HOW PERVASIVE IS CORPORATE FRAUD?

Alexander Dyck
University of Toronto

Adair Morse
University of California at Berkeley & NBER

Luigi Zingales*
University of Chicago, NBER, & CEPR

February 2021

ABSTRACT

We provide an estimate of the undetected share of corporate fraud. To identify the hidden part of the 'iceberg,' we exploit Arthur Andersen's demise that triggered added scrutiny on AA's former clients, thereby increasing the detection likelihood of preexisting frauds. Our evidence suggests that in normal times only one third of corporate frauds are detected. We estimate that on average 11% of large publicly-traded firms are committing securities fraud every year. Combining this figure with existing estimates of the costs of detected and undetected fraud, we estimate that corporate fraud destroys 1.7% of equity value each year, equal to $744bn in 2020.

* We thank Patricia Dechow, Michael Guttentag, Katherine Guthrie, Phil McCollough, Frank Partnoy, Richard Sloan, Joseph P. Weber, Michael Weisbach, and anonymous referee, as well as participants at the NBER Crime Working Group, the University of Illinois Symposium on Auditing Research, the European Finance Association, the University of Chicago, USC, the Rotman School at the University of Toronto, Berkeley-Haas Accounting, Berkeley-Haas Finance, and Queens University. Alexander Dyck thanks the Connaught Fund of the University of Toronto, and Adair Morse and Luigi Zingales, the Center for Research on Security Prices, the Stigler Center, and the Initiative on Global Financial Markets at the University of Chicago for financial support. Luigi Zingales has worked as a consultant for the PCAOB.

Starting with Jensen and Meckling (1976), a very large literature documents the agency costs of public ownership. One of the less emphasized of these costs is fraud. To enrich themselves, managers who own a relatively small fraction of the stock, might be willing to break the law, even when the cost of breaking the law from the company's point of view exceeds its benefits. Is corporate fraud a major component of the agency costs of public ownership? Should regulation do something to reduce this cost?

Before we even try to answer these questions, we need to establish how pervasive corporate fraud is: is the fraud we observe the whole iceberg or just its visible tip? The answer to this question requires an estimate of the ratio of the exposed tip to the submerged portion, also known as the detection likelihood. Thus far in the literature, approaches to estimate the detection likelihood (and by implication the hidden prevalence of fraud) have included (i) accounting misconduct prediction models (Dechow, Ge, Larson and Sloan (1996) and Beneish (1999)), (ii) corporate surveys (Graham, Harvey and Rajgopal (2013)), and (iii) structural, partial-observability approaches (Wang (2013), Winton, Wang, and Yu (2010), Hahn, Murray, and Manolopoulou (2016), and Zakolyukina (2018)). In this paper, we introduce a fourth approach, which relies on a natural experiment: the 2001 demise of Arthur Andersen (AA).

The simple idea is that after the AA demise, former AA clients were subject to vastly increased scrutiny. They found themselves in the spotlight of the media, investment intermediaries, short-sellers and their internal gatekeepers. In addition, they were forced to seek a different audit firm. Given the extreme cloud of suspicion that was covering AA clients immediately after the Enron scandal exploded (Chaney and Philipich (2002), Krishnamurthy, Zhou, and Zhou (2006)), the new auditors as well as all other fraud detectors had strong incentives to uncover any fraud committed by former AA clients. Even if this increased scrutiny does not reveal all the existing fraud, the Kolmogorov axiom of conditional probability allows us to derive an upper bound (and thereby, conservative) estimate of the detection likelihood, which in turn provides us with a lower bound estimate of the pervasiveness of corporate fraud.

As we explain in Section I, we use the term fraud loosely, since what we measure is some form of misconduct or alleged fraud. To estimate the detection likelihood, we use several different measures. The first measure is auditor-detected securities fraud from Dyck, Morse and Zingales (2010) (hereafter,

"DMZ"). The second is SEC Accounting and Auditing Enforcement Releases (AAERs) from Dechow et al. (2011). The third is financial misreporting not due to simple clerical errors. The last one is all securities fraud under *SEC section 10b-5*.

We find that the detection likelihood varies depending on the measure used: 0.29 for auditor-detected securities fraud, 0.34 for accounting restatements, 0.47 for securities fraud at large, and 0.52 for AAERs. All of the estimates are statistically significant and robust to a host of potential omitted differences between the firms audited by Arthur Andersen and the rest (e.g. industry focus, regional variation, and importance of IPOs). Our estimates represent an upper bound of the detection likelihood, which implies that our estimates of unobserved fraud are conservatively low. The most plausible estimate of the detection likelihood is 0.33, with 95% confidence interval between 0.24 and 0.45. Fraud is indeed like an iceberg with significant undetected fraud beneath the surface.

With our estimates of the detection likelihood, we can calculate how pervasive corporate fraud is, as a function of the definition of fraud we adopt. Accounting violations are widespread: in an average year 43% of companies misrepresent their financial reports, even when we ignore simple clerical errors. Fortunately, securities fraud is less pervasive. In an average year, 11% of all large public corporations commit (alleged) securities fraud (with a confidence interval between 8% and 14%).

Having estimated the pervasiveness of fraud, we can calculate the losses produced by fraud, using cost estimates from the prior literature. For detected frauds, we use equity drop at the revelation of the fraud as estimated by Karpoff, Lee and Martin (2008), henceforth KLM). For undetected frauds, we apply the average value lost due to cover-up, as estimated by Beneish, Lee and Nichols (2013). Combining these detected and undetected fraud losses – 25% and 11% of the equity value, respectively – to the proportion of frauds detected (0.33) and undetected (0.67) results in an average fraud cost of 16% of equity value. Putting together our estimate of the pervasiveness of fraud (11%) with the per-firm cost of fraud (16%), we arrive to an expected annual cost of fraud of 1.7% of the total equity value of U.S. public firms (with a 95 percent interval between 1.2% and 2.3%). Hence, the annual cost of fraud among US corporations at the end of our sample period (2004) is $265 billion or -- bringing our cost calculation forward to 2020 -- $744

billion. If we compare the 2004 expected cost of fraud with the $19.9 billion of annual SOX compliance cost (as estimated by Hochberg et al. (2009)), for the benefits of SOX to exceed its costs it would suffice that SOX reduced the probability that a fraud is initiated by 1 percentage points (equal to 9% of the baseline probability).

Our paper builds on a large literature seeking to understand the incidence of undetected corporate fraud or misconduct. Dechow, Ge, Larson and Sloan (1996) and Beneish (1999) estimate the likelihood of accounting misconduct through patterns in financial disclosure. Wang (2013) uses a bivariate probit to identify factors affecting the propensity to commit fraud and the vulnerability of fraud to detection. Hahn, Murray, and Manolopoulou (2016) integrate Wang's bivariate probit with a flexible Bayesian prior. Finally, Zakolyukina (2018) uses a structural model of earnings manipulation. Instead of relying on statistical or economic models, our approach is based on a natural experiment, with all the pluses and minuses of natural experiments.

Our paper complements Krishnan, Raghunandan and Yang (2007), who document how after AA's demise Big 4 auditors were more likely to issue a going concern qualification when they audited large former AA clients than when they audited similar non-AA clients. They do not find a similar result for small firms, likely due to a greater rejection of small AA clients with dubious accounting by big 4 auditors. We find that when we look at all auditors (not just Top 4), the revelation of misconduct is more likely in small former AA clients than in similar-sized non-AA clients.

Finally, our paper also builds off the literature seeking to understand the costs associated with corporate fraud, both in direct reputation effects (Karpoff and Lott (1993) and KLM), and in cover-up costs of avoidance (Beneish, Lee and Nichols (2013)).

The rest of the paper proceeds as follows. Section I describes the data and presents summary statistics on caught frauds. Section II explains our methodology. Section III presents the detection likelihood and the pervasiveness of corporate fraud results. Section IV provides estimates of the cost of such fraud, and Section V concludes.

I. Observed Corporate Fraud Measures & Incidence

We start by defining corporate fraud. As one court observed, "[t]he law does not define fraud; it needs no definition; it is as old as falsehood and as versatile as human ingenuity."[1] In a similar vein, the Fourth Circuit noted "fraud is a broad term, which includes false representations, dishonesty and deceit." [2] Thus, it is not easy to define corporate fraud precisely. Securities law defines securities fraud as "To make any untrue statement of a material fact or to omit to state a material fact necessary in order to make the statements made, in the light of the circumstances under which they were made, not misleading."[3] Since our primary measure derives from securities class actions, we assume that definition and we regard corporate fraud primarily as misrepresentation.

In spite of this ambiguity, the established legal practice requires the presence of three elements to label an act of misconduct fraud: misrepresentation, materiality, and intent. As we will explain below, all our measures of misconduct contain an element of misrepresentation. We do our best to set a minimum threshold to ensure materiality. Where we cannot deliver is on intent. Intent can only be proven in court and all these alleged fraud cases are settled before they reach the final verdict, since directors' and officers' insurance does not indemnify executives if they are convicted in court (see, for example, the data and discussion in Black, Cheffens and Klausner (2006)). As a result, all the corporate fraud we discuss is *alleged* fraud that was settled out of court.

To capture alleged fraud we resort to four different measures. The first measure is financial misrepresentations exposed by auditors. To construct "auditor-detected fraud", we start with SEC *Rule 10b-5* securities fraud from DMZ's dataset of class actions suits. To ensure comprehensiveness, DMZ focus on large firms (those with over $750 million in assets) so that a potential payoff from a suit would be

[1] Weiss v. United States, 122 F.2d 675, 681 (5th Cir. 1941), cert. denied, 314 U.S. 687 (1941).
[2] See United States v. Grainger, 701 F.2d 308, 311 (4th Cir. 1983), cert. denied, 461 U.S. 947 (1983).
[3] https://ecfr.federalregister.gov/current/title-17/chapter-II/part-240/subpart-A/subject-group-ECFR271f57d9adc8b47/section-240.10b-5.

sufficiently attractive to mobilize monitoring by legal actors.[4] To remove frivolous suits, DMZ impose a series of filters, building on the existing legal literature.[5] The final step is to restrict the DMZ sample to frauds that are revealed directly or indirectly by an auditor. Directly revealed by the auditors are frauds where the auditor in DMZ was the whistleblower. Indirectly revealed by the auditor are frauds where the auditor issued a qualified opinion and the whistleblower in DMZ is either the company itself or an analyst. For these latter cases, we re-read the auditor's qualifying opinion in the DMZ cases where the company itself or an analyst was the whistleblower and dropped two cases where the fraud could not be plausibly related to the reason why the audit opinion was qualified.

The second measure is financial misrepresentation that lead to an Accounting and Auditing Enforcement by the SEC, as contained in the AAER dataset created by Dechow, Ge, Larson and Sloan (2011). Since 1982, the SEC has issued Accounting and Auditing Enforcement Releases (AAERs) during or at the conclusion of an investigation against a company, an auditor, or an officer for alleged accounting and / or auditing misconduct. The benefit of this sample is that all AAERs are likely to be material violations, since an AAER release by the SEC is considered a severe step meant to attract attention. Dechow et al (2011) refers to AAER as misstatements "although fraud is often implied by the SEC's allegations" because "firms and managers typically do not admit or deny guilt with respect to the SEC allegations." We will refer to them simply as AAERs.

