher θ). Formally, this means that the second partial of P_I with respect to θ and α should be positive.

The partial derivative of (imperfectly audited) security prices in auditee quality, θ, is

$$\frac{\partial P_I}{\partial \theta} = -\frac{\gamma(1-\gamma)(1-\alpha)}{\left[\gamma + (1-\gamma)(1-\alpha)\theta\right]^2}$$

so an increase in θ, i.e., a *drop* in auditee quality (holding audit quality constant) leads to a *drop* in price.[17] The second partial with respect to an increase in audit quality (α),

$$\frac{\partial^2 P_I}{\partial \alpha \partial \theta} = \frac{\gamma(1-\gamma)\theta}{\left[\gamma + (1-\gamma)(1-\alpha)\theta\right]^2}\left[1 - \frac{2(1-\gamma)(1-\alpha)\theta}{\gamma + (1-\gamma)(1-a)\theta}\right]$$

is positive if and only if

$$\gamma > \gamma^* \equiv \frac{(1-\alpha)\theta}{1 + (1-\alpha)\theta}.$$

As $(1-\alpha)\theta$, the audit failure rate times the likelihood of a Bad project, i.e., the realized failure rate, goes to zero, so does γ^*. In other words, when audit failure rates are low, auditing is more valuable for lower quality projects even when the population proportion of good projects in the economy is very small.

Observation. Francis (2004) reviews the evidence on U.S. audit failure rates and concludes they are very low. If, for instance, $(1-\alpha)\theta = 0.01$, then $\gamma > 0.0099$ suffices for the proposition to hold.

[17] We work with partial derivatives since we are modeling the short-run effect of a shock to investors perception about the quality of past audits, i.e., the effect of the shock before audit quality and costs have had time to adjust to the shock itself.

We therefore also assume that the proportion of Good projects in the economy exceeds the threshold γ^*, i.e.,

(A6) $\gamma > \gamma^*$.

Collectively, assumptions A1-A6 suffice to yield the desired result, i.e.,

Proposition. Under assumptions A1-A6, a *decline* in auditor credibility (α) has a *more* negative impact on the security prices of *lower* quality projects (projects with higher θ).

Empirically, the observed pattern of security price changes in response to an impairment of auditor credibility can therefore be expected to look much like a flight to quality: the prices of low quality securities will drop more than prices of high quality securities and demand pressure may well lead some high quality auditee securities to experience positive returns.[18]

IV. Empirical Methods

Our approach to detecting spillovers is built on two maintained assumptions. First, when investors are presented with information that impairs auditor credibility, their faith in the auditor's ability or willingness to moderate management misrepresentations will decline, leading them to distrust reported accounting estimates (accruals) more than reported cash flow numbers. Investors will rebalance their investment portfolios, seeking out more desirable securities and exiting less desirable ones (the *flight-to-quality* effect). Second, this *flight-to-quality* will lead investors to disfavor auditees with *suspicious*[19] accruals and favor auditees with more positive cash flows.[20] Stocks that have "small" suspicious accruals and higher (more positive) cash flows

[18] Chaney and Philipich document that a substantial fraction of their sample experiences positive abnormal returns during the bad news windows they study.

[19] We focus on *suspicious* accruals, rather than on all accruals since total accruals reflect both the business fundamentals as well as the effects of management manipulations that investors now believe auditors were less able or willing to thwart. We define more precisely in Section IV the measures of suspicious accruals.

[20] The phenomenon we have in mind is captured by the aphorism "Accruals are opinions, cash is a fact."

(*high quality stocks*) are, therefore, expected to be more desirable than stocks with "large" suspicious accruals and low (more negative) cash flows (*low quality stocks*).

The overall approach we adopt to spillover detection involves four steps. These are depicted in Exhibit 1. In step 1 we identify using Lexis-Nexis dates on which major business news sources (e.g., *The New York Times*, *The Wall Street Journal*, *The Washington Post, Barrons*, *Business Week*) featured coverage of key events involving Andersen's Waste Management, Sunbeam and Enron audits. For an event to be selected, an article must (1) mention by name one or more of the three auditees as well as Andersen, (2) include at least one term stemming from the search root "account*" or "audit*" and (3) include disclosure of information about the auditee's financial statements or the conduct of the audit that was not already reported. This procedure yields 25 bad news events related to Andersen's audits of Waste Management, Sunbeam and Enron between November 1997 when accounting improprieties at Waste Management first surfaced and March 2002 when Andersen was indicted.

Table 1 lists the 25 events. During the period from November 1997 to April 2001 Waste Management and Sunbeam, two prominent Andersen auditees, disclosed fraudulent accounting in their audited financial statements and issued restatements. The Securities and Exchange Commission (SEC) launched investigations into Andersen's audits and Andersen was sued and settled stockholder lawsuits pertaining to these two engagements.

On May 15, 2001, the SEC sued the Andersen partner in charge of the Sunbeam engagement for civil fraud. On June 19, 2001, the SEC censured Andersen and imposed an unprecedented monetary fine on the firm for negligent audits of Waste Management. On June 26, 2002, the acting SEC chairman characterized Waste Management as Andersen's "smoking gun." Contemporary press coverage of these events unequivocally interpreted them as unfavorable to

Andersen's reputation as well as to that of auditors in general.[21] The tenor of this coverage (as well as the breadth of coverage of the sanctions in U.S. and international media) was highly critical of auditor conduct and, therefore, could well have impaired investors' trust in auditors and in audited financial statements.

The period from August 2001 to March 2002 covers Enron's failure and Andersen's mounting problems over its conduct as Enron's auditor. In August 2001, Enron CEO Skilling resigned in frustration over his inability to reverse a significant decline in the firm's stock price. Skilling's resignation was described as a "bombshell" and immediately raised concerns among analysts and in the business press about potential accounting irregularities at Enron (Harbrecht 2001).[22] The day after Skilling's resignation was announced, an Enron employee, Sharon Watkins, warned Enron Chairman Ken Lay that "we will implode in a wave of accounting scandals." She also suggested that Skilling's departure reflected his concerns that Enron's accounting manipulations had become unsustainable (Watkins 2001).

Public revelations about Enron's financial distress and Andersen's questionable conduct as Enron's auditor started to emerge by late October 2001.[23] By the end of November 2001, Enron-

[21] A June 20, 2001, Wall Street Journal article noted that "The Waste Management accounting scandal stands out for its size and breadth. … Within the SEC, the Arthur Andersen investigation became a centerpiece of the commission's aggressive campaign to demonstrate that conflicts of interest can be caused by consulting and other nonauditing services that numerous accounting firms now routinely offer audit clients." (Schroeder 2001). The SEC's director of enforcement was quoted in the New York Times on June 19, 2001, as saying that "Arthur Andersen and its partners failed to stand up to company management and thereby betrayed their ultimate allegiance to Waste Management's shareholders and the investing public. … Given the positions held by these partners and the duration and gravity of the misconduct, the firm itself must be held responsible for the false and misleading audit reports." The Times also noted that the SEC action was "a warning to all accounting firms that going along with management in allowing improper accounting presents major risks for a firm." (Norris 2001b) The Economist magazine titled its June 21, 2001, report of the sanctions as "Andersen's Fairy Tales: A $7m Fine From The SEC Again Raises Doubts About The Independence Of Auditors." A July 2, 2001, Business Week article titled "Andersen: Too Cozy With The Client" also noted that the sanctions amounted to "sounding a warning to the accounting profession." (Roman 2001).

[22] Skilling's abusive response to an analyst who questioned him about a lack of transparency in Enron's accounting and poor financial disclosure had been reported in the business press circa April 2004 (Reuters 2001).

[23] A November 12, 2001, article in the Wall Street Journal on Andersen's questionable approval of certain Enron transactions noted that "For Andersen, Enron joins a list including Waste Management Inc. and Sunbeam Corp. While the names of the clients may change, the issues remain largely the same." (Weil 2001).

related coverage of Andersen more often than not also referenced Sunbeam and Waste Management.[24] By December 2001, press coverage had begun to refer to Enron as the third in a series of major audit failures involving Andersen (Deener 2001).[25] Once Andersen admitted to potentially illegal document shredding on January 10, 2002, the firm's survival became questionable (Norris 2002a and b). On February 4, 2004, the Powers Committee (a special committee appointed by Enron's Board of Directors) released a report outlining in considerable detail Andersen's intimate involvement in Enron's misleading financial reports. This report was perceived as further impairing Andersen's credibility (Smith and Emshwiller 2002). On March 14, 2002, the U.S. Justice Department indicted Andersen for criminal obstruction of justice, effectively sealing the firm's eventual demise (Norris 2002c).

In steps 2 and 3 of Exhibit 1, we test whether, in each of the 25 bad-news windows, (1) mean cumulative abnormal returns (mean *CARs*), and (2) hedge cumulative abnormal returns (hedge *CARs*) are significantly negative. Hedge *CARs* are computed as the difference in returns between two portfolios: one that goes short in stocks with small absolute abnormal accruals and high cash flow from operations (high quality portfolio) and another that goes long in stocks with large absolute abnormal accruals and low cash flow from operations (low quality portfolio). A

[24] A November 23, 2001, New York Times article questioned Andersen's credibility in light of its recent history of accounting failures: "This has been the worst year ever for Arthur Andersen, the accounting firm that once deserved the title of conscience of the industry. The Securities and Exchange Commission filed civil fraud complaints against the Andersen partner who audited Sunbeam and against the firm itself in the Waste Management case. Now Enron has repudiated the financial statements that were certified by Arthur Andersen, in the process shaving more than half a billion dollars from the company's reported profits in recent years. All of which raises the question: Has Arthur Andersen become the black sheep of the accounting industry?" (Norris 2001a). Other press coverage during this period by and large adopted a similar tone.

[25] A December 24, 2001, Business Week article titled "Arthur Andersen: How Bad Will It Get?" played up this connection, noting that "Andersen's reputation has taken hits before. Prior to Berardino's appointment, the firm slogged through a string of high-profile scandals – most notably Sunbeam Corporation and Waste Management Inc." The article also quoted Alan Bromberg, a securities law professor, as saying "If this were the first big hassle Andersen had with regulators, I think most people would shrug it off. But this is the third time, so it'll be hard for people to ignore" (Weber et al. 2001). The "three strikes" argument became a persistent refrain thereafter and was a prominent feature of the Justice Department's arguments in its subsequent lawsuit against Andersen (Murphy 2002; Farrell 2002; Weil and Barrionuevo, 2002).

negative return to the hedge portfolio means that low quality auditees are, on average, penalized more than high quality auditees. In step 4 we investigate whether mean *CARs* are significantly negative for solvent auditees (auditees with Altman Z-scores in excess of 3) and distressed auditees (auditees with Altman Z-scores of less than 3).[26] We also estimate a series of regression models that shed light on the relation between auditee abnormal returns, accruals and cash flows.

V. Data and Measures

We obtain our data from the 2003 Compustat annual files and the 2003 CRSP daily files. Our sample for *each* of the 25 bad news events includes all firms with required Compustat and CRSP data in the year of the event. Our methods and data requirements permit us to use a more comprehensive sample than those used in prior published studies.[27] For the purpose of investigating market-wide spillovers, a more comprehensive sample is desirable not only for the extra coverage that it provides but also because it reduces the likelihood of smaller firms—which are more likely to need credible auditing—being eliminated from the analysis.[28]

We measure the market response to the event using the four-day cumulative abnormal returns during day 0 to day 3 [0, +3] where day 0 is the event date. We calculate daily abnormal returns during the 4-day event window in a 2-step procedure. First, we estimate, for each firm *i*, parameters $\alpha_{i,}$ β_{1i}, β_{2i}, β_{3i}, and β_{4i} from the following four-factor model:

$$R_{it} = \alpha_i + \beta_{1i} RM_t + \beta_{2i} SBM_t + \beta_{3i} HML_t + \beta_{4i} UMD_t + u_{it}, \quad (1)$$

[26] Changing the Z-score cutoff to 1.8 does not materially alter any of the principal inferences.
[27] For example, Chaney and Philipich (2002) study all domestic (US incorporated) auditees in the S&P1500 list for a sample size of about 1300 auditees. Callen and Morel (2003) examine 648 matched pairs of Andersen and non-Andersen auditees. Our sample by contrast comprises, on average, of over 4000 Big Five auditees and also includes non-Big Five clients in the analysis.
[28] In particular, the mean abnormal returns test to identify spillovers requires no further data beyond the computation of abnormal returns while the analysis of mean abnormal returns by auditor type (Andersen, other Big Five or non-Big-Five) requires only a match to Compustat to discover auditor identity. It is the data requirement for the cross-sectional association tests that reduces sample sizes quite considerably.

where R_{it} is the daily stock return for firm i in day t, RM_t is the daily value-weighted NYSE/AMEX market index return in day t, SBM_t and HML_t are the returns on the Fama and French (1992) factor-mimicking portfolios for size and book-to-market in day t, respectively, and UMD_t is Carhart's (1997) momentum factor in day t.[29] We estimate equation (1) for each stock in the 2003 CRSP universe over a 12-month period (with at least 150 daily returns) ending in the month immediately before the event we examine.[30] Second, we calculate daily abnormal returns in the event window as follows:

$$AR_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_{1i} RM_t - \hat{\beta}_{2i} SMB_t - \hat{\beta}_{3i} HML_t - \hat{\beta}_{4i} UMD_{4i},$$

where R_{it}, RM_t, SBM_t, HML_t and UMD_t are the daily stock returns of firm i, the daily value-weighted NYSE/AMEX market index returns, and the daily returns on the size, book-to-market and momentum factors, respectively, on day t in the event window and the parameters $\hat{\alpha}_i$, $\hat{\beta}_{1i}$, $\hat{\beta}_{2i}$, $\hat{\beta}_{3i}$ and $\hat{\beta}_{4i}$ are estimated in the first stage. For each of the 25 events, the four-day cumulative abnormal return (CAR) in the event window is computed as

$$CAR = \prod_{t=0}^{3} (AR_{it} + 1) - 1.$$

Our hedge returns tests require partitioning the sample on cash flows and accruals. The results reported in the paper are obtained by using the net cash from operating activities from the statement of cash flows (Compustat item #308) as the measure of cash flows (CFO) and the absolute value of abnormal accruals estimated using the Jones (1991) model, as modified by

[29] We obtain the daily SBM, HML and UMD factors from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[30] For example, when examining the market response to the event on November 14, 1997 when Waste Management announced new charges and a restatement of its third quarter operations, we estimate equation (1) using daily returns from November 1, 1996, to October 31, 1997. When examining the market response to the event on January 10, 2002, when Enron admitted document shredding, we estimate equation (1) using daily returns from January 1, 2001, to December 31, 2001.

Dechow, Sloan and Sweeney (1995), as the measure of accruals (|*AbAcc*|).[31] Results using alternative definitions of cash flows and accruals are discussed as part of the sensitivity analyses reported in Section VII.

For each of the 25 events, we compute *CFO* and |*AbAcc*| from annual financial statements for the most recent fiscal year ending *at least* three months *before* the month of the event. For example, Waste Management announced new charges and a restatement on November 14, 1997. For this event, we compute *CFO* and |*AbAcc*| for all auditees using accounting information from annual financial statements with fiscal year-ends before August 1997. More specifically, for auditees with fiscal year-end months in June and July (their 1997 fiscal year ends in June and July 1997, respectively), we use accounting information from fiscal year 1997 financial statements since these statements are likely to be publicly available by the event date. However, for auditees with fiscal year-end months between January and May (their 1997 fiscal year ends in January-May 1998) or with fiscal year-end months between August and December (their 1997 fiscal year ends in August-December 1997), we use fiscal year 1996 financial statements, since the fiscal 1997 statements for these auditees are either unavailable or may not be available to investors as of the event date.

---

[31] Our approach is similar to that of Xie (2001). Specifically, we initially estimate for each industry-year combination the following cross-sectional model:

$$ACCR/ATA = a_1[1/ATA] + a_2[\Delta REV/ATA] + a_3[PPE/ATA] + e_t, \quad (1a)$$

where *ACCR* is total accruals calculated from the statement of cash flows as earnings before extraordinary items (Compustat item #123) minus net cash flow from operating activities (Compustat item #308) following Hribar and Collins (2002), *ATA* is the average total assets (Compustat #6), ΔREV is the change in sales revenues (Compustat item #12), and *PPE* is gross property, plant, and equipment (Compustat item #7). We estimate equation (1a) in each year and 2-digit SIC code combination with at least 20 observations using all firms on the 2003 annual files. Following the prior literature, we winsorize all variables in equation (1a), including $1/ATA_{t-1}$, at their respective 1st and 99th percentiles. The coefficient estimates from equation (1a) are used to estimate normal accruals ($NAcc_t$) using Dechow, Sloan and Sweeney's (1995) modification of the Jones (1991) model:

$$NAcc = \hat{a}_1[1/ATA] + \hat{a}_2[(\Delta REV - \Delta AR)/ATA] + \hat{a}_3[PPE/ATA],$$

where ΔAR is the change in accounts receivable (Compustat #2). Abnormal accruals (*AbAcc*) are simply the difference between total accruals and normal accruals, i.e., $AbAcc = ACCR/ATA - NAcc$.

VI. Results

Table 2, panel A, reports the mean *CARs* for each of the 25 event windows. The sample size for each window ranges from about 4200 to about 5200 and, thus, is more comprehensive than the samples used in prior related research. Results for the full sample, Andersen auditees, Other Big Five auditees, all Big Five auditees and Non-Big Five auditees are presented in the columns titled All, AA, OB5, B5 and NB5, respectively. For expositional convenience, in the remainder of this section, we switch freely between the subgroup of auditees being referred to and the respective column titles.