The main shortcoming of this measure is that the sample may be overly restrictive because the SEC does not have the budget to go after all frauds (e.g. Dechow, Sloan and Sweeney (1996) and Miller (2006)). Furthermore, as the SEC website notices: "[AAER] only highlights certain actions and is not meant to be a complete and exhaustive compilation of all of the actions that fall into this category."[6] For example,

[4] The legal industry has set up an automatic process whereby every time a stock price experiences a large drop, specialist attorneys file suits and scour financial reports for misrepresentations. Large, publicly-traded companies offer lucrative possibilities for claims on a suit. For these, it is unlikely that the class action legal structure would miss an opportunity to detect a securities fraud (Coffee, 1986).

[5] As described in Appendix I, DMZ restrict attention: first, to cases after a 1995 change in the law which forced courts to become more stringent about evidence to certify a case; second, to cases that are not dismissed; and third, to cases without low settlement amounts (below $3 million).

[6] https://www.sec.gov/divisions/enforce/friactions.shtml

during the period 2001-2014 there were on average 54 AAERs per year versus 220 securities class actions (Soltes, 2019).

The third measure does not capture fraud, but only restatements. To obtain such a measure we start from the most widely employed data on restatements: the Audit Analytics database. We then use Audit Analytics classification to separate clerical errors from intentional implementation of misstatements.

Finally, the fourth measure of fraud is the full set of *SEC section 10b-5* securities fraud cases. While all 10(b)-5 cases arise from some material misstatements or omissions in providing material information, they include also non auditing related misrepresentations. For example, *"The Department of Justice indicts Sotheby's for price fixing after a long investigation of the art industry. Shareholders sue the firm for failing to disclose that a large portion of their revenues are being derived from the unsustainable high prices achieved through the pricing arrangement with Christie's*." (DMZ Appendix, 2010). In this example, no *accounting* violation has occurred, but omitting the economic implications of fixing prices represents an SEC rule 10b-5 violation. In DMZ, 35% of the 10b-5 violations are omissions of information concerning an underlying legal violation, as in the Sotheby's example. Rather than using the original DMZ set of 10b-5 class actions, we use the updated dataset from the Securities Class Action Clearinghouse, as compiled by Kempf and Spalt (2019). We refer to this measure as *SCAC securities fraud*, again noting that these are alleged frauds.

Figure 1 plots the fraction of Compustat large U.S. corporations engaged in (alleged) fraud or misconduct (henceforth for simplicity "fraud") according to the various measures. The data are monthly and a company is classified as committing a fraud during a certain month if that month falls within the fraud period determined after detection. When the measure we use is Accounting Restatements, the fraction of firms involved in fraud is large (on average 13% per year) and increasing over time (which might reflect the increase complexity of accounting rules). All the other measures of misconduct produce a much lower incidence, between 1% and 4% of firms, and a pronounced hump shape centered around the turn of the

millennium. This pattern is consistent with Wang, Winton and Yu (2010), who find firms start fraud more in booms because monitoring is lower and managerial reward from cheating is higher.

II. Methodology: Inferring the Detection Likelihood using the Arthur Andersen Experiment

II.1. Experiment Overview

Figure 1 depicts the incidence of corporate frauds that are eventually caught. Our main agenda is to estimate what percentage of initiated frauds are caught. We refer to this percentage as the detection likelihood.

Our methodology relies on two key ingredients – the Kolmogorov axiom of conditional probability and the Arthur Andersen natural experiment. The pervasiveness of corporate fraud is the probability of a firm engaging in fraud (denoted *F*), regardless of whether it is caught or not:$\Pr(F)$. What we observe instead is the joint event of a firm engaging in fraud and being caught: $\mathbf{Pr(}\boldsymbol{F}\mathbf{,}\boldsymbol{caught}\mathbf{)}$. (We will use the convention of bolding the variables we observe). By the law of conditional probability, the unconditional probability of engaging in a fraud can be written as:

$$\Pr(F) = \frac{\boldsymbol{Pr(F,caught)}}{Pr(caught|F)}. \quad (1)$$

Thus, if we knew the denominator, $\Pr(caught \mid F)$, we could calculate $\Pr(F)$.

Our experiment derives an upper bound of $\Pr(caught \mid F)$ and thus a lower bound of $\Pr(F)$, by comparing the frequency of revealed fraud in "normal" times with the frequency of revealed fraud in a sample where scrutiny increased substantially. In our experiment, the sample of firms under enhanced scrutiny is the set of Arthur Andersen (AA) clients after the AA demise following the Enron scandal. On October 22, 2001 Enron acknowledged an SEC inquiry concerning possible conflicts of interest in various partnerships (Barton, 2005). On November 23, 2001, the NYT ran an article with the headline: "From Sunbeam to Enron: Andersen's Reputation Suffers". Thus, we assume that the period of enhanced scrutiny

started after November 30, 2001, a date to which we refer as the watershed date.[7] Since roughly one fifth of all large publicly traded firms had AA as their auditor in 2001, this event provides a natural experiment of increased scrutiny of firms. We look at the revelation of fraud that started being perpetrated before the watershed date. Thus, we can ignore any deterrence effect, since the enhanced scrutiny produced by the Enron's crisis was unexpected.

One source of the increased scrutiny are the new auditors (Chen and Zhou (2007)). Kealy et al. (2007) find that the audit fees charged by AA successors to AA former clients when they switched after Enron are positively related to the length of AA tenure with those clients, suggesting that there is a perceived risk associated with AA clients, which needs to be mitigated with additional monitoring. This enhanced scrutiny is also consistent with Nagy's finding (2005) that smaller ex-AA clients have lower discretionary accruals after switching to a new auditor.

Nevertheless, new auditors are not the only source of enhanced scrutiny. As Chaney and Philipich (2002) and Krishnamurthy, Zhou, and Zhou (2006)) show, after the Enron scandal exploded all AA clients got under a cloud of suspicion. Investment intermediaries and internal gatekeepers of AA clients had stronger incentives to scrutinize their firms thoroughly to clear them of the shadow of suspicion. Likewise, short sellers and the media had special incentives to focus their attention on AA clients in search for the next big scandal.

II.2. Identification Assumption 1: AA firms are not different in 1998-2000

Our first assumption is that in the period before AA's demise (specifically 1998-2000), fraud (F) was equally likely in AA firms and non-AA firms $(\overline{AA})$:[8]

Assumption 1: $$Pr(F \mid \overline{AA}) = Pr(F \mid AA)$$

[7] As a robustness test, we shifted the beginning of the detection period one month backward or one month forward: the results are substantially unchanged.

[8] To keep notation simple, we do not include time subscripts. We make the timing clear in the text.

AA's indictment by the Department of Justice and its initial conviction for obstruction of justice may make this a surprising assumption. Yet, the initial conviction of AA was for obstruction of justice, not for being a bad auditor, and it was unanimously overturned upon appeal.

More convincingly, the accounting literature has concluded that there was no difference between AA and other auditors in terms of auditing rigor. In a matched sample, Agrawal & Chada (2005) find that the existence of AA as the auditor is not associated with firms having more restatements. Likewise, controlling for client size, region, time and industry, Eisenberg & Macey (2004) find that AA clients did not perform any better or worse than other firms. Nevertheless, we re-test this assumption for assurance and because our sample of large U.S. corporations is different from the aforementioned studies.

Table 1 reports summary statistics, comparing the characteristics and industry of AA clients and non-AA clients during the period 1998-2000. The set of firms are Compustat large (>$750 million in assets) U.S. incorporations. We identify auditors using the Compustat "AU" item, mapping the auditor who signs a financial statement to the calendar month-year of the report. We use two sets of non-AA clients: all non-AA clients and the subset of all non-AA clients that are audited by another of the Big Five audit firms. Statistics are collapsed to one observation per firm over the time period. A firm appears in multiple columns only if it switches auditors. Because of auditor transitions and firms' delisting during the 3-year period, the sample of AA firms is larger than in the subsequent tables, where we require a firm to be audited by AA in 2001.

Panel A reports that AA clients are statistically similar to other Big Five clients in assets, sales, and EBITDA, but have higher leverage (long term debt-to-assets). The leverage difference suggests a difference in industry composition, which is what we present in Panel B. For example, compared to Big 5 auditors that are not AA, AA have at least 2 percentage points less clients in the industries of Banks & Insurance, Retail & Wholesale, and Computers, while having 2 percentage points more clients in the industries of Utilities, Refining & Extractive, Communication & Transport, and Services & Healthcare. We return to these industry compositional differences in our empirical specifications.

Were AA clients more fraudulent? In Table 2, we consider this question using eight measures of

fraud, collapsed to one observation per firm for the 1998 – 2000 period before AAs demise (allowing, as before, a firm to have two observations if it switched to/from AA in this period). For the discrete variables of caught fraud – *auditor-caught frauds, AAERs, restatements*, and *SCAC securities fraud*— we define a fraud event if the event ever occurs in the three-year period. Thus, to compare the detected fraud in this table to those in Figure 1 (and to those in the subsequent tables) we need to divide the reported number by three. In addition to these measures of fraud, we include measures of fraud likelihood based on a continuous measure from the accounting literature. We include both the probability of accounting manipulation score (Beneish, 1999) [*ProbM*][9] and the probability of fraud score of Dechow, Ge, Larson and Sloan (DGLS 2011) [*FScore*].[10] Following Beneish (1999) and DGLS (2011), we also consider a discrete version of the *ProbM* and *FScore* in the form of a threshold dummy variable with the authors considering a firm above the threshold as indicating a strong likelihood of being a manipulator.

Panel A of Table 2 reports that for all eight measures of fraud, AA clients are never more likely to be committing fraud. In fact, based on the measures from the accounting literature, AA clients show lower scores, indicating they are less likely to be committing fraud. These results do not account for the above-mentioned leverage and industry differences.

In Panel B, Table 2 we report estimates from multivariate estimations of fraud likelihood using accounting literature measures of *probM*, the indicator *probM > -2.2*, *FScore*, and the indicator *FScore >1.4* . For each estimation, our variable of interest is an AA dummy, and we include a set of controls for size (log assets, and sales/assets) along with debt (LT Debt/Assets), and profitability (EBITDA/Sales). When we estimate the threshold models (columns 3, 6, 9, and 12), we estimate via logit instead of OLS. All columns include year fixed effects. Columns 2, 3, 5, 6, 8, 9, 11, and 12 include industry fixed effects

[9] Appendix II details the calculations for the ProbM score. ProbM is a score with no natural scale. The mean and standard deviation of ProbM in our sample are -2.325 and 1.357, respectively. According to Beneish (1999), a score greater than -2.22 indicates a strong likelihood of a firm being a manipulator.

[10] We thank Weili Ge for the FScore data. We use their FScore measure from model 2 of the paper, which includes financial statement variables and market data. FScore is a score variable, scaled to imply that a score of 1.00 indicates that the firm has no more probability of AAER fraud than the unconditional probability. In our sample, the mean and standard deviation of Fscore are 1.785 and 10.57, respectively.

(where the industries are those of Table 1). Columns 2, 5, 8, and 11 include industry*year fixed effects. (We do not include this further interaction for the logit, as the estimation does not consistently converge uniquely.) In columns 4-6 and 10-12, we restrict the non-AA firms to be Big 5 accounting firm clients only.

Across all specifications, the coefficient on the AA dummy is never positive and significant. In columns 1 and 3, the AA dummy is negative and significant (i.e., AA firms are associated with *less* probability of manipulation). Being an AA client does not significantly affect the positive likelihood of manipulation or fraud in the pre-demise period.

In Panel C, we repeat these tests using as dependent variables *auditor-caught fraud*, *AAERs*, *restatements*, and *SCAC securities fraud*. In all even-numbered specifications, we control for industry crossed with year fixed effects, except for auditor-detected fraud whose small sample forces us to include only industry fixed effects. As in Panel B, the coefficient on the AA dummy is never positive and significant.

Krishnan and Visvanathan (2008) find that earnings management was more pronounced in the AA Houston office than in the Houston offices of other Big Five auditors. Since Enron was headquartered in Houston, this is not that surprising. Nevertheless, to address the concern that in certain regions AA audited firms may have been more likely to commit fraud, in Appendix Table A1, we restrict our sample to large corporations located in the state of Texas. For each AA client, we find a match among other Big Five clients within the two-digit SIC code, based on a propensity score to be an AA client. We generate the propensity score based on assets, sales, EBITDA, and leverage within industry. In Table A1 (columns 1 and 2), the dependent variable is the ProbM score. An AA dummy variable (equal to one for an AA client) is not significant either in the Texas sample (Panel A) or in the United States sample (Panel B). Across the remaining columns, we repeat the same tests using other measures of uncaught (*FScore*) and caught fraud *auditor-caught fraud*, *AAERs*, *restatements*, and *SCAC securities fraud*). The results are unchanged. We find no evidence that AA clients are statistically or economically different from other auditors' clients, consistent with Assumption 1.

II.3. Identification Assumption 2

Our methodology is based on the assumption that after the watershed date, AA clients were subject to an enhanced level of scrutiny not only by the new auditors, but also by gatekeepers, short sellers, investment intermediaries, and the media. This assumption implies that $Pr(caught|\, F, AA) > Pr(caught|\, F, \overline{AA})$. While we cannot test this inequality directly, the hypothesis is consistent with the decline in the stock price of AA clients at the announcement of AA problems (Chaney and Philipich (2002)) and indictment (Krishnamurthy, Zhou, and Zhou (2006)). Furthermore, – as we will show in Table 3 – this inequality is true in the data for all our measures of fraud.

To obtain a point estimate on the amount of undetected fraud we further assume:

Assumption 2: $$Pr(caught|\, F, AA) = 1.$$

When it comes to purely financial fraud or financial restatements, Assumption 2 may be plausible, since this was the fraud discovered in Enron (and later in WorldCom) and this is the type of fraud every gatekeeper was fearful of missing. When we use as a measure of fraud all SCAC cases (including for instance price fixing), however, the level of alertness was less. The reason is that some of the detectors with enhanced incentives (like the new auditors) have no role in uncovering these types of fraud. Thus for some types of fraud it is more likely that $Pr(caught|\, F, AA) > Pr(caught|\, F, \overline{AA})$, but $Pr(caught|\, F, AA) < 1$. When this is the case– as we explain below – our estimate will yield an upper bound of the detection likelihood and a lower bound on the amount of undetected fraud.

In some cases (Kohlbeck, Mayhew, Murphy and Wilkins (2009) estimate 25% of the cases), AA clients switched audit firms but not audit partners: the former AA engagement partner moved to another audit firm, bringing the account. This continued relationship might reduce the new auditor's willingness to clean the dirty laundry; yet it is hard to imagine that it will completely eliminate the power of the experiment. Speaking about a firm that went from AA to Grand Thornton, the CEO of Grand Thornton testified at trial that "We converted it to a Grant Thornton audit approach and Grant Thornton audit-

quality controls, and we had other people review the engagement."[11] Thus, even when the audit partner remains the same, the new audit firm performs an extra screening.

A more general concern is that the increased attention on corporate fraud that followed the Enron and WorldCom scandals might have prompted the SEC (or the other auditors not affected by the turnover) to become more active in detecting fraud. In the limit, if this enhanced scrutiny exposed all fraud in all firms, there will be no difference between the amount of fraud revealed in AA clients and the amount of fraud revealed in clients of other audit firms, invalidating our experiment. Yet, as long as some enhanced scrutiny affects all firms, but AA firms are affected more, then our methodology will work, but will underestimate undetected fraud. Thus, our results should be interpreted as a lower bound on the pervasiveness of fraud.

Finally, the AA-demise was completely unexpected, so it could not have altered the ex-ante incentives to commit fraud. In this sense, it is different from a mandatory turnover, which is anticipated.

II.4. Conditional Probabilities within the AA Experiment

With Assumptions 1 and 2 in hand, we can use the law of conditional probability to derive an estimator for the detection likelihood of fraud. If we write down two versions of equation (1) (one for AA and one for non-AA), bringing the denominators to the other side, and dividing one by the other, we have a relationship that puts only observable detections of fraud on the right-hand side:

$$\frac{Pr(caught|\ F,\ \ \overline{AA})\cdot Pr(F|\ \overline{AA})}{Pr(caught|\ F,\ \ AA)\cdot Pr(F|\ AA)} = \frac{\boldsymbol{Pr(F,caught|\ \overline{AA})}}{\boldsymbol{Pr(F,caught|\ AA)}}. \tag{2}$$

Substituting Assumptions 1 and 2 into the left-hand side of equation (2) implies:

$$\Pr(caught|F,\overline{AA}) = \frac{\boldsymbol{Pr(F,caught|\ \overline{AA})}}{\boldsymbol{Pr(F,caught|\ AA)}}. \tag{3}$$

Since the normal level of scrutiny is the one experienced by non AA clients, then $\Pr(caught|F,\overline{AA}) = \Pr(caught|F)$. Thus, Equation (3) provides an estimator for the detection likelihood of fraud that is based

[11] https://www.wsj.com/articles/SB112951490246670395.

solely on two observables: the emergence of fraud in non-AA firms and the emergence of fraud in former AA firms.

Note that to the extent that not all fraud emerges in former AA clients (Assumption 2 becomes $Pr(caught|\ F, AA) < 1$), then our estimate $Pr(\ F)$ of the pervasiveness of fraud represents a lower bound of the real amount of fraud. To see this, consider that

$$\Pr(F) = \frac{\mathbf{Pr}(\mathbf{F}, \mathbf{caught})}{\Pr(caught \mid F)} > \frac{\mathbf{Pr}(\mathbf{F}, \mathbf{caught})}{\mathbf{Pr}(\mathbf{caught} \mid \mathbf{F})},$$

where the inequality derives from (3) and $Pr(caught|\ F, AA) < 1$).

Note also that we derive the detection likelihood by comparing fraud detection in two groups (former AA clients and non-AA clients) at the same time. Thus, this estimate should not be affected by fluctuations in the probability of committing fraud, as long as these fluctuations are similar in the two groups. Yet, these fluctuations will matter for the level of fraud pervasiveness at any point in time, since this is likely to vary over the business cycle as shown by Wang et al. (2010).

II.5. Empirical Methodology

Having explained why the detection likelihood given in Equation (3) is the number we want to estimate, we now describe how we will estimate it. For each company we observe whether fraud is discovered during the so-called detection period, both for companies that were AA clients and for companies that were not. This is tantamount to observing whether a subject gets infected with Covid-19 depending on whether he was treated with a vaccine or a placebo. This experiment can be thought of as a sum of Bernoulli trials, which can be modeled using a Poisson probability distribution, where the mean λ is the proportion of firms discovered as fraudulent per unit of time. The maximum likelihood estimate of λ in a Poisson distribution is simply the frequency of fraud during the detection period. To facilitate a system estimation (see below), we will estimate this mean λ with a Poisson regression, where $E(Y \mid X) = \lambda = e^{\theta x}$, and report both the coefficient θ.

In this framework, the object of our interest (i.e., the detection likelihood ($\Pr(caught|F,\overline{AA}) = \frac{\boldsymbol{Pr(F,caught|\ \overline{AA})}}{\boldsymbol{Pr(F,caught|\ AA)}}$) is simply the Relative Risk (RR) of a firm being caught committing a fraud if it is not an AA former clients versus if it is an AA former client (to map our case into the vaccine-testing language we should define the treated group as the non-AA clients). The detection likelihood, thus, is simply the ratio of the frequency of fraud in the two samples. Given the paucity of fraud observations, we compute the exact p-values of this ratio in finite samples. Following Agresti (1992) we do so by enumerating all possible draws in the sample and then computing the frequency of outcomes at least as different from the null hypothesis as the one observed.

III. Results

The timing of the experiment is as follows. We define a company as having AA as an auditor if AA signed a financial report anytime in the calendar year 2001, irrespective of the firm's fiscal year. All companies without AA as auditor during this period are non-AA clients. We consider that a fraud was revealed during the detection period if the fraud started before the watershed date of November 30, 2001, and it came to light between the watershed date and the end of 2003 (the detection period). Note that our estimates are obtained from the ratio of the fraud caught in AA and non-AA firms. Thus, as long as the amount of fraud committed in AA and not AA clients respond to cyclical fluctuations in a similar fashion, our estimates are not affected by business cycle fluctuations in the *aggregate* level of fraud, an important phenomenon documented by Wang, Winton and Yu (2010).

The count of observations is shown at the bottom of Table 3, Panel A. The natural experiment includes 353 AA clients in the pre-period who survive at least one year in the post-period, and 2404 non-AA client firms. The number of detected fraud events depends on the fraud measure: 168 restatements, 63 SCACs, 59 AAERs, and 21 auditor-detected fraud.

III.1 Detection Likelihood Results

Table 3, Panel A reports the main detection likelihood estimates across the four fraud measures – *auditor-detected fraud*, *AAER*s, *restatements*, and *SCAC securities fraud*. For each of these measures, we

report the estimated coefficient of the Poisson regression, the implied detection likelihood, and both the asymptotic and the exact p-values for the hypothesis that the detection likelihood is equal to 1. In each sample, the estimated coefficient of the Poisson regression we report is the coefficient of a dummy variable equal to one if a company was a former AA client. Thus, if AA former clients have an average frequency of fraud $\lambda_{AA} = e^{\theta_{AA} x}$ and non-AA clients of $\lambda_N = e^{\theta_N x}$. The estimated coefficient we report is $\theta_{AA} - \theta_N$. Since the detection likelihood ratio is nothing but $\frac{\lambda_N}{\lambda_{AA}} = \frac{e^{\theta_N}}{e^{\theta_{AA}}} = e^{\theta_N - \theta_{AA}}$, the detection likelihood ratio can be obtain as $\frac{1}{e^{\theta_{AA} - \theta_N}}$, or the inverse of *e* raised it to the power of the estimated coefficient.

Thus, the test that the detection likelihood is less than one verifies the assumption that after the demise of AA more fraud are caught in AA firms than in non AA firms ($\Pr(caught \mid F, AA) > \Pr(caught \mid F, \overline{AA})$). As we can see in Table 3, the hypothesis that the detection likelihood is bigger or equal to 1 is rejected at least at the 10% level in all four measures and in three out of four cases at the 5% level or better.

More specifically, when we look at *auditor-detected fraud* the estimated detection likelihood is 0.29, different from 1 at the 3% level. For *AAERs*, the detection likelihood estimate is 0.52, which is different from 1 at the 7% level. The detection likelihood estimate for accounting violations is 0.34, different from one at the 0.1% level. For the *SCAC securities fraud* measure, the detection likelihood is 0.47, different from 1 at the 2% level.[12] In sum, despite the different sources and the different definitions of fraud, there is a clear result: a substantial amount of corporate fraud remains undetected: if the detection likelihood is between 29% and 52%, between 48% and 71% of total fraud goes undetected.