Three features of Table 2, panel A, are particularly worthy of note. The first key feature is the considerable heterogeneity in investors' reactions across the twenty-five event windows. This heterogeneity implies that either some events were not interpreted as market-wide bad news and/or that other confounding events, a common threat to the validity of any event study, affected stock returns during the test windows. Absent a theory that predicts which news stories would be expected to create spillovers (and which stories would *not*), we follow prior studies in treating windows in which the overall market reaction is negative and significant as potential spillover windows. Twelve such potential spillover windows are identified in boldface in Table 2, panel A: the mean *CAR* in each window is significantly negative for All auditees at the 10% level or better, suggesting that the event, on average, affected all auditees adversely. To address the possibility that confounding events drive the reaction during such windows, we conduct both pooled and cross-sectional tests spanning multiple windows of interest.[32]

[32] While a pooled analysis of mean *CARs* eliminates systematic confounds with alternative explanations for the results, it still leaves open the possibility that the observed results arise from random unobserved events. Our cross sectional analysis, by contrast, exploits the theoretical prediction that during such windows one group (low quality auditees) is expected to be differentially treated than is another group (high quality auditees). Thus, relative to an analysis of mean *CARs* alone, a finding that, on average across multiple windows, investors penalize low quality auditees more than they penalize high quality auditees, provides stronger evidence of auditor credibility impairment.

Second, during the twelve potential spillover windows, mean abnormal *CARs* are generally negative and significant for every sub-group of auditees (and *never* positive and significant for *any* sub-group). These twelve windows can be classified into three types of events: four windows involve *regulatory actions* (*SBM Acct Inquiry*, *AA Partner Sued*, *WMI Penalty*, *Smoking Gun*), four involve the disclosure of *details* of accounting improprieties and/or auditor conduct (*Barrons, Berardino*, *Shredding* and *Powers*), while the remaining four involve *other* events such as auditee admissions of accounting fraud (*WMI Charge* and *SBM Restatement Needed*), executive turnover (*Skilling*) or financial failure (*Enron Bankruptcy*).

The third interesting feature of Table 2, panel A, is that eight of the twelve potential spillover events occurred before Enron's financial troubles became a matter of public record (pre-Enron events). Two of these eight events (*SBM Acct Inquiry* and *WMI penalty*) feature the most negative mean *CARs* for the market as a whole (-1.02% and -1.27%, respectively). In addition, the single most negative mean *CAR* for both Andersen (-1.37%) and Other Big Five (-1.24%) auditees occurs when the SEC penalizes Andersen over Waste Management (*WMI penalty*).

The mean *CARs* across the 25 windows reported at the bottom of panel A are significantly negative at the 5% level for All auditees, and for each sub-group of auditees. Formal tests of differences in the mean *CARs* between AA and OB5 auditees and AA and NB5 auditees show that in each case the differences are not significant (p-values of 0.37 and 0.15). Thus, irrespective of auditor identity, *all* auditees are, on average, penalized about equally. Overall, the results reported in panel A of Table 2 are consistent with pre-Enron events having impaired investor confidence in audits in general. This finding provides a potential explanation for why studies that focus only on Enron-related events find little evidence of significant spillovers.

Table 2, panel B, is organized along the same lines as panel A and reports cumulative mean *CARs* for all twelve potential spillover windows and for the three types of events discussed above. The first row of panel B shows that AA and OB5 auditees are penalized about equally (AA about -8.4%, OB5 about -6.7%). Consistent with the notion that non-Big-Five auditors are perceived to be of lower quality, NB5 auditees are penalized about twice as much as Big Five auditees (-15.7% versus -7%).

The disaggregated results reported in the next six lines of panel B show the following patterns. First, about 40% (46%) of the total penalty experienced by AA and OB5 (NB5) auditees occurs during the four *regulatory action* windows. In these four windows B5 auditees lose about 2.8% while NB5 auditees lose about 7%. Of the three types of events, the *regulatory action* windows are associated with the largest investor loss for Big Five auditors and their relative impact (% of total loss) is about the same for Big Five auditors as well. For Non-Big Five auditors, however, the largest investor loss occurs in the four *other* windows.[33]

Second, across the three types of events, the penalties experienced by AA and OB5 auditees vary in lock-step: about 42% (40%) of the total experienced by AA (OB5) auditees occurs during the four *regulatory action* windows. The corresponding percentages for the *details* and *other* windows are about: 33% and 25% for AA, 32% and 28% for OB5. Overall, these results suggest a relatively homogeneous market reaction to AA and OB5 auditees. By contrast, about 94% of the penalty on NB5 auditees occurs (about equally) during the *regulatory action* and *other* windows.

Third, during the four *details* windows, B5 auditees (clients of *higher quality* auditors) were penalized about twice as much as NB5 auditees (clients of *lower quality* auditors). The higher

[33] We do not conduct statistical tests for the mean *CARs* reported in Table 2, panel B, because the Fama-MacBeth (1973) p-values would be based on only four mean observations.

penalty on B5 auditees during these four windows suggests that when investors learned about the specific nature of accounting improprieties and auditor misjudgments that Andersen had been party to, they lost more confidence in the Big Five auditors as a group (AA lost about -2.8%, OB5 about -2.2%). Overall, the evidence in panel B suggests that each of the three types of events has the potential to create market-wide spillovers.

In panel C of Table 2 we examine the overall mean *CAR* (across all 25 windows) to each Big Five auditor's clients. We also compare the mean *CAR* for each OB5 auditor's clients to that for AA auditees. The mean *CARs* are significantly negative for the auditees of each Big Five auditor. Moreover, the mean *CAR* comparisons between clients of AA and each of the other Big Five auditors are insignificant (p-values range from 0.27 to 0.48). Taken together, these results suggest, on average over the 25 event windows, the presence of *potential* spillovers.

Table 3 presents the results of the cross-sectional analysis. The structure of Table 3 is similar to that of Table 2 except that the returns reported are mean hedge *CARs*.[34] Panel A reports mean hedge *CARs* by event window. Like Table 2, panel A, mean hedge *CARs* display considerable heterogeneity across the 25 windows, reflecting once again possible contamination arising from other, potentially confounding, events. Three windows, in particular, however, show negative and significant mean hedge *CARs* for both AA and OB5 auditees. These are the *AA Partner sued, WMI Penalty* and *Skilling* windows. It is worthy of note that all three windows predate the usual set of Enron-related events examined in prior research. The *WMI Penalty* window, in particular, stands out: it is the only window in which, for all key subgroups of auditees (AA, OB5 and NB5), mean abnormal returns are significantly negative in Table 2 *and* the mean hedge *CAR* is negative and significant for AA and OB5 auditees. The overall mean hedge *CAR* shown

[34] Recall that we have defined hedge *CARs* as the difference in mean *CARs* between the portfolio of low quality auditees and the portfolio of high quality auditees. A negative hedge *CAR*, thus, indicates that the market penalizes low quality auditees to a greater extent than it does high-quality auditees.

at the bottom of Table 3, panel A, is negative and significant (at the 6% level) for AA auditees only. This result suggests that over all 25 windows taken as a group, there is some evidence that AA's credibility was impaired, but not that of other auditors. However, as we show next, analyzing separately the twelve potential spillover windows identified in Table 2, reveals a dramatically different picture.

Table 3, panel B, reports mean hedge *CARs* for the twelve potential spillover windows and the thirteen non-spillover windows identified in Table 2. The first two rows of the panel display strikingly different patterns. Mean hedge *CARs* for non-spillover windows (top row) are positive and significant for OB5 auditees but insignificant for AA and NB5 auditees. However, across the twelve potential spillover windows (second row) mean hedge *CARs* are large (in excess of -1% per window) negative and significant for AA, OB5 and NB5 auditees. Further, mean hedge *CARs* for the potential spillover windows are *not* significantly different between AA (the impugned auditor) and either OB5 or NB5 auditees (insignificant p-values in the last row). Overall, the pattern of results in Table 3, panel B, greatly increases our confidence that, on average, the twelve potential spillover events impaired investors' confidence in audits in general (and not just in AA audits).

Table 4 reports, separately for financially distressed and non-distressed auditees, mean *CARs* for the thirteen non-spillover windows (panel A) and the twelve potential spillover windows (panel B).[35] The difference in the results between the two panels is striking. In panel A the mean *CARs* are non-negative and thus do not show any signs of impairment. Also, returns to AA auditees and to OB5 or NB5 auditees are statistically indistinguishable (p-values of 0.12 and 0.22 for OB5 auditees and 0.39 and 0.33 for NB5 auditees). In panel B, by contrast, mean *CARs* during the twelve potential spillover windows are significantly negative across all auditors for

[35] Auditees are classified as solvent (distressed) if their Altman Z-score is at least 3 (less than 3).

both solvent and distressed auditees. In addition, the mean *CAR* for AA auditees in panel B is not significantly different from that for OB5 auditees (p-values for solvent and distressed auditees of 0.37 and 0.19, respectively). Consistent with the proposition that the Big Five auditors are viewed as "higher quality" auditors than non-Big Five auditors, NB5 auditees are penalized more than AA auditees with p-values for solvent and distressed auditees of 0.05 and 0.04, respectively.