If we want to move beyond this simple result and estimate more precisely the amount of undetected fraud, we need to pool together all the possible observations toward the estimation of a single detection

[12] To show that our results do not depend upon IPO-induced securities class actions, we also report SCAC securities fraud purged of three IPO frauds. Using this measure, the detection likelihood estimate is 0.53, very close to the 0.47 reported in the text.

likelihood. This is what we do in Table 3, Panel B. The first column reports the estimates obtained by estimating simultaneously the four Poisson regressions, with the restriction that the coefficient should be the same in all four. When we pool the four measures of fraud, the estimated detection likelihood is 0.38, with a 95% confidence interval, between 0.30 and 0.49. Thus, we can say with a 95% confidence that between 51% and 70% of all fraud is undetected or, similarly, that the total amount of corporate fraud is between 2.0 and 3.3 times the amount of corporate fraud observed.

These estimates of undetected fraud are downward biased because Assumption 2 (that after the Enron scandal all the fraud in former AA firms was revealed) is unlikely to be satisfied. This is true for all the measures, but particularly so for SCACs and AAERS. The SCAC securities fraud measure includes fraud that are not misrepresentation of accounting information, but failures to disclose material information. It is unlikely that the additional scrutiny triggered by the AA demise would expose all these failure-to-disclose cases. In such cases, the experiment design implies that 0.47 detection likelihood is only an upper bound (biasing downward the proportion of undetected fraud). Similarly, the detection of an AAER-type of fraud depends on the willingness of the SEC to bring an enforcement action. In any economic analysis of crime and punishment (see Becker, 1968), the SEC's incentives to bring a case against a defunct firm like AA are small. Thus, the 0.52 AAER estimate is also likely an upper bound.

By contrast, Assumption 2 is more likely to hold for misconducts that auditors are more likely to catch, such as accounting restatements and auditor-detected fraud. Thus, in column two of Table 3.B we report the detection likelihood estimate obtained by pooling together only the two fraud measures for which Assumption 2 is more likely to hold. The estimated detection likelihood is 0.33, with a 95% confidence interval of 0.24-0.45. Thus, only one-third of corporate fraud is detected and the total amount of corporate fraud is three times the corporate fraud observed. In the rest of the paper we will treat this as our best point estimate, with a 95% confidence interval between 2.2 and 4.1.

III.2. Detection Likelihood Robustness

Small Firms

Krishnan, Raghunandan and Yang (2007) document that top 4 auditors were much more likely to issue a going concern qualification in their audits of large former AA clients than in their audits of large non-former AA clients (specifically, 5.6% versus 2.3%). In our experiment, this ratio would imply a detection likelihood of 0.41, quite consistent with our results. Yet, Krishnan et al (2007) find the opposite for small firms: former AA clients received less than half of the going concern opinions of non-AA clients (6.2% to 14.5%). As the authors recognize, this difference in the results for large and small firms is driven by sample selection. Krishnan et al (2007) restrict their attention to AA firms that are subsequently audited by Top-4 auditors. It is likely that Top-4 auditors accept all the large former AA clients, but only the best small ones, leaving the rest to less prestigious auditors. In this paper we can test this conjecture by looking at all former AA clients, not just those subsequently audited by a top-4 auditor. To do so, however, we need to expand the sample to small firms.

Table 4, panel A reproduces our detection likelihood estimation setup for smaller public corporations – those that were never over $750 million in assets in the pre-period years. We run the experiment in the same setup as Table 3. Since DMZ does not collect auditor-detected data for small firms, the first measure of misconduct is not available for this sample.

First of all, for all three measures the detection likelihood is less than 1 and for two out of three significantly so at the 1% level. This analysis, suggested by an anonymous referee, provides an out-of-sample validation of our methodology to use the AA forced turnover to identify the misconduct that generally remains undetected.

For the AAER and restatement measures, the detection likelihood estimates obtained using small firms are very similar to the ones obtained using large firms: 0.64 vs. 0.52 for AAERs and 0.42 vs. 0.34 for restatements. By contrast, for securities fraud cases, the detection likelihood falls drastically down from 0.47 (large firms) to 0.29 (smaller firms). Most likely, this drop is due to

the fact that the probability of a class action suit among small non-AA clients is only 0.5% vs. the 2% observed in large firms. This is hardly surprising. The main reason why DMZ restrict their sample to large companies is that the discovery of a fraud will always lead to a class action suit in large companies, where there is plenty of money to pay the lawyers, but not necessarily in small companies. Thus, the lower detection likelihood reflects a real difference between large and small companies: more fraud go undetected in small companies.

Placebo

It is possible that AA clients were in businesses intrinsically more prone to corporate fraud or its detection. One way to help to rule out this possibility is to observe these firms in a different time period and check that they are not behaving differently. To this purpose, we reproduce the same experiment comparing former AA and non-AA clients (minus a few firms that do not survive) in the two years after the detection period (i.e., 2004-2006) and measure fraud detection in these years. Since the enhanced scrutiny following the demise of AA cannot last forever, we regard this exercise as a placebo test. This placebo period coincides with the beginning of the implementation of SOX. A large literature has tried to establish what the effects of the introduction of SOX are (see Coates and Srinivasan (2014)). Our test, however, is unaffected by any impact of SOX, since it compares the detection rate in former AA clients and former non- AA clients at the same time.

Table 4 panel B reports the results. The percentage of firms caught committing fraud is very similar between former AA clients and former non-AA clients (detection likelihood not different from 1), suggesting that there is not a natural proclivity of former AA clients to commit more fraud. In addition, the similarity between the fraud revealed in AA and non-AA clients suggest that the enhanced disclosure of fraud during the treatment period is not just an acceleration of the discovery that would have taken place regardless, but a net increase in discovery.

Industry and Geography

In section *II.2,* we showed that AA had more clients in Communications & Transport, Refining & Extractive, Services & Healthcare, and Utilities, while less clients in Banks & Insurance, Retail & Wholesale and Computers. If – for unspecified reasons – corporate fraud was more prevalent among sectors in which AA was over-represented just prior to the detection period and not afterward, then our detection likelihood estimate could be biased. To address this concern, we report the detection likelihood estimates in various sub-samples that remove industries where AA was either over- or under-represented. (Appendix Table A2, Panels B and C report the industry and regional distribution of fraud for all the fraud measures.) As Table 5 panel A shows, the detection likelihoods remain substantially unchanged.

The same concern could arise from regional variation in AA clients. For this reason, in Table 5 panel B, we repeat the same exercise excluding some regions or some large states. Once again, the detection likelihood results appear stable.

III.3. Pervasiveness of Corporate Fraud Results

We now can use the detection likelihood estimates in equation (1): $\Pr(F) = \frac{\boldsymbol{Pr(F,caught)}}{Pr(caught|F)}$ to estimate the overall pervasiveness of corporate fraud. The numerator in equation (1) is the observable incidence of fraud that is caught. The denominator in equation (1) is the detection likelihood. Table 6 reports observed caught frequencies in panel A, detection likelihood estimates from Table 3 in panel B, a baseline estimate of the pervasiveness of fraud across the measures of misconduct and alleged fraud in panel C, and a best estimate of the pervasiveness of fraud across the measures of misconduct and alleged fraud using the best estimate detection likelihood in panel D.

Focusing first on panel A, recall that Figure 1 showed that observed incidences of misconduct varies widely depending on the definition of corporate fraud and the time period of reference. Since fraud may be cyclical (Wang, Winton and Yu (2010)), we do not want to rely on a specific point in time, instead preferring to estimate pervasiveness over a full cycle of boom and bust years. The start in January of 1998

and the end point in December of 2005 are each almost exactly halfway through the respective expansion period; thus the period covers one full business cycle from mid-point to mid-point.[13] Panel A reports that auditor-detected securities frauds hover around 1%, with a peak of 1.2% in 2000-2001. AAER investigations average 2.7%, with a peak of 3.7% in the 2000-2001 period and a trough of 2.0% in 1998-99. Accounting violations measured by non-clerical restatements average 14.2% during the entire sample period, with a peak of 18.5% in 2002-2003 and a trough of 13.2% in 1998-1999. The broader SCAC securities fraud averages at 3.5%, with a peak of 5% in 2000-2001 and a trough of 2.9% in 1998-1999.

Using these observed misconduct averages in panel A and detection likelihood estimates from Table 3 reproduced in panel B, Table 6 panel C reports estimates of the pervasiveness of fraud. A concern with Panel C is that, as we discussed with regards to Table 3, some of the detection likelihood estimates are clearly more biased than others. The biases are always conservative, but such biases do not help with the precision of our fraud pervasiveness estimates. In particular, our experimental design of using AA's demise has the most power to identify the whole iceberg of hidden fraud (Assumption 2 holding as an equality) for auditor-detected securities fraud and restatements, as opposed to AAERs and the broader SCAC securities fraud cases, since no matter how rigorous the monitoring of the new auditors is, they will find it difficult to identify cases of price fixing such as Sotheby's. Thus, the best estimate detection likelihood is the average of the detection likelihood of auditor-detected and restatement measures, or 0.33, as highlighted before. We use this detection likelihood to calculate the best estimate of corporate misconduct across all of our fraud measures in panel D.

Our findings as to the pervasiveness of corporate fraud depend on the measure of misconduct we use. We find that in any year averaged across the business cycle, 3% of large corporations are committing severe financial misreporting that lead to an alleged securities fraud. Auditor-detected securities fraud is a subcategory of SCAC alleged securities fraud; thus it is not surprising that it has a low frequency. Such a

[13] According to the NBER Business Cycle Dating Committee, the two periods of expansion are March 1991- March 2001 and November 2001 – December 2007, while the one recession is March 2001-November 2001.

measure is interesting for the detection likelihood estimation, given that it maps well to our AA demise design, but of more interest here are the SCAC securities fraud at large.

We find that, during an average year over the business cycle, 11% of large corporations are committing a misrepresentation, an information omission, or another misconduct that lead to an alleged securities fraud claim settled for at least $3 million (with a 95% interval between 8 and 14%). This result from the SCAC data is our main estimate of the pervasiveness of corporate fraud, since SCAC are indeed (alleged) securities fraud. By using the AAER measure, we arrive to similar estimates: 8% of fraud pervasiveness, with a 95% range of 6%-11%. This magnitude is similar to the SCAC estimate, even though AAERs have lower observed frequencies because their existence requires the SEC to act. (Recall that the SEC failed to act on Madoff, despite six substantive complaints).[14]

Accounting violations, less severe than alleged securities fraud, are more prevalent, with an average annual pervasiveness of 43% (95% interval between 32% and 58%). We do not want to conclude from this estimate that each year 43% of large corporations commit a severe misreporting. To reach this conclusion we would need some more substantive filters to eliminate inconsequential misreporting. Nevertheless, this estimate does not bode well for the US auditing system. In spite of all the regulation, roughly half of the US financial statements suffer of misreporting more serious than pure clerical errors.

Comparison with Pervasiveness Estimates in the Literature

Taking the estimate of 11% as our main estimate of the pervasiveness of corporate fraud, we now ask how our estimates lines up with the literature. As the summary of the literature reported below shows, our estimate is at the low end of the pervasiveness of corporate fraud found in the literature. This is not surprising, since if Assumption 2 is violated, our estimate represents a lower bound.

We begin with evidence concerning the appetite for corporate misdoing for personal gain. Prior to Lie (2005), the existence of the options backdating practice had not been understood, and thus firms wanting to commit fraud in this manner could do so with little detection threat. Bebchuk, Grinstein and Peyer (2010)

[14] https://www.nytimes.com/2009/09/03/business/03madoff.html.

look back over the pre-2005 period and identify the percentage of publicly-traded firms from 1996-2005 in which CEOs or directors were 'lucky' in that they received option grants on the day of the month when the stock price was the lowest. By their estimate 12% of firms had such lucky CEOs, suggesting that the appetite for fraudulent behavior was present in at least 12% of firms.

That at least 12% of firms are committing some fraud is supported also by survey evidence and accounting prediction models. Dichev, Graham, Harvey and Rajgopal (2013) survey 169 CFOs of public companies, finding that 18% of firms manage earnings to misrepresent performance. Beneish, Lee and Nichols (2013) build a model based on Beneish (1999), a model that out of sample was able to predict successfully 71% of the most important accounting violations. They apply this model to estimate the pervasiveness of accounting manipulation and find that 18% of firms are fraudulent each year during the period 1997-2005.

How do the structural approaches to modelling the hidden iceberg of corporate misconduct line up with the evidence reported thus far? Wang, Winton and Yu (2010) examine financial fraud among the 3,297 IPOs from 1995-2005. While their main goal is to show that fraud is procyclical, their bivariate probit model produces predicted probabilities of engaging in fraud of 10-15%, very much in line with our estimates.[15] Similarly, in their flexible Bayesian priors approach to partial observability, Hahn, Murray, and Manolopoulou (2016) estimate a pervasiveness of SEC-investigated accounting misconduct to be 15%. Zakolyukina (2018) uses a structural model to explore detected and undetected GAAP manipulation (more in line with our accounting violation measure) and reports that 73% of CEOs manipulate their financials, with a detection likelihood of only 0.06.

IV. How Expensive Is Corporate Fraud?

IV.1 – Expected Cost of Fraud

[15] We infer this from Figure 1, predicted probability of fraud, and summary statistics on the distribution of industry EPS growth available in the internet appendix.

The finding that 11% of large corporations are committing fraud does not necessarily imply the economic cost of fraud is large. To draw this conclusion we need to estimate the cost of these corporate frauds. We do not innovate here, but we rely on existing estimates of the cost of fraud. As in the prior literature, we focus on the cost of fraud borne by the firms involved in fraud, ignoring spillovers to other firms.[16] We build our calculations off two prior papers.

Using an event-study methodology, KLM estimate the reputational loss of detected fraud at 25% of the equity value of a fraudulent firm. This cost, which is almost entirely due to a loss in reputation, represents the present value of the decline in expected cash flows as firms' investors, suppliers, and customers change the terms at which they interact with a fraudulent firm.

KLM's cost estimates are for detected fraud and do not necessarily apply to undetected fraud. Even when a fraud is not yet in the public domain, the firm incurs costs for two reasons. First, the fraud is unlikely to remain a secret for customers and employees, who will seek business relationships or employment elsewhere, demand a premium to remain, or take advantage of the fraud themselves. For example, Bernie Madoff's employees like Frank Di Pasquale were lavishly paid to ensure their silence. In addition, they stole money for themselves.[17] Second, the biggest cost is often the cover up. For 20 years, Japanese company Olympus was able to hide a $730m financial loss incurred in 1990, with a series of bad acquisitions and accounting tricks. The bad acquisitions alone cost $300m.[18]

It is hard to put a number on these costs. Yet, if we assume that, at least in the medium term, the stock market is strong-form efficient, the abnormal low returns of companies that are likely to have committed a fraud, but were never exposed as fraudulent, can provide an estimate of these hidden costs.

---

[16] Fraud has positive spillovers (when revealed it benefits competitors of the fraudulent firm) and negative ones (for example, employees who lose their jobs and are unable to find a job at an equal pay somewhere else). These spillovers are even more difficult to estimate that the costs we consider. We ignore them not because they are unimportant, but because we do not need them to reach our conclusion on the benefit of financial regulation from a purely investors' perspective. Since the negative spillovers are likely to be much larger than the positive ones, if we take a social welfare perspective our policy implications are likely to be strengthened if we add spillover effects to the cost-benefit analysis.

[17] https://www.justice.gov/sites/default/files/usao-sdny/legacy/2012/04/16/20090811dipascaliinformationsigned.pdf.

[18] https://www.nytimes.com/2011/12/09/business/deep-roots-of-fraud-at-olympus.html

Beneish, Lee and Nichols (2013) perform this exercise. They compare the annual buy-and-hold return of firms with a high probM score with that of firms with a low probM score. After controlling for a four-factor model, they estimate an annual 10.9% difference in returns. We take this underperformance as an estimate of the costs of undetected fraud.[19]

We are now in the position to compute the total cost of fraud. Our estimates suggest that 11% of firms are committing fraud. If the detected fraud (representing 33% of cases) costs the firm 25% of market value and the undetected fraud (representing 69% of cases) costs 10.9%, then the cost of an average fraud is 15.6% of firms' market capitalization. Thus, the cost of fraud is 1.7% of firm's equity value per year (i.e. fraud pervasiveness (11%) * loss for fraud (15.6%)). The 95% range is between 1.2% and 2.2%.

In 2004, the total capitalization of the U.S. equity market was $16 trillion.[20] Since on average 11% of firms are engaged in fraud, the annual fraud cost is $265 billion a year (between $195bn and $358bn with a 95% confidence). If we repeat the calculation at the end of 2020, the expected annual cost of fraud is $744 billion (between $548bn and $1,007bn with a 95% confidence).

IV.2 – An Application to Cost-Benefit Analysis

When a firm is 100% owned by one individual, the cost of fraud is fully internalized by the owner. In publicly traded companies, where equity is dispersed and insiders often own only a tiny fraction of the outstanding equity, this is not the case. These agency costs are one of the justifications for the introduction of regulations like Sarbanes Oxley (SOX). As an illustration of the wide applicability of our estimates, we sketch how it can be used for a cost-benefit analysis of SOX. Note that reducing agency costs is only one of the benefits of SOX. We are completely silent on other potential benefits, like reducing the lemon discount due to the asymmetry of information on the financial accounts.

[19] We use Beneish et al.'s figures as the cost of undetected fraud, conditional on committing fraud.
If Beneish et al.'s numbers are both selection and treatment, they represent the unconditional effect of fraud, not the conditional one. In such a case, our method would underestimate the reputational cost of fraud.
[20] https://statistics.world-exchanges.org/.

Hochberg, Sapienza and Vissing-Jorgensen (2009) exploit survey data collected by Finance Executives International (FEI) to arrive to an estimate of $3.8 million of compliance costs per firm, with costs increasing in the issuer's size.[21] To calculate the total compliance costs, we multiply this average compliance costs by the number of publicly traded firms in 2004 (5,226) to obtain an annual compliance cost of SOX of $19.9 billion per year. Note that faced with a regulatory mandate to increase monitoring in some ways, firms may cut down some of the monitoring they were previously doing in other ways (Karpoff and Lott (1993)). Thus, $19.9 billion represents an overestimate of SOX compliance costs.

The potential benefit of SOX is equal to the net reduction in the probability of fraud times the cost of fraud when it occurs. If instead of 11% of firms engaged in fraud, a policy change resulted in only 10% of firms engaged in fraud, the cost of fraud will drop by $25 billion (1% times 15.6%, which is the average cost of fraud, times 16 trillion, which is the 2004 total equity market capitalization). Thus, for the benefits of SOX regulation to exceed its costs, it would suffice that SOX were able to reduce the probability of fraud by one percentage point (a 9% reduction). Establishing whether this was indeed the case is beyond the scope of our paper. The purpose here is simply to illustrate how important is having an estimate of the pervasiveness of fraud, which enables us to derive an estimate of the total cost of fraud. With this estimate, we can easily assess what is the minimum level of effectiveness (in terms of reduction in the probability of starting a fraud) that any system of controls needs to achieve to justify its cost.

V. Conclusion

In this paper, we use the natural experiment provided by the sudden demise of a major auditing firm like Arthur Andersen to infer the fraction of corporate fraud that goes undetected. This detection likelihood is essential to quantify the pervasiveness of corporate fraud in the United States and to assess the costs that this fraud imposes on investors. We find that two out of three corporate fraud go undetected,

[21] See Hochberg, Sapienza and Vissing-Jorgensen (2009), Table 11 on page 571 for the costs by size categories that we use in our estimates.

implying that pre Sox 43% of large public firms were misreporting their financial accounts in a material way and 11% of the firms were committing securities fraud, imposing a cost of $265bn on investors.

These figures project a dismal picture of the effectiveness of financial auditing pre-SOX. Whether SOX reduced or eliminated the problem, our paper is unable to answer. Yet, the magnitude of the problem suggests that some action was warranted. Based on our estimates, we can also infer that if the new regulation reduced by 9% the probability of starting a fraud, its cost would be fully justified.

Appendix I: Dyck, Morse, and Zingales (2010) Filters to Eliminate Frivolous Fraud

First, DMZ(2010) restrict attention to alleged frauds in the period of 1996 -2004, specifically excluding the period prior to passage of the Private Securities Litigation Reform Act of 1995 (PSLRA) that was motivated by a desire to reduce frivolous suits and among other things, made discovery rights contingent on evidence. Second, they restrict attention to large U.S. publicly-traded firms, which have sufficient assets and insurance to motivate law firms to initiate lawsuits and do not carry the complications of cross-border jurisdictional concerns. In particular, they restrict attention to U.S. firms with at least $750 million in assets in the year prior to the end of the class period (as firms may reduce dramatically in size surrounding the revelation of fraud). Third, they exclude all cases where the judicial review process leads to their dismissal.[22] Fourth, for those class actions that have settled, they only include those firms where the settlement is at least $3 million, a level of payment previous studies suggested to divide frivolous suits from meritorious ones.[23] Fifth, they exclude those security frauds that Stanford classifies as non-standard, including mutual funds, analyst, and IPO allocation frauds.[24] The final filter removes a handful of firms that settle for amounts of $3 million or greater, but where the fraud, upon their reading, seems to have settled to avoid the negative publicity.[25]

[22] They retain cases where the reason for dropping the suit is bankruptcy for in this instance the cases could still have had merit but as a result of the bankruptcy status, plaintiff lawyers no longer have a strong incentive to pursue them.

[23] Grundfest (1995), Choi (2007) and Choi, Nelson, and Pritchard (2009) suggest a dollar value for settlement as an indicator of whether a suit is frivolous or has merit. Grundfest establishes a regularity that suits which settle below a $2.5 -$1.5 million threshold are on average frivolous. The range on average reflects the cost to the law firm for its effort in filing. A firm settling for less than $1.5 million is most almost certainly just paying lawyers' fees to avoid negative court exposure. To be sure, we employ $3 million as our cutoff.

[24] Stanford Class Action Database distinguishes these suits for the reason that all have in common that the host firm did not engage in wrongdoing. IPO allocation cases focus on distribution of shares by underwriters. Mutual fund cases focus on timing and late trading by funds, not by the firm in question. Analyst cases focus on false provision of favorable coverage.

[25] The rule they apply is to remove cases in which the firm's poor ex post realization could not have been known to the firm at the time when the firm or its executives issued a positive outlook statement for which they are later sued.

Appendix II: Calculation of Beneish's Probability of Manipulation Score (ProbM Score)

The components in the ProbM Score include days sales in receivables, gross margin, asset quality index, sales growth index, depreciation index, SGA index, leverage, and the ratio of accruals to assets. (Please refer to Beneish (1999) for motivation of how each of these subindices captures an aspect of manipulation.) To construct the ProbM Score, we use Compustat, data to construct the variable components following Beneish (1999) and apply his estimated coefficients.