Finally, in panel C of Table 4, we report, separately for non-distressed and distressed auditees of each Big Five auditor, tests of mean *CARs* during the non-spillover and potential spillover windows. The mean *CARs* during the non-spillover windows are, with one exception (PWC, solvent auditees, -0.16%, significant at the 10% level), either positive or not significant. Mean *CARs* for AA auditees are, with one exception (PWC, as noted) not different from those for auditees of any other Big Five auditor. By contrast, during the potential spillover windows, mean *CARs* are, again with one exception (KPMG, -0.25%, not significant), all reliably negative. Finally, during the potential spillover windows, returns to AA auditees are, without exception, not different from those to auditees of any other Big Five auditor.

Overall, the results of Table 4 indicate no systematic differences in the pattern of investor responses for solvent and distressed auditees either during the twelve potential spillover windows or during the thirteen non-spillover windows. The reliably negative returns to *solvent* auditees—auditees relatively less likely to need auditor-provided insurance (auditees whose auditor-provided insurance option was more likely to be *out-of-the-money*)—during the twelve potential spillover windows provides strong support for the view that the events in question impaired auditor credibility. We also find no substantial differences between auditees of different Big Five auditors, suggesting that the events in question impaired the credibility of all auditors.

Table 5 reports for the twelve potential spillover windows, cumulative mean *CARs* for solvent (panel A) and distressed (panel B) auditees of every Big Five auditor and for non-Big Five auditees. Two patterns in the data are particularly noteworthy. First, the effects are substantial: solvent auditees lose, cumulatively, about 5% of their value during these twelve windows and financially distressed auditees about 10%. These amounts are well in excess of any reasonable amount of insurance cover any auditor could conceivably provide.[36] Third, by the end of the *WMI Penalty* window, losses to all Big Five auditors are quite substantial (and for solvent auditees about equal—between 2% and 2.5% for every Big Five auditor). These results further increase one's confidence that the events examined impaired the credibility of all auditors.

VII. Sensitivity Analysis

To investigate the sensitivity of the results reported thus far to measurement error, we also redid the analyses after forming portfolios based on (a) free cash flow in lieu of cash flow from operations,[37] and (b) signed abnormal accruals, accrual quality (Dechow and Dichev 2002) and, finally, total accruals. The use of free cash flow yielded very similar and, in hedge returns analyses, stronger results than the reported ones (based on *CFO*). The use of signed accruals yielded smaller excess returns than the ones reported but did not otherwise alter our principal conclusions. The use of Dechow and Dichev's accrual quality measure yielded very similar results to the ones reported in terms of both magnitude and significance. Thus, the use of

[36] This finding must be interpreted in light of our earlier discussion that empirically, distinguishing between insurance and assurance impairment in our setting may be problematic (cf. p. 8).

[37] Free cash flow (*FCF*) is calculated, following Penman and Sougiannis (1998), as net cash flow from operating activities (Compustat item #308) plus interest paid (#315) plus net cash flow from investing activities (#311) minus capitalized interest (#147). If interest paid is missing but interest expense (Compustat #15) is not, then interest expense is used in lieu of interest paid. When both interest paid and interest expense are missing, interest paid is set to zero. Capitalized interest is set to zero when missing. We scale our measures of cash flows by average total assets (Compustat item #6).

alternative measures of accruals and cash flows does not affect our principal conclusions.[38] The use of total accruals, however, did not give us any clearly interpretable results. Thus the market reaction seems to be associated with the magnitude of abnormal or suspicious accruals and not with the magnitudes of total accruals.

We also conducted three additional analyses worthy of reporting. First, wherever possible, we selected for each of our 25 event windows a 4-day *non-event* window within the same calendar month ending at least two days before or beginning at least two days after the event window. During these matched non-event windows, the overall pattern of results is strikingly different from those for the 25 bad news windows. Across all non-event windows, mean *CARs* and hedge *CARs* are almost never negative and never both negative at the same time. Across all control windows, for the overall sample as well as by auditor type, the pattern of abnormal returns is *never* suggestive of auditor credibility impairment.

In a second sensitivity analysis we restricted the sample to smaller auditees (auditees with market capitalization of less than $1 billion). Smaller auditees are less likely to be widely followed by analysts and market participants are likely to have less current information about these auditees available from non-accounting channels. Consequently, investors may rely on audited financial statements to a greater extent for smaller auditees than for larger ones. Results from the restricted analyses are so similar to those already discussed both in pattern and magnitude that no new insights emerge.

In a third series of sensitivity analyses we estimated several regression models, four of which are reported in Table 6.[39] These regression models examine the cross-sectional association at the

[38] Results of these and other sensitivity analyses are available upon request.
[39] The results reported in Table 6 are based on the full sample including Andersen auditees. Repeating the analyses reported in Table 6 after restricting the sample to Other Big Five auditees (non-Andersen Big Five auditees) does not materially alter any of our key inferences.

firm level between abnormal returns and auditee characteristics. All standard errors (and therefore significance levels) are based on clustered standard errors corrected for cross-sectional correlation of abnormal returns within windows.[40] Model 1 of Table 6 shows that over all 25 events, neither accruals nor cash flows significantly explains the cross-sectional variation in abnormal returns: the intercept is not significantly different from zero, nor is the coefficient of cashflow while that of |*AbAcc*| is significant at the 10% level. There is, therefore, at best a weak negative association between the magnitude of accruals and abnormal returns.[41]

In Model 2 we decompose the 25 windows into the twelve spillover windows ($S = 1$) and thirteen non-spillover windows ($S = 0$). A radically different picture from Model 1 emerges. We find, as hypothesized, that (1) consistent with a market-wide penalty for all firms, the intercept for the S = 1 windows is negative (-0.007) and highly significant (at the 1% level of significance or better), and (2) consistent with investors being concerned about the quality of accounting estimates, the coefficient on |*AbAcc*| is negative (-0.02) and significant (at the 5% level) and

[40] We also estimated each of these models using the Schipper-Sefcik-Thompson approach outlined in Schipper and Thompson (1985) and Sefcik and Thompson (1986) and found virtually identical results (available on request). This approach attempts to correct for cross-sectional correlation of returns during event and non-event windows by estimating OLS regression coefficients in a first stage regression and, then, in a second-stage procedure regressing the estimated coefficients on a dummy variable that takes the value one during the windows of interest. This procedure is formally identical to testing whether the mean of the first-stage coefficients estimated from event windows differs significantly from the mean of the coefficients estimated during non-event windows. This represents another application of the idea behind the Fama-MacBeth (1973) test we use in the portfolio tests. However, since the Schipper-Sefcik-Thompson approach throws away *all* information about the pattern of cross-sectional variation, its finite sample properties are problematic (MacKinlay 1987 and 1997, Campbell, Lo and MacKinlay 1997: 167). A comparison of standard errors of the coefficient estimates from the clustered OLS estimation and the Schipper-Sefcik-Thomson estimation shows that neither produces uniformly larger (more conservative) estimates. The exact number of cases in which the standard errors produced by one method or the other are larger varies with the model, the average is about 50% for each approach. However we did not find material differences in the significance levels using either approach and therefore report the results of the fully-nested-OLS-with-clustered-standard-error procedure since this procedure throws away less information and may be easier to implement in most modern econometric packages.

[41] Interestingly, however, when we use signed accruals, we find no relationship between signed accruals and abnormal returns. This suggests that the accrual mispricing correction explanation in which investors are thought to correct the initial mispricing of accruals (Sloan 1995, Xie 2001) when forced to pay attention to the quality of accounting, say due to bad news about the auditor, is unlikely to explain our findings. The accrual mispricing story is that investors initially would have over-valued positive accruals (and under-valued negative accruals). The correction would then take the form of a penalty for positive accruals firms and reward for negative accruals firms and, on average no significant penalty associated with the *magnitude* of accruals *per se*. We, however, find that the reaction *is* associated in the cross-section with the *magnitude* of the accruals and *not* with the signed accruals.

consistent with investors seeking out safer more viable firms during the spillover windows, the coefficient for cashflow is positive (0.031) and significant (at the 1% level or better). Model 3 shows that the overall tenor of these findings is not materially altered by introducing additional controls for (a) auditee membership of the manufacturing industry, (b) auditee industry R&D intensity and (c) auditee financial distress scores. Finally, in Model 4 we replace the continuous measures of accruals and cash flows by an indicator variable that takes the value one if the auditee is a high quality auditee, 0 otherwise. This model reconciles the regression approach with the portfolio-based findings in the previous tables by showing that relative to the other eight portfolios, and after controlling for other factors including auditee financial distress, high quality auditees are rewarded (or penalized less): the coefficient of *Port*S* (*Port* = 1 for high-quality auditees, 0 for all others) is positive and highly significant: on average over the 12 windows, high-quality auditees earn about 0.6% higher returns, a total of about 7% in all windows. Overall, the regression analysis complements our portfolio analysis and increases our confidence in inferring auditor credibility impairment spillovers.