The probability of manipulation, ProbM Score, of Beneish (1999) is calculated as follows:

$$\begin{aligned}\text{ProbM} = & -4.84 + 0.92 * \text{DSR} + 0.528 * \text{GMI} + 0.404 * \text{AQI} + 0.892 * \text{SGI} + 0.115 * \text{DEPI} \\ & + 0.172 * \text{SGAI} + 4.679 * \text{ACCRUALS} - 0.327 * \text{LEVI}\end{aligned}$$

The variable codes are defined as follows:

DSR = Days Sales in Receivables

GMI = Gross Margin Index

AQI = Asset Quality Index

SGI = Sales Growth Index

DEPI = Depreciation Index

SGAI = Sales, General and Administrative expenses Index

ACCRUALS - Total Accruals to total assets

LEVI = Leverage Index

References

Agresti, Alan, 1992, "A Survey of Exact Inference for Contingency Tables," *Statistical Science* Volume 7, Number 1 (1992), 131-153.

Agrawal, Anup and Sahiba Chada, (2005), "Corporate Governance and Accounting Scandals." *Journal of Law and Economics*, 48(2): 371-406.

Baker, Malcolm and Jeffrey Wurgler, (2006), "Investor Sentiment and the Cross-Section of Stock Returns," *Journal of Finance*, (61): 1645-1680.

Bebchuk, Lucien, Olav Grinstein and Urs Peyer, (2010). "Lucky CEOs and Lucky Directors," *Journal of Finance*, (65): 2363-2401.

Becker, Gary S., 1968, "Crime and Punishment: An Economic Approach," Journal of Political Economy, Vol. 76, No. 2 (Mar. - Apr.), pp. 169-217.

Beneish, M.D., (1999), "The Detection of Earnings Manipulation, *Financial Analysts* Journal, 24-36.

Beneish, M. D., Lee, C. M. C., & Nichols, D. (2013). Earnings manipulation and expected returns. Financial Analysts Journal, 69(2), 57-82.

Berger, Philip G. and Eli Ofek. (1995). "Diversification's Effect on Firm Value." *Journal of Financial Economics*, vol. 37(1), 39-65.

Blouin, J., B. Grein, B. Rountree, (2005). "An analysis of forced auditor rotation: the case of former Arthur Andersen clients." *Accounting Review*.

Burns, N. and S. Kedia. (2006). "The impact of performance-based compensation on misreporting." *Journal of Financial Economics* 79: 35-67.

Cahan and Zhang, (2006). "After Enron: Auditor Conservatism and Ex-Andersen Clients." *The Accounting Review*, 81: 49-82

Chen, Ken Y. and Jian Zhou. (2007). "Audit Committee, Board Characteristics and Auditor Switch Decisions by Andersen's Clients." *Contemporary Accounting Research*, 24 (4): 1085-1117.

Choi, Stephen J. (2007), "Do the Merits Matter Less after the Private Securities Litigation Reform Act?" *Journal of Law, Economics and Organization,* 23 (3): 598-626.

Choi, Stephen J., Karen K. Nelson and A.C. Pritchard. (2009). "The Screening Effect of the Securities Litigation Reform Act." *Journal of Empirical Legal Studies,* 6(1): 35-68.

Coates John C. and Suraj Srinivasan, 2014, "SOX after Ten Years: A Multidisciplinary Review", *Accounting Horizons*, September 2014, Vol. 28, No. 3, pp. 627-671

Coates IV, John C., 2015, Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications, *124 Yale Law Journal* 882 (2015).

Coffee, John C. (1986). "Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law through Class and Derivative Actions," *Columbia Law Review*, (86): 669-727.

Coffee, John C. (2006), *Gatekeepers: The Professions and Corporate Governance,* Oxford University Press: New York.

Dechow, P. M., R. G. Sloan, and A. Sweeney. (1996). "Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC." *Contemporary Accounting Research,* 13 (1): 1-36.

Dechow, Patricia M., Weili Ge, Chad R. Larson, Richard G. Sloan (2011), Predicting Material Accounting Misstatements," *Contemporary Accounting Research, 28: 17–82.*

Dichev, Ilia, Joh Graham, Campbell Harvey and Shiva Rajgopal, (2013). "Earnings Quality: Evidence from the Field," Working paper,

Dyck, Alexander, Adair Morse and Luigi Zingales. (2010). "Who Blows the Whistle on Corporate Fraud?" *Journal of Finance*. 65 (6): 2213-2254.

Dyck, Alexander, Natalya Volchkova, and Luigi Zingales (2008), "The Corporate Governance Role of the Media: Evidence from Russia," *Journal of Finance,* 63 (3): 1093-1136.

Efendi, Jap, Anup Srivastava, and Edward Swanson, (2007). "Why Do Corporate Managers Misstate Financial Statements? The Role of in-the-money Options and Other Incentives," *Journal of Financial Economics*, 85(3): 667-708.

Eisenberg, Macey, (2004). *"*Was Arthur Andersen Different? An Empirical Examination of Major Accounting Firm Audits of Large Clients" *Journal of Empirical Legal Studies,*

Feroz, E., K. Parek and V. Pastena, (1991). "The Financial and Market effects of the SEC's Accounting and Auditing Enforcement Releases," *Journal of Accounting Research,* 29: 107-148.

General Accounting Office, 2003, "Public Accounting Firms. Required Study on the Potential Effects of Mandatory Audit Firm Rotation," United States General Accounting Office, 2003.

General Accounting Office, (2008). "Audits of Public Companies: Continues Concentration in Audit Market for Large Public Companies Does Not Call for Immediate Action'," United States General Accounting Office.

Grundfest, Joseph A. (1995). "Why Disimply?" *Harvard Law Review*, (108): 740-741.

Hahn, P. Richard, Jared S. Murray, and Ioanna Manolopoulou (2016). "Bayesian Partial Identication Approach to Inferring the Prevalence of Accounting Misconduct." *Journal of the American Statistical Association*, 111 (513): 14-26.

Hennes, K, Leone, A., B. Miller, (2008). "The Importance of Distinguishing Errors from Irregularities in Restatement Research: The Case of Restatements and CEO/CFO Turnover," *The Accounting Review*, 83(6): 1487-1519.

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

Jensen, Michael and William Meckling, (1976). "The Theory of the Firm: Managerial Behavior, Agency Cost, and Ownership Structure." *Journal of Financial Economics* 3, no. 4 (October): 305-60

Karpoff, J.M. and J.R. Lott, Jr., (1993). "The Reputational Penalty Firms Bear from Committing Criminal Fraud." *Journal of Law and Economics,* (36): 757-802.

Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin, (2008). "The Cost to Firms of Cooking the Books," *Journal of Financial and Quantitative Analysis* 43 (3): p. 581–612.

Karpoff, Jonathan, Allison Koester, D. Scott Lee and Gerald Martin, (2017). "Proxies and databases in financial misconduct research," The Accounting Review, November 2017, Vol. 92, No. 6, pp. 129-163.

Kealey, Burch T, Ho Young Lee, and Michael T. Stein (2007) The Association between Audit-Firm Tenure and Audit Fees Paid to Successor Auditors: Evidence from Arthur Andersen. Auditing: A Journal of Practice & Theory: November 2007, Vol. 26, No. 2, pp. 95-116.

Kempf, Elisabeth and Spalt, Oliver G. (2019). Litigating Innovation: Evidence from Securities Class Action Lawsuits. European Corporate Governance Institute - Finance Working Paper

Krishnamurthy, S, J Zhou, N Zho (2006). "Auditor Reputation, Auditor Independence and the Stock Market Impact of Andersen's Indictment on Its Client Firms," Contemporary Accounting Research 23 (2), 465-490, 2006.

Krishnan, Gopal V., (2007). "Did earnings conservatism increase for former Andersen clients?" *Journal of Accounting, Auditing an*d *Finance*. 22(2): 141-163.

Krishnan, Jagan, K. Raghunandan, and Joon S. Yang (2007). "Were Former Andersen Clients Treated More Leniently Than Other Clients? Evidence from Going-Concern Modified Audit Opinions *Accounting Horizons*, 21 (4): 423–435.

Lie, Eric (2005), "On the Timing of CEO Stock Option Awards." *Management Science*, 51(5): 802-812.

Miller, Gregory S. (2006). "The Press as a Watchdog for Accounting Fraud." *Journal of Accounting Research* 44(5): 1001-1033.

Nagy, Albert L. (2005). Mandatory Audit Firm Turnover, Financial Reporting Quality, and Client Bargaining Power: The Case of Arthur Andersen. Accounting Horizons: June 2005, Vol. 19, No. 2, pp. 51-68.

Palmrose, Z.-V.; V. Richardson; and S. Scholz,. (2004). "Determinants of Market Reactions to Restatement Announcements." *Journal of Accounting and Economics* 37: 59–89.

Palmrose, Z-V., and S. W. Scholz. (2004). "The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements." *Contemporary Accounting Research*. 21 (1): 139-180.

Povel, Paul, Rajdeep Singh and Andrew Winton, 2007. "Booms Busts and Fraud," *Review of Financial Studies*, (20): 1219-1254.

Roe, Mark, (2002), "Corporate Law's Limits." *Journal of Legal Studies*, 31:.

Soltes, Eugene F., (2019), "The Frequency of Corporate Misconduct: Public Enforcement versus Private Reality Journal of Financial Crime (2019). Vol. 26 No. 4, p. 923-937.

Wang, Tracy Yue, (2013). "Corporate Securities Fraud: Insights from a New Empirical Framework", *Journal of Law, Economics and Organization*, 29 (3): 535-568.

Wang, Tracy Yue, Andrew Winton and Xiaoyun Yu, (2010). "Corporate Fraud and Business Conditions: Evidence from IPOs," *Journal of Finance,*

Yu, Frank (2008). "Analyst Coverage and Earnings Management,'' *Journal of Financial Economics (88):* 245–71.

Zakolyukina, Anastasia, (2018) "How Common Are Intentional GAAP Violations? Estimates from Dynamic Model," *Journal of Accounting Research*. 56(1). March 2018.

FIGURE 1: Frequency of Observed Corporate Fraud

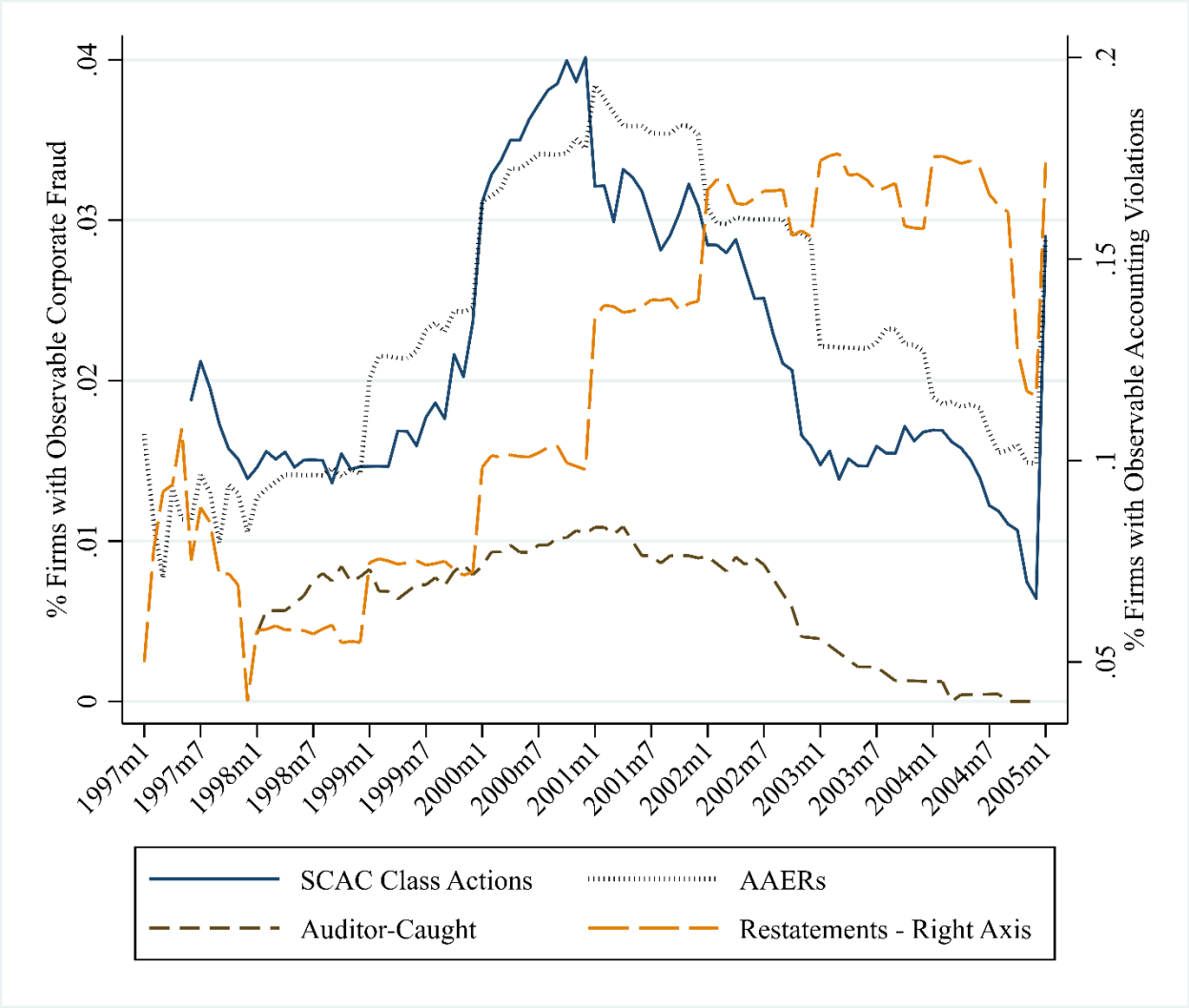


FIGURE 1: Frequency of Observed Corporate Fraud

Plotted are the prevalence rates of corporate fraud, for cases of corporate fraud that are eventually caught. The magnitudes represent the percentage of all U.S. large public corporations committing corporate fraud that is caught. Auditor-detected frauds are frauds in the DMZ (2010) sample of SEC 10(b) securities class actions which were triggered by an auditor, either by an auditor resignation or by the auditor issuing a qualified opinion and either the firm or analysts revealing the fraud. Restatements are from AuditAnalytics and refer to restatements triggered by accounting misapplication. AAERs are the SEC investigation releases used in Dechow, et al. (2011). The sample SCAC securities fraud is from Kempf-Splat (2019). Restatements are plotted with magnitudes on the right scale.

Table 1: Statistics: Arthur Andersen Clients (AA) and Non-AA Clients, 1998-2000

The sample is all Compustat firms with $750 million in assets in the 1998-2000 period. The designation of auditor is based on the calendar time designation in financial statements, not the experiment designation starting in Table 3. All statistics are collapsed to a single observation per firm, except where a firm switches auditors per the column designation. In those cases, we allow the firm to appear in multiple columns. Because of this time series collapsing, this table identifies more AA firms than in subsequent tables that focus on the collapse of AA. In Panel A, the p-values are from t-tests comparing means, and the z-values are from ranksum tests for the medians. Leverage is long term debt divided by assets. In Panel B, we present the industry distribution of firms and present Pearson distribution equivalence tests.

Panel A: Firm Characteristics

	AA Firms		All Non-AA				Big 5 Non-AA			
	Observations = 483		Observations = 1,792				Observations = 1,426			
	Mean	Median	Mean	p-value	Median	z-value	Mean	p-value	Median	z-value
Assets	6,998	1,991	9,802	0.120	1,982	0.920	9,088	0.148	2,040	0.594
Sales	3,645	1,471	4,297	0.212	1,384	0.784	4,431	0.144	1,433	0.593
EBITDA	568.6	196	669	0.318	182	0.174	677	0.233	201	0.794
Leverage	0.34	0.311	0.279	0.000***	0.256	0.000***	0.279	0.00***	0.260	0.000***

Panel B: Industry Distribution

	AA Firms Observations = 483 Percent of Distribution (*100)	All Non-AA Observations = 1792 Percent of Distribution (*100)	Big 5 Non-AA Observations = 1,426 Percent of Distribution (*100)
Agriculture	0.41	0.17	0.14
Banks & Insurance	12.42	21.09	19.57
Chemicals	3.11	4.58	4.70
Communication & Transport	16.15	10.44	9.40
Computers	4.35	10.04	9.68
Durable Manufacturing	10.35	12.39	13.32
Food & Tobacco	1.24	2.85	2.59
Lumber, Furniture, Printing	4.14	4.30	4.42
Mining & Construction	2.07	1.56	1.68
Pharmaceuticals	1.66	2.90	3.16
Refining Extractive	3.93	1.79	1.82
Retail & Wholesale	7.25	9.49	9.33
Services & Healthcare	11.39	7.37	7.36
Textile & Apparel	1.04	1.06	0.98
Utilities	20.50	9.99	11.85
	100	100	100
Pearson's Chi-Square test of Distribution Equivalence of non-AA Samples to AA Sample			
Statistic		105.2	85.8
P-value		0.000	0.000

Table 2: Did Arthur Andersen Clients Exhibit Excess Corporate Fraud in the Pre-period of 1998 - 2000?

Panel A: Univariate tests: Probability of manipulation scores (ProbM Score) are calculated using Compustat data following Beneish (1999), winsorized at 0.025 for extremes outside Beneish's range. Fraud score (FScore) are from Dechow-Ge-Larson-Sloan (2011), auditor-caught frauds are from Dyck-Morse- Zingales (2010), AAERs are from the CFRM at Berkeley-Haas, accounting restatements are from AuditAnalytics, SCACs are from Kempf-Splat (2019). The designation of auditor is the calendar time designation in the financial statements. Panel A reports cross sectional t-tests of differences in mean fraud rates for large corporations, comparing those firms with AA as auditor with all other firms, and all other firms with another Big 5 auditor.

	AA Firms		All Non-AA		p-value	Big 5 Non-AA		p-value
	Mean	Obs.	Mean	Obs.		Mean	Obs.	
ProbM Score (Beneish)	-1.711	435	-1.548	1,603	0.014	-1.601	1,269	0.096
ProbM > -2.2 threshold	0.585	435	0.624	1,603	0.078	0.613	1,269	0.217
Fscore (Dechow, et al. 2011)	3.217	420	4.378	1,559	0.301	3.664	1,246	0.650
Fscore > 1.4 threshold	0.323	420	0.413	1,559	0.000	0.391	1,246	0.005
Auditor-Caught Fraud	0.0186	483	0.0201	1,792	0.839	0.021	1,426	0.747
AAERs	0.0352	483	0.0458	1,792	0.313	0.0463	1,426	0.302
Restatements	0.164	483	0.137	1,792	0.134	0.139	1,426	0.183
SCAC Securities Fraud	0.0704	483	0.0854	1,792	0.287	0.0799	1,426	0.498

Panels B & C: Multivariate Tests: The dependent variables are accounting scores predicting fraud (Panel B) or realized corporate fraud (panel C). In columns (3), (6), (9) and (12), the dependent variable is an indicator score that takes the value of one when the indicated accounting score is above the relevant threshold of concern as identified in the original literature. Industry fixed effects, when included, are those of Table 1. Estimation is either by ordinary least squares (OLS) or by logit, as designated. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively. Standard errors in brackets are clustered at the firm level.

Panel B: Unobservable Measures of Fraud Dependent Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent var.:	Accounting Manipulation (ProbM Score)						Fraud Score (FScore)					
Estimation:	OLS ProbM	OLS ProbM	Logit ProbM>-2.2	OLS ProbM	OLS ProbM	Logit ProbM>-2.2	OLS FScore	OLS FScore	Logit FScore >1.4	OLS FScore	OLS FScore	Logit FScore >1.4
Arthur Andersen	-0.142***	-0.0224	0.11	-0.104*	0.00648	0.158	-0.384	0.179	0.0229	-0.155	0.318	0.0683
	[0.0514]	[0.0509]	[0.0960]	[0.0530]	[0.0518]	[0.0990]	[0.516]	[0.546]	[0.120]	[0.515]	[0.554]	[0.124]
Log Assets	-0.106***	-0.0938***	-0.109***	-0.112***	-0.106***	-0.113***	-0.184	0.00543	0.0411	-0.0758	0.117	0.0663
	[0.0164]	[0.0157]	[0.0341]	[0.0176]	[0.0166]	[0.0371]	[0.277]	[0.266]	[0.0388]	[0.296]	[0.283]	[0.0428]
Sales / Assets	0.352***	0.115**	0.366***	0.342***	0.121**	0.364***	-1.025***	-1.608***	-0.0978	-0.772***	-1.252***	-0.0859
	[0.0411]	[0.0498]	[0.108]	[0.0427]	[0.0512]	[0.117]	[0.252]	[0.373]	[0.0806]	[0.246]	[0.356]	[0.0879]
EBITDA / Sales	-2.495***	-2.849***	-3.155***	-2.394***	-2.604***	-3.345***	-23.28***	-22.19***	-2.909***	-19.25***	-17.70***	-3.280***
	[0.332]	[0.390]	[0.623]	[0.360]	[0.426]	[0.700]	[4.199]	[4.779]	[0.565]	[4.143]	[4.566]	[0.650]
LT Debt/Assets	-0.189	-0.107	0.00651	-0.234	-0.145	-0.203	-7.047***	-5.010***	-0.345*	-5.521***	-3.355***	-0.432*
	[0.141]	[0.145]	[0.188]	[0.156]	[0.157]	[0.209]	[1.486]	[1.208]	[0.209]	[1.481]	[1.147]	[0.236]
Sample		All Clients		Clients of Big 5 Auditors				All Clients		Clients of Big 5 Auditors		
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
Industry*Year FE	N	Y	N	N	Y	N	N	Y	N	N	Y	N
Observations	4,963	4,963	4,963	4,167	4,167	4,167	4,855	4,855	4,855	4,121	4,121	4,121
R2/ Pseudo R2	0.082	0.18	0.123	0.085	0.193	0.125	0.031	0.066	0.172	0.025	0.061	0.183

Table 2 (Continued)

Panel C: Observable Measures of Fraud Dependent Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent var.:	Auditor-Caught Securities Fraud				AAERs			
Non-AA sample:	--	--	Big 5	Big 5	--	--	Big 5	Big 5
Estimation:	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Arthur Andersen	-0.0375	0.158	-0.0865	0.0369	-0.143	0.0227	-0.145	-0.0213
	[0.395]	[0.426]	[0.397]	[0.430]	[0.290]	[0.306]	[0.297]	[0.313]
Log Assets	0.547***	0.619***	0.540***	0.594***	0.471***	0.562***	0.460***	0.536***
	[0.106]	[0.122]	[0.117]	[0.130]	[0.0884]	[0.0910]	[0.0965]	[0.0967]
Sales / Assets	0.297*	0.0184	0.282*	0.0406	0.419***	0.309**	0.425***	0.330**
	[0.152]	[0.212]	[0.167]	[0.215]	[0.103]	[0.131]	[0.110]	[0.138]
EBITDA / Sales	-0.883	-1.898	-1.494	-2.163	-0.409	-0.925	-0.653	-0.99
	[2.324]	[1.431]	[2.286]	[1.472]	[1.043]	[0.910]	[1.152]	[1.014]
LT Debt/Assets	0.715	0.873	0.305	0.368	0.48	0.879*	0.259	0.612
	[0.728]	[0.775]	[0.860]	[0.846]	[0.489]	[0.477]	[0.558]	[0.510]
Fixed Effects:								
Industry	N	Y	N	Y	N	Y	N	Y
Year	Y	Y	Y	Y	Y	Y	Y	Y
Industry*Year	N	N	N	N	N	Y	N	Y
Observations	5,465	4,993	4,641	4,228	5,465	5,105	4,641	4,331
R-squared	0.0586	0.103	0.0597	0.117	0.062	0.109	0.0602	0.0989