VIII. Summary and Conclusions

We analyze abnormal stock returns around twenty-five key events involving Andersen's questionable audits of Waste Management, Sunbeam, and Enron, to test whether these events resulted in significant intra- and inter-firm (auditor credibility impairment) spillovers. In twelve of the twenty-five windows, we detect substantial externalities to distressed as well as solvent clients of other auditors. Our inferences are robust to several sensitivity analyses such as the use of alternative measures of financial reporting quality and financial health and to calibration

against the pattern of stock returns observed during twenty-five matched *non-event* windows (i.e., we do not observe similar results during these non-event windows as the event windows).

Our results cast new light on the findings of prior studies of auditor reputation that have used key events related to the implosion of Enron and the demise of Andersen to detect inter-firm spillovers. We present evidence that substantial credibility impairments may already have occurred several months before Enron's financial troubles and shenanigans became a matter of public knowledge. These findings can explain why prior studies of Enron-related events may have found it hard to document substantial evidence of inter-firm spillovers resulting from those events. In particular, we find that news of SEC litigation against Andersen's Sunbeam engagement partner, of SEC sanctions of Andersen for its Waste Management audit and the resignation of Enron CEO Jeff Skilling triggered declines in investor confidence in audited financial statements and, thus, in auditors.

Our findings also cast fresh light on two recent studies of auditor-client realignments in the wake of Andersen's demise. Barton (2005) notes that only 5% of Andersen auditees switched to other auditors before Andersen was indicted. He conjectures that "This low initial defection rate could reflect among other things, a general lack of concern about auditor reputation, high auditor-switching costs, and uncertainty about Andersen's culpability in the Enron scandal or its ability to continue in existence." (Barton, *ibid* 554). If, however, as our results suggest, antecedent events had already significantly impacted other leading auditors' credibility, the low rate of early defections from Andersen might simply reflect relatively low *benefits* to Andersen auditees of an early defection.

In a related study, Krishnamurthy et al. (2006, p. 468) find that Andersen auditees experienced a significant negative abnormal returns on the indictment date. The penalty reverses

if the auditee quickly finds a replacement Big Four auditor but intensifies if the auditee does not. Our finding that substantial credibility impairment appears to already have happened well before the indictment date (and our failure to find any credibility impairment effects on the indictment date) is consistent with the conjecture that the negative reaction to Andersen auditees around the indictment date reflects investor expectations regarding auditor replacement costs.

Overall, we find that bad news about key events related to Andersen's Waste Management, Sunbeam and Enron audits imposes significant externalities on both its own auditees and those of other auditors, i.e., on the entire market. The research methods in this study may be useful for future research into spillovers in audit markets. Last, but not the least, our findings suggest interesting opportunities for future research. Prior research provides little guidance as to events likely to cause market-wide spillovers. Our finding that investors react strongly to reports of regulatory sanctions and to revelations of inadequate auditing of financial statements provides a starting point for future research into the causes of market-wide spillovers. Future research might profitably investigate whether other types of events such as controversies about tax shelters or expense billing or violations of SEC norms for investments in auditees also result in spillovers. Additional research into the dynamics of auditor-client realignments in the wake of Andersen's demise and their effect on capacity reallocation, audit firm clienteles and audit fees also present interesting avenues of future research.

References

Akerlof, G.A. 1970. The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84: 488-500.

Altman, E. 1983. *Corporate Financial Distress.* Wiley, New York.

Antle, R. 1984. The auditor as an economic agent. *Journal of Accounting Research* 20: 503-527.

Antle, R. 1984. Auditor independence. *Journal of Accounting Research* 22: 1-20.

Asthana, S., S. Balsam and J. Krishnan. 2003. Audit firm reputation and client stock price reactions: Evidence from the Enron experience. Working Paper, Temple University.

Baber, W.R.; K.R. Kumar; and T. Verghese. 1995. Client security price reactions to the Laventhol and Horwath bankruptcy. *Journal of Accounting Research* 33: 385-395.

Barton, J. 2005. Who Cares about Auditor Reputation. *Contemporary Accounting Research* 22: 549-86.

Buckless, F., and R Peace. 1993. The influence of the source of professional standards on juror decision making. *The Accounting Review* 68(1): 164-175.

Cahan, S., D. Emanuel, and J. Sun. 2005. Shredded reputation: Did Andersen's Non-US clients suffer and why? Working paper. University of Auckland.

Callen, J.L., and M. Morel. 2003. The Enron-Andersen debacle: Do equity markets react to auditor reputation? *Finance Letters*: 1-5.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52: 57-82.

Chaney, P.K., and K.L. Philipich. 2002. Shredded reputation: The cost of audit failure. *Journal of Accounting Research*, 40: 1221-1245.

Datar, S.M., G.A. Feltham and J.S. Hughes. 1991. The role of audits and audit quality in valuing new issues. *Journal of Accounting and Economics* 14: 3-49.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3: 183-199.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review,* 77: 71-114.

Dechow, P.M., R.G. Sloan, and A.P. Sweeney. 1995. Detecting earnings management. *The Accounting Review,* 70: 193-225.

Deener, W. 2001. Arthur Andersen may take years to recover from Enron debacle, analysts say. *The Dallas Morning News*, December 7.

Dopuch, N., and D. Simunic. 1980. Competition in auditing: an assessment. Symposium on Auditing Research IV. University of Illinois, Urbana-Champaign, pp. 401-450.

Dye, R.A. 1993. Auditing standards, legal liability and auditor wealth. *Journal of Political Economy*, 101 (5): 887-914.

Economist, The. 2001. Andersen's fairy tales: A $7m fine from the SEC again raises doubts about the independence of auditors. June 21.

Fama, E.F., and J.D. MacBeth. 1973. Risk, return and equilibrium: Empirical tests/ *The Journal of Political Economy* 81: 607-636.

Fama, E.F., and K.R. French. 1992. The cross-section of expected stock returns. *Journal of Finance*, 74: 427-465.

Farrell, G. 2002. Jury will hear of Andersen's past scandals. *USA Today*, May 8.

Firth, M. 1990. Auditor reputation - The impact of critical reports issued by government inspectors. *Rand Journal of Economics* 21: 374-387.

Francis J. 2004. What do we know about audit quality? *The British Accounting Review* 34(4): 345-368.

Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* 19: 208-231.

Harbrecht, D. 2001. Enron's Ken Lay: "There's No Other Shoe to Fall." *Business Week* August 24, 2001.

Hillegeist, S.A. Financial reporting and auditing under alternative damage apportionment rules. *The Accounting Review*, 74 (3): 347-379.

Hribar, P., and D. W. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research*, 40: 105-134.

Jennings, M., D. Kneer and P Reckers. 1993. The significance of audit decision aids and pre-case jurists' attitudes on perceptions of audit firm culpability and liability. *Contemporary Accounting Research* 9(2): 489-507.

Johnson, V., I. Khurana, and J. Reynolds. 2002. Audit-firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19 (4): 637-660.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research,* 29: 193-228.

Klein, A. 2002. Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics* 33 (3): 375-400.

Krishnamurthy, S., J. Zhou, and N. Zhou. 2006. Auditor reputation, auditor independence, and the stock-market impact of Andersen's indictment on its client firms. *Contemporary Accounting Research* 23: 465-490.

Latham, C.K., and M Linville. 1998. A review of the literature in audit litigation. *Journal of Accounting Literature* 17: 175-213.

Leland, H.E., and D.H. Pyle. 1977. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance* 32: 371-387.

MacKinlay, A.C. 1987. On multivariate tests of the CAPM. *Journal of Financial Economics* 18: 341-372.

MacKinlay, A.C. 1997. Event studies in economics and finance. *Journal of Economic Literature* 35: 1-39.

Menon, K., and D.D. Williams. 1994. The insurance hypothesis and market prices. *The Accounting Review* April 69: 327-342.

Miller, G. 2006. The press as a watchdog for accounting fraud. Journal of Accounting Research 44: 1001-1033.

Murphy, K. 2002. Down home at the Andersen trial. *Business Week*, May 9.

Myers, J., L. Myers, and T. Omer. 2003. Exploring the term of auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (3): 779-799.

Norris, F. 2001a. From Sunbeam to Enron, Andersen's reputation suffers. *The New York Times*, November 23: C1.

Norris, F. 2001b. A top five accounting firm to pay $7 million over fraud. *The New York Times*, June 19.

Norris, F. 2002a. Did Enron's auditors think they had something to hide? *The new York Times*, January 11.

Norris, F. 2002b. For Andersen and Enron, the questions just keep coming. *The New York Times*, January 16.

Norris, F. 2002c. Execution before trial for Andersen. *The New York Times*, March 15.

Pae, S. and S-W. Yoo. 2001. Strategic interaction in auditing: An analysis of auditors' legal liability, internal control system quality, and audit effort. *The Accounting Review*, 76 (3): 333-356.

Palmrose, Z. 1987. Litigation and independent auditors: The role of business failures and management fraud. *Auditing: A Journal of Practice and Theory* : 90-103.

Penman, S., and T. Sougiannis. 1998. A Comparison of dividend, cash flow, and earnings approaches to equity valuation. *Contemporary Accounting Research,* 15: 343-383.

Reuters (New Services). 2001. Skilling, analyst verbally butt heads. April 18, 2001.

Roman, M. 2001 Andersen: Too cozy with the client. *Business Week*, July 2: 44.

Schroeder, M. 2001. SEC fines Arthur Andersen in fraud case, Big 5 firm to pay $7 million after inquiry of audits for Waste Management. *The Wall Street Journal*, June 20: A3.

Schwartz, R. 1997. Legal regimes, audit quality and investment. *The Accounting Review* 72 (3): 385-406.

Simunic, D. 2003. Audit Quality and Audit Firm Size: Revisited. University of British Columbia.

Smith, R. and J. Emshwiller. 2002. Internal probe of Enron finds wide-ranging abuses -- Unanswered in board report are some big questions regarding legal liability. *The Wall Street Journal* March 04.

Watkins, S. 2001. Untitled communication. Archived at http://energycommerce.house.gov/reparchives/107/hearings/02072002Hearing485/tab17.pdf.

Weber, J., D. Little, D. Henry, and L. Lavelle. 2001. Arthur Andersen: How bad will it get? *Business Week*, December 24.

Weil, J. 2001. Basic principle of accounting tripped Enron. *The Wall Street Journal*, November 12: C1.

Weil, J., and A. Barrionuevo. 2002. Andersen hits major setback one day into criminal trial – Federal judge rules government can introduce past misconduct as evidence of obstruction motive. *The Wall Street Journal*, May 8, C1.

Xie, H. 2001. "The mispricing of abnormal accruals." *The Accounting Review,* 76: 357-373.

Exhibit 1

Procedure to identify credibility impairment spillovers

Step Number and objective	Procedure or Test
1. Identify unfavorable press coverage about Andersen's Waste Management, Sunbeam and Enron audits.	Search Lexis-Nexis database for press coverage
2. Examine investor perceptions of the news as reflected in stock price reactions to the events identified in step 1.	Test whether the mean cumulative abnormal returns (mean *CARs*) are negative for clients of the impugned and non-impugned auditors. If mean *CARs* <0 then there is evidence of potential spillover.
3. Examine hedge return sign and significance to determine whether reported accruals and cash flows were systematically repriced during the event windows.	Test whether mean hedge *CARs* (CAR for low quality auditees *minus* CAR for high quality auditees) are negative for clients of the impugned and non-impugned auditors. If mean hedge *CARs* are negative for both groups then there is evidence of auditor credibility impairment spillover.
4. Examine return sign and significance for both distressed and non-distressed auditees.	Test whether mean *CARs* are negative not only for distressed but also for non-distressed clients of the impugned and non-impugned auditors. If that is the case then there is evidence of assurance impairment (and not just insurance impairment)

Table 1
Bad News Events Regarding Andersen and Its Audits during 1997-2002

These 25 bad news events are identified using the following procedure. First, we search, using Lexis-Nexis, dates on which major business news sources featured stories involving Andersen's Waste Management, Sunbeam and Enron audits. To be selected, a story must (1) mention by name one or more of the three auditees as well as Andersen, (2) include at least one term stemming from the search root "account*" or "audit*" and (3) include disclosure of information about the auditee's financial statements or the conduct of the audit that was not already known. Second, we also identified and included key dates on which controversial accounting charges or need for restatements were first announced as well as dates on which amended financial statements were filed or released. Finally, we included all dates identified in prior research as being potentially relevant dates on which auditor credibility may have been impaired due to Enron-related announcements.

Window	Date	Event
1 (WMI Charge)	11/14/1997	Waste Management announces charges
2 (WMI Restatement)	02/24/1998	Waste Management files restated 10-Ks
3 (WMI Audit Review)	04/06/1998	SEC announces review of Andersen's Waste Management audit
4 (Barrons)	06/08/1998	Barrons' questions Sunbeam accounting
5 (SMB CEO Fired)	06/16/1998	Sunbeam fires CEO Dunlap
6 (SBM Acct Inquiry)	06/22/1998	SEC announces inquiry into Sunbeam's accounting
7 (SBM Restatement Possible)	06/30/1998	Sunbeam announces possible restatement
8 (SBM Probe)	07/14/1998	SEC announces Sunbeam probe
9 (SBM Restatement needed)	08/06/1998	Sunbeam announces material restatement will be needed
10 (SBM Restatement)	10/20/1998	Sunbeam files restated financial statements
11 (WMI Settlement)	12/08/1998	Andersen settles Waste Management class action lawsuit for $220M
12 (WMI Probe)	08/25/2000	SEC probes Andersen's conflicts on Waste Management audits
13 (SBM Settlement)	04/26/2001	Andersen settles Sunbeam class action lawsuit for $110M
14 (AA Partner sued)	05/15/2001	SEC sues Andersen engagement partner over Sunbeam audit
15 (WMI Penalty)	06/19/2001	SEC censures Andersen, fines $7 million for Waste Management
16 (Smoking Gun)	06/26/2001	SEC chief makes "smoking gun" remark
17 (Skilling)	08/15/2001	Enron CEO Skilling resigns
18 (Enron Loss)	10/16/2001	Enron announces loss and charges
19 (Enron Investigation)	10/24/2001	SEC announces Enron investigation
20 (Enron Restatement)	11/08/2001	Enron restates, *Wall Street Journal* impugns Andersen's audits
21 (Enron Bankruptcy)	12/03/2001	Enron Bankruptcy filing
22 (Berardino)	12/12/2001	Andersen CEO Berardino testifies before Congress
23 (Shredding)	01/10/2002	Andersen shredding admission
24 (Powers)	02/04/2002	Powers report issued
25 (Indictment)	03/14/2002	Andersen indicted

Table 2
Mean Abnormal Returns by Event Window and Auditor Type and Comparisons of Returns Between Andersen and Other Auditors' Auditees

Abnormal returns are four-day [0, +3] cumulative abnormal returns (*CARs*) around each of the 25 news events described in the first column (and in Table 1), computed using a four-factor model as in Carhart (1997). The sample comprises of all domestic auditees in the CRSP-Compustat universe for which relevant data items are available. N indicates the sample size in each window. All = All auditors, B5 = All Big Five auditors (AA + OB5), AA = Arthur Andersen, OB5=Other Big Five (DT = Deloitte and Touche, EY = Ernst & Young, KPMG = KPMG, PwC = PricewaterhouseCoopers), and NB5 = non-Big-Five auditors. The windows in boldface are identified as *potential spillover* windows because the mean *CAR* in each of these twelve windows is significantly negative for All at the 10% level or better. p-values for the overall mean *CAR* and for the differences between mean *CARs* of AA, OB5 and NB5 in panel A as well as p-values in Panel C are based on *Fama-MacBeth* (1973) *t* statistics of the mean *CARs* across the 25 event windows.
* (**, ***) denotes significantly different from zero at 10% (5%, 1%) levels.

Panel A: Window-by-window mean *CARs*

Event	N	Type of Auditee				
		All	AA	OB5	B5	nB5
1 (WMI Charge)	**5217**	**-0.0068 *****	**-0.0066 *****	**-0.0049 *****	**-0.0052 *****	**-0.0181 *****
2 (WMI Restatement)	5093	-0.0006	-0.0022	0.0002	-0.0003	-0.0027
3 (WMI Audit Review)	5245	0.0033 ***	0.0016	0.0037 ***	0.0033 ***	0.0039
4 (Barrons)	**5237**	**-0.0015 ***	**-0.0042 ***	**-0.0012**	**-0.0018 ***	**0.0007**
5 (SMB CEO Fired)	5226	-0.001	0.0028	-0.0019	-0.0009	-0.0016
6 (SBM Acct Inquiry)	**5223**	**-0.0102 *****	**-0.0083 *****	**-0.0084 *****	**-0.0084 *****	**-0.023 *****
7 (SBM Restatement Possible)	5191	-0.0008	0.0007	0.0001	0.0002	-0.0083 *
8 (SBM Probe)	5239	-0.0011	-0.0031	0.0002	-0.0005	-0.0052
9 (SBM Restatement needed)	**5238**	**-0.006 *****	**-0.0069 *****	**-0.0034 ****	**-0.0042 *****	**-0.0192 *****
10 (SBM Restatement)	5110	0.0027	0.0081 *	0.0024	0.0036	-0.0038
11 (WMI Settlement)	5046	0.0009	-0.004	0.0027	0.0013	-0.0021
12 (WMI Probe)	4752	0.008 ***	0.0131 ***	0.0065 ***	0.0079 ***	0.0085 *
13 (SBM Settlement)	4611	0.0052 ***	-0.0017	0.0064 ***	0.0047 **	0.0083 *
14 (AA Partner sued)	**4598**	**-0.0032 ****	**-0.0031**	**-0.0023**	**-0.0025**	**-0.0079 ***
15 (WMI Penalty)	**4566**	**-0.0127 *****	**-0.0137 *****	**-0.0124 *****	**-0.0127 *****	**-0.0131 *****
16 (Smoking Gun)	**4552**	**-0.0076 *****	**-0.01 *****	**-0.0033**	**-0.0047 ****	**-0.0272 *****
17 (Skilling)	**4497**	**-0.0079 *****	**-0.0034**	**-0.0088 *****	**-0.0077 *****	**-0.0096 ****
18 (Enron Loss)	4420	0.002	0.0019	-0.0003	0.0001	0.015 *
19 (Enron Investigation)	4415	0.0053 ***	0.0099 ***	0.004 *	0.0052 ***	0.0058
20 (Enron Restatement)	4392	0.0003	0.0065	-0.0014	0.0002	0.001
21 (Enron Bankruptcy)	**4351**	**-0.0057 *****	**-0.0046**	**-0.0018**	**-0.0023**	**-0.0287 *****
22 (Berardino)	**4344**	**-0.0085 *****	**-0.0096 *****	**-0.0077 *****	**-0.0081 *****	**-0.0113 ***
23 (Shredding)	**4317**	**-0.0033 *****	**-0.0083 *****	**-0.0038 *****	**-0.0047 *****	**0.0067**
24 (Powers)	**4290**	**-0.0079 *****	**-0.0057 ****	**-0.0089 *****	**-0.0083 *****	**-0.0057**
25 (Indictment)	4230	0.0018	-0.0049 *	0.002	0.0006	0.0102 **
Overall Mean *CAR*		-0.0022 **	-0.0022 **	-0.0017 **	-0.0018 **	-0.0051 **
p-value		0.02	0.05	0.05	0.04	0.02
p-value for difference from AA				0.37		0.15

Table 2 (continued)

Panel B. Cumulative Mean *CARs* for twelve potential spillover windows by event type

Event	Type of auditee				
	All	AA	OB5	B5	nB5
Total (all twelve windows)	-8.13%	-8.43%	-6.69%	-7.05%	-15.63%
Four "regulatory action" windows	-3.37%	-3.51%	-2.64%	-2.82%	-7.12%
% of total	41.40%	41.70%	39.51%	40.05%	45.56%
Four "details" windows	-2.12%	-2.77%	-2.16%	-2.29%	-0.96%
% of total	26.12%	32.86%	32.34%	32.47%	6.11%
Four "other" windows	-2.64%	-2.14%	-1.88%	-1.94%	-7.56%
% of total	32.48%	25.45%	28.15%	27.48%	48.33%

Panel C. Mean *CAR* comparisons between AA and OB5 auditees - all 25 windows.

	AA	DT	EY	KPMG	PWC
Mean abnormal return per window	-0.0022 ***	-0.0012 **	-0.0015 **	-0.0017 **	-0.0021 ***
p-value for difference from AA		0.27	0.35	0.41	0.48

Table 3
Mean Abnormal *Hedge* Returns by Event Window and Auditor Type and Comparisons of Returns Between Andersen and Other Auditors' Auditees

The hedge is formed by sorting auditees (independently) into terciles based on the absolute value of abnormal accruals (|*AbAcc*|) and cash from operations (*CFO*). Mean abnormal *hedge* returns (hedge *CARs*) are computed by comparing, for each event window, cumulative abnormal returns to the portfolio of auditees with high |*AbAcc*| and low *CFO* (low quality auditees) with corresponding cumulative abnormal returns to the portfolio of auditees with low |AbAcc| and high *CFO* (high quality auditees). Abnormal returns are four-day [0, +3] cumulative abnormal returns (*CARs*) around news events described in Table 1, computed using a four-factor model as in Carhart (1997). The sample comprises of all domestic auditees in the CRSP-Compustat universe for which relevant data items are available. N indicates the sample size in each window. All = All auditors, B5 = All Big Five auditors (AA + OB5), AA = Arthur Andersen, OB5 = Other Big Five (DT= Deloitte and Touche, EY = Ernst & Young, KPMG = KPMG, PwC = PricewaterhouseCoopers), and NB5 = non-Big-Five auditors. During the windows in boldface, mean hedge *CARs* are significantly negative for both AA and OB5 at the 10% level or better. p-values for the overall mean hedge *CAR* and for the differences between mean hedge *CARs* of AA, OB5 and NB5 are based on *Fama-MacBeth* (1973) *t* statistics of the mean hedge *CARs* across the 25 event windows (panel A), the 12 potential spillover windows or the 13 non-spillover windows (panel B) identified in Table 2.
* (**, ***) denotes significantly different from zero at 10% (5%, 1%) levels.

Panel A. Mean Hedge *CARs* by event window

Event	N	Type of Auditee				
		All	AA	OB5	B5	nB5
1 (WMI Charge)	1577	-0.0097 **	-0.01 *	-0.005	-0.0061 *	-0.0202
2 (WMI Restatement)	1555	-0.0024	-0.0166 **	0.0028	-0.0017	-0.0063
3 (WMI Audit Review)	1589	0.0188 ***	0.0193	0.0202 ***	0.02 ***	0.0226
4 (Barrons)	1593	0.0097 **	0.0027	0.0142 ***	0.0118 ***	0.0003
5 (SMB CEO Fired)	1590	0.0002	-0.0009	0.0027	0.0016	-0.0035
6 (SBM Acct Inquiry)	1589	-0.0126 ***	-0.0056	-0.008 *	-0.0075 **	-0.0237 *
7 (SBM Restatement Possible)	1579	-0.0049	-0.0048	-0.0055	-0.0052	0.0077
8 (SBM Probe)	1603	0.0063	0.0036	0.0095 *	0.0083 *	0.0023
9 (SBM Restatement needed)	1607	-0.0106 **	0.0015	-0.013 **	-0.0099 **	-0.005
10 (SBM Restatement)	1573	0.0176 **	0.0241	0.0164 *	0.018 **	0.0241
11 (WMI Settlement)	1552	-0.0027	-0.016 *	0.0021	-0.0018	-0.0059
12 (WMI Probe)	1420	0.0097 **	-0.0017	0.0108 *	0.0083 *	0.0073
13 (SBM Settlement)	1457	0.0004	0.003	-0.001	-0.0001	-0.0064
14 (AA Partner sued)	**1456**	**-0.02 *****	**-0.0257 ****	**-0.0228 *****	**-0.0234 *****	**-0.0128**
15 (WMI Penalty)	**1438**	**-0.0123 *****	**-0.024 ****	**-0.0122 ****	**-0.0144 *****	**-0.0073**
16 (Smoking Gun)	1435	-0.0195 ***	-0.0339 **	-0.011	-0.0154 **	-0.027
17 (Skilling)	**1419**	**-0.0292 *****	**-0.03 *****	**-0.0303 *****	**-0.0305 *****	**-0.0123**
18 (Enron Loss)	1382	0.0157 **	0.0131	0.0117 *	0.012 **	0.0312
19 (Enron Investigation)	1380	-0.0012	0.015	-0.0013	0.0015	-0.0254
20 (Enron Restatement)	1377	0.011 **	0.0054	0.0122 **	0.0109 **	0.0133
21 (Enron Bankruptcy)	1362	-0.0145 **	-0.0087	-0.0115 **	-0.0109 *	-0.0238
22 (Berardino)	1359	-0.0003	0.009	0.0013	0.0027	-0.0187
23 (Shredding)	1347	-0.0007	-0.0175 *	-0.0014	-0.0042	0.0194
24 (Powers)	1340	-0.026 ***	-0.0136	-0.0329 ***	-0.0296 ***	-0.0055
25 (Indictment)	1318	-0.0057	-0.0066	-0.0062	-0.0057	-0.0176
Overall Mean Hedge *CAR*		-0.0033	-0.0048 *	-0.0023	-0.0028	-0.0037
p-value		0.11	0.06	0.2	0.15	0.13
p-value for difference from AA				0.28		0.41

Table 3 (continued)

Panel B. Mean hedge *CARs* comparisons between AA and OB5 auditees pooled over all potential spillover and non-spillover windows identified in Table 2, panel A.

	All	AA	OB5	B5	NB5
Non-spillover windows	0.0048 **	0.0028	0.0057 **	0.0051 **	0.0033
Potential spillover windows	-0.0122 ***	-0.013 ***	-0.011 ***	-0.0114 ***	-0.0114 ***
p-value for difference from AA (Non-Spillover windows)			0.25		0.47
p-value for difference from AA (Potential spillover windows)			0.36		0.38

Table 4
Mean Abnormal Returns by Auditor Type for Potential Spillover and Non-spillover Windows and for Solvent and Financially Distressed Auditees

Abnormal returns are four-day [0, +3] cumulative abnormal returns (*CARs*) around news events described in Table 1, computed using a four-factor model as in Carhart (1997). The sample comprises of all domestic auditees in the CRSP-Compustat universe for which relevant data items are available. All = All auditors, B5 = All Big Five auditors (AA + OB5), AA = Arthur Andersen, OB5 = Other Big Five (DT = Deloitte and Touche, EY = Ernst & Young, KPMG = KPMG, PwC = PricewaterhouseCoopers), and NB5 = non-Big-Five auditors. The twelve potential spillover windows are those identified in Table 2, panel A. Z is Altman's bankruptcy z-score. p-values in all panels are based on *Fama-MacBeth* (1973) *t* statistics of the mean *CARs* across the 12 potential spillover or the 13 non-spillover windows.

* (**, ***) denotes significantly different from zero at 10% (5%, 1%) levels.

Panel A: Mean *CAR* comparisons for 13 non-spillover windows

Non-spillover windows (13)	All	AA	OB5	B5	NB5
Solvent auditees (Z≥3)	0.0006	0.0022	-0.0002	0.0002	0.003 *
Distressed auditees (Z<3)	0.0035 ***	0.0024	0.0044 ***	0.0039 ***	0.0008
p-value for difference from AA	Solvent		0.12		0.39
p-value for difference from AA	Distressed		0.22		0.33

Panel B: Mean *CAR* comparisons for 12 potential spillover windows

Potential spillover windows (12)	All	AA	OB5	B5	NB5
Solvent auditees (Z≥3)	-0.0047 ***	-0.0045 ***	-0.0038 **	-0.0039 **	-0.0099 ***
Distressed auditees (Z<3)	-0.009 ***	-0.009 ***	-0.0072 ***	-0.0076 ***	-0.0172 ***
p-value for difference from AA	Solvent		0.37		0.05
p-value for difference from AA	Distressed		0.19		0.04

Panel C: Mean *CAR* auditor-by-auditor comparisons for non-spillover and potential spillover windows

Non-spillover windows (13)	AA	DT	EY	KPMG	PWC
Solvent auditees (Z≥3)	0.0022	0.0016	0.0000	-0.0002	-0.0016 *
Distressed auditees (Z<3)	0.0024	0.0031 *	0.0054 **	0.0037 *	0.0046 *
p-value for difference from AA (Solvent)		0.41	0.17	0.17	0.04
p-value for difference from AA (Distressed)		0.41	0.17	0.34	0.27

Potential spillover windows (12)	AA	DT	EY	KPMG	PWC
Solvent auditees (Z≥3)	-0.0045 ***	-0.0044 **	-0.0035 **	-0.0025	-0.0046 **
Distressed auditees (Z<3)	-0.009 ***	-0.0062 ***	-0.0062 ***	-0.0088 ***	-0.0075 ***
p-value for difference from AA (Solvent)		0.47	0.32	0.22	0.48
p-value for difference from AA (Distressed)		0.16	0.12	0.47	0.30

Table 5
Cumulative Mean Abnormal Returns by Auditor Type for Solvent and Financially Distressed Auditees Across Twelve Potential Spillover Windows

Abnormal returns are four-day [0, +3] cumulative abnormal returns (*CARs*) for twelve potential spillover windows identified in Table 2, panel A, computed using a four-factor model as in Carhart (1997). The sample comprises of all domestic auditees in the CRSP-Compustat universe for which relevant data items are available (4566 auditees). The window mean *CARs* are aggregated across the twelve windows. AA = Arthur Andersen, OB5 = Other Big Five (DT = Deloitte and Touche, EY = Ernst & Young, KPMG = KPMG, PwC = PricewaterhouseCoopers), and NB5 = non-Big-Five auditors. Z is Altman's financial distress z-score.
* (**, ***) denotes significantly different from zero at 10% (5%, 1%) levels.

Cumulative Mean *CARs*: Solvent auditees (Altman Z-score >= 3)

Event	AA	DT	EY	KPMG	PWC	NB5
1 (WMI Charge)	-0.52%	-0.34%	-0.28%	0.24%	-0.19%	-1.21%
4 (Barrons)	-0.82%	-0.85%	-0.77%	0.13%	-0.28%	-1.62%
6 (SBM Acct Inquiry)	-1.74%	-2.08%	-1.36%	-0.45%	-0.97%	-3.97%
9 (SBM Restatement needed)	-2.20%	-2.12%	-1.36%	-0.84%	-0.59%	-5.25%
14 (AA Partner sued)	-1.78%	-1.71%	-0.66%	-0.82%	-1.01%	-5.69%
15 (WMI Penalty)	**-2.33%**	**-2.66%**	**-2.16%**	**-2.68%**	**-2.37%**	**-6.87%**
16 (Smoking Gun)	-3.43%	-1.89%	-1.79%	-1.64%	-2.53%	-9.39%
17 (Skilling)	-3.76%	-3.44%	-2.38%	-2.26%	-3.42%	-10.64%
21 (Enron Bankruptcy)	-3.62%	-3.37%	-2.24%	-1.16%	-2.56%	-11.80%
22 (Berardino)	-4.13%	-4.06%	-2.93%	-1.76%	-3.67%	-12.50%
23 (Shredding)	-4.89%	-4.58%	-3.24%	-2.07%	-4.77%	-10.94%
24 (Powers)	-5.41%	-5.23%	-4.22%	-2.97%	-5.56%	-11.84%

Cumulative Mean *CARs*: Financially distressed auditees (Altman Z-score < 3)

Event	AA	DT	EY	KPMG	PWC	NB5
1 (WMI Charge)	-1.01%	-0.64%	-1.57%	-1.26%	-1.41%	-2.57%
4 (Barrons)	-1.37%	0.43%	-1.53%	-0.33%	-1.64%	-2.41%
6 (SBM Acct Inquiry)	-2.14%	-0.76%	-1.83%	-2.02%	-3.09%	-4.63%
9 (SBM Restatement needed)	-3.10%	-1.95%	-2.27%	-2.81%	-4.02%	-7.75%
14 (AA Partner sued)	-3.75%	-1.93%	-2.28%	-4.06%	-4.45%	-8.45%
15 (WMI Penalty)	**-5.88%**	**-3.72%**	**-3.36%**	**-5.46%**	**-5.53%**	**-9.78%**
16 (Smoking Gun)	-6.66%	-5.13%	-3.20%	-5.52%	-7.71%	-12.84%
17 (Skilling)	-6.76%	-5.95%	-3.72%	-7.04%	-7.86%	-13.46%
21 (Enron Bankruptcy)	-7.32%	-6.10%	-4.85%	-8.54%	-7.53%	-18.00%
22 (Berardino)	-8.55%	-6.93%	-5.78%	-9.66%	-7.22%	-19.96%
23 (Shredding)	-9.92%	-7.61%	-6.04%	-9.80%	-7.74%	-20.12%
24 (Powers)	-10.75%	-7.47%	-7.42%	-10.52%	-9.00%	-20.66%

Table 6
Determinants of Abnormal Return Variation over 25 Bad News Windows

The sample comprises of all domestic auditees in the CRSP-Compustat universe for which relevant data items are available. Coefficients reported are obtained from OLS regressions of cumulative abnormal returns on the vector of explanatory variables after clustering standard errors by window to account for within-window cross-sectional correlation in security returns. *|AbAcc|* is the absolute value of abnormal accruals and *CFO* is cash flow from operations computed as described in the text. *S* takes the value 1 during the twelve potential spillover windows identified in Table 2, panel A, 0 otherwise. *Manuf* takes the value 1 if the observation is from an industry with SIC Code between 2000 and 3999 (inclusive), 0 otherwise. *R&D* takes the value 1 if the observation is from an industry with SIC Code between 2800 and 2899 (both inclusive) or between 3500 and 3899 (both inclusive) or between 7370 and 7379 (both inclusive), 0 otherwise. *ZScore* is Altman's bankruptcy score. *Zdum* takes the value 1 if *ZScore* ≤ 1.8. In each window, observations are partitioned into nine portfolios based on independent sorts into terciles based on absolute values of abnormal accruals and cash flows as described in the text. *S* takes the value of 1 if an observation is in one of the twelve spillover windows and 0 if in one of the thirteen non-spillover windows. *Port* takes the value 1 if an observation is in the low-accrual, high-cash-flow (high quality) portfolio, 0 otherwise.

Variable	*Coefficient (significance levels)*			
(Expected	*1*	*2*	*3*	*5*
Sign)	*Model 1*	*Model 2*	*Model 3*	*Model 4*
Constant	-0.001	0.002	0.004 ***	0.003 **
|AbAcc|	-0.01 *	0.001	0.001	
CFO	0.002	-0.014 ***	-0.014 ***	
Port				-0.002 *
Manuf			-0.004 **	-0.004
R&D			0.001	0.002
Zscore			< 0.001	
Zdum				0.002
S[§] (-)		-0.007 ***	-0.009 ***	-0.006 ***
*|AbAcc|*S*[§] (-)		-0.02 **	-0.02 *	
*CFO*S*[§] (+)		0.031 ***	0.029 ***	
*Port*S*[§] (+)				0.006 ***
*Manuf*S*			0.004 **	0.005 *
*R&D*S*			-0.002	-0.006 **
*ZScore*S*			< 0.001	
*Zdum*S*				-0.01 ***
R^2	< 0.001	0.003	0.003	0.003
N	119400	119400	101403	101403

[§] Test variables associated with directional hypotheses, one-tail significance levels reported.
* (**, ***) denotes significantly different from zero at 10% (5%, 1%) levels.