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Dependent var.:	Restatements				SCAC Securities Fraud 10b-5			
Non-AA sample:	--	--	Big 5	Big 5	--	--	Big 5	Big 5
Estimation:	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Arthur Andersen	0.119	0.189	0.127	0.181	-0.103	0.0118	-0.076	0.023
	[0.149]	[0.156]	[0.153]	[0.160]	[0.208]	[0.221]	[0.213]	[0.226]
Log Assets	-0.104*	-0.0548	-0.0818	-0.0428	0.127**	0.243***	0.141*	0.255***
	[0.0541]	[0.0550]	[0.0584]	[0.0585]	[0.0627]	[0.0676]	[0.0730]	[0.0775]
Sales / Assets	0.109	-0.0824	0.0979	-0.0836	-0.0947	-0.234	-0.0934	-0.205
	[0.0769]	[0.106]	[0.0875]	[0.117]	[0.116]	[0.157]	[0.129]	[0.167]
EBITDA / Sales	-1.208*	-1.359*	-1.214*	-1.193	-4.751***	-5.059***	-4.158***	-4.224***
	[0.643]	[0.711]	[0.713]	[0.784]	[1.154]	[1.268]	[1.251]	[1.442]
LT Debt/Assets	-0.0784	0.0165	0.0233	0.0821	-0.810**	-0.231	-0.878*	-0.206
	[0.262]	[0.281]	[0.293]	[0.310]	[0.411]	[0.406]	[0.478]	[0.461]
Fixed Effects:								
Industry	N	Y	N	Y	N	Y	N	Y
Year	Y	Y	Y	Y	Y	Y	Y	Y
Industry*Year	N	Y	N	Y	N	Y	N	Y
Observations	5,465	5,465	4,641	4,641	5,465	5,237	4,641	4,152
R-squared	0.0133	0.0402	0.0134	0.0388	0.0493	0.114	0.0434	0.11

Table 3: Detection Likelihood Estimates

The table presents the main Poisson-Binomial estimates of the detection likelihood (odds-ratio) using the AA experiment. The sample is U.S. publicly traded corporations with more than $750 million in assets. A firm is identified as a former Arthur Andersen (AA) client if it was audited by AA in the year 2001. In all cases, the frauds considered are those that started prior to the watershed date of December 1, 2001 and which are caught after the watershed date and before the end of 2003. Auditor-detected frauds are frauds in the DMZ (2010) sample of class action frauds which were detected by an auditor, either by an auditor resignation or by the auditor issuing a qualified opinion and either the firm or analysts revealing the fraud. Restatements are from AuditAnalytics and refer to restatements triggered by accounting mis-application. AAERs are the SEC investigation releases used in Dechow, et al. (2011). The sample SCAC Securities Fraud is from Kempf-Splat (2019). Panel A reports Poisson-Binomial coefficient estimates and the asymptotic and exact p-values for the AA coefficient. The detection likelihood is calculated directly in Poisson estimation as *odds ratio = 1/exp*(*coefficient*). Panel B reports maximum likelihood estimates of the system of all Poisson equations estimated simultaneously, with the coefficient on AA constrained to be the same. The first column of Panel B includes all 4 fraud definitions, and the second column, only auditor-detected fraud and restatements, per the discussion in the text that the experiment is most precise for these definitions. Also included in Panel B are the confidence intervals for the detection likelihood estimates.

Panel A: Exact Poisson Estimation of Poisson-Binomial Draws

	Severe Financial Reporting Violations		Accounting Violations	Securities Fraud
	Auditor-Detected	AAERs	Restatements	SCAC
Former Arthur Andersen Clients				
Coefficient	1.225	0.655	1.088	0.755
Detection Likelihood	0.294	0.520	0.337	0.470
Asymptotic p-value	0.008	0.037	0.000	0.011
Exact p-value	0.026	0.067	0.000	0.024
Observations (# of clients)	2,757	2,757	2,757	2,757
AA clients	353	353	353	353
Non-AA clients	2,404	2,404	2,404	2,404
Total Number of Frauds	21	59	168	63
Pseudo R-squared	0.0067	0.0067	0.0278	0.0094

Panel B: Simultaneous, Maximum Likelihood Estimation of System of Poisson-Binomial Equations, Constrained to Have Equal Coefficients across Equations

	Equation Dependent Variables: Auditor-Detected Fraud, AAERs, Restatements, SCAC		Equation Dependent Variables: Auditor-Detected Fraud, Restatements	
Former Arthur Andersen Clients				
Coefficient	0.957		1.104	
Detection Likelihood	0.384		0.332	
95% Confidence Interval	0.300	0.493	0.243	0.452
Asymptotic p-value	0.000		0.000	
Observations (# of clients)	2757 in each of 4 equations		2757 in each of 2 equations	

Table 4: Detection Likelihood Robustness: Small Corporations & Placebo

The table presents Poisson-Binomial estimates of the detection likelihood (odds-ratio) using the AA experiment. Panel A reports estimates based on frauds in U.S. publicly traded corporations with *less* than $750 million in assets, and thus differs from Table 3 results that are based on firms with *more* than $750 million in assets. Panel B is similar to Table 3 in using as a sample U.S. publicly traded corporations with more than $750 million in assets, but differs from Table 3 in that the detection period is a placebo period of frauds detected between 2004-2006 (Table 3 Panel A focused on frauds caught after the watershed date and before the end of 2003). A firm is identified as being a former Arthur Andersen (AA) client if it was audited by AA in the year 2001. Auditor-detected frauds are frauds in the DMZ (2010) sample of class action frauds which were detected by an auditor either by an auditor resignation or by the auditor issuing a qualified opinion and either the firm or analysts revealing the fraud. Restatements are from AuditAnalytics and refer to restatements triggered by accounting mis-application. AAERs are the SEC investigation releases used in Dechow, et al. (2011). The sample SCAC Securities Fraud is from Kempf-Splat (2019). Both panels report Poisson-Binomial coefficient estimates and asymptotic p-values for the AA coefficient. Panel B also reports exact p-values. (The sample is too large for computation of this statistics for restatements in Panel B.) The detection likelihood is calculated directly in Poisson estimation as *odds ratio = 1/exp* (*coefficient*).

Panel A: Small Corporations

	AAERs	Restatements	SCAC
Former Arthur Andersen Clients			
Coefficient	0.438	0.867	1.241
Detection Likelihood	0.645	0.420	0.289
p-value	0.322	0.000	0.000
Observations (# of clients)	7,553	7,553	7,553
AA clients	772	772	772
Non-AA clients	6,781	6,781	6,781
Total Number of Frauds	40	394	46
Pseudo R-squared	0.0018	0.0134	0.0134

Panel B: Placebo of Detection Time Period = 2004-2006

	AAERs	Restatements	SCAC
Former Arthur Andersen Clients			
Coefficient	0.021	-0.117	0.460
Detection Likelihood	0.979	1.124	0.631
Asymptotic p-value	0.961	0.429	0.190
Exact p-value	1.000	--	0.267
Observations (# of clients)	2,445	2,445	2,445
AA clients	314	314	314
Non-AA clients	2,131	2,131	2,131
Total Number of Frauds	46	450	53
Pseudo R-squared	0.0000	0.0003	0.0030

Table 5: Detection Likelihood Robustness: Industry and Region Subsampling

The table presents, for robustness, Poisson-Binomial estimates of the detection likelihood (odds-ratio) using the AA experiment for various sub-samples. In Panel A, we present results where we exclude, one-by-one, industries in which AA was less represented or more represented, as indicated by the row labels. In Panel B, we present results where we exclude firms from certain regions or states, as indicated in the row labels. The overall sample is U.S. publicly traded corporations with more than $750 million in assets. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level from Poisson p-values, testing that the AA coefficient is significantly different from zero, or equivalently, that the odds ratio is different from unity.

Panel A: Robustness in Excluding Industries

	Counts: AA [Non-AA]	Detection Likelihood			
		Auditor Detected	AAERs	Restatements	SCAC
Full Sample (Table 3)	353[2404]	0.294**	0.520*	0.337***	0.470**
<u>Excluding Less Represented Industries</u>					
Banks & Insurance	313[1592]	0.294**	0.608	0.355***	0.544*
Computers	340[2236]	0.329**	0.593	0.320***	0.468**
Retail &Wholesale	332[2231]	0.276***	0.458**	0.313***	0.456***
<u>Excluding More Represented Industries</u>					
Refining & Extractive	338[2360]	0.286***	0.496**	0.327***	0.481**
Communication &Transport	306[2230]	0.594	0.656	0.357***	0.524**
Utilities	276[2214]	0.229***	0.536*	0.374***	0.499**
Services & Health	313[2261]	0.237***	0.426***	0.334***	0.425***

Panel B: Robustness in Excluding Regions and States

	Counts: AA [Non-AA]	Detection Likelihood			
		Auditor Detected	AAERs	Restatements	SCAC
Full Sample (Table 3)	353[2404]	0.294**	0.520*	0.337***	0.470**
<u>Excluding Regions:</u>					
Southeast	267[1928]	0.277**	0.615	0.353***	0.396***
Northeast	290[1822]	0.398	0.385***	0.311***	0.525*
Midwest	258[1918]	0.247***	0.523*	0.399***	0.553
Mountain	333[2325]	0.266***	0.496**	0.349***	0.430***
Southwest	307[2132]	0.247***	0.537*	0.301***	0.422***
West	323[2067]	0.365**	0.568*	0.316***	0.529**
<u>Excluding Large States:</u>					
Texas	320[2172]	0.253***	0.549*	0.323***	0.442***
California	328[2137]	0.358**	0.558*	0.310***	0.519**
New York	335[2186]	0.306**	0.460**	0.338***	0.507**

Table 6: How Pervasive is Corporate Fraud?

The table presents estimates of corporate fraud pervasiveness in publicly traded US firms with more than $750 million in assets, along with the inputs to generate these estimates. Panel A presents observable caught fraud frequencies. These are calculated based on firms that engaged in misconduct that took place sometime in the indicated data range specified by the columns and were caught, using the corporate fraud or misconduct measure indicated in the rows. The percentages in Panel A are based on all corporate frauds, not just those aligning with the AA experiment timing. Panel B presents the detection likelihood estimates from Table 3. Panels C and D use the Panel A and B estimates to calculate the prevasiveness of corporate fraud defined as the observed caught fraud divided by the detection likelihood. Panel C provides baseline estimats using the detection liklihood for that fraud measure. Panel D presents our best estimate, using the average of the detection likelihood of auditor-detected and restatement measures.

Panel A: Annual Ongoing Misconduct Eventually Caught as % of Firms

	All		*1998-1999*	*2000-2001*	*2002-2003*	*2004 -2005*
Auditor-Detected Fraud	1.00%	[n=56]	1.07%	1.21%	0.71%	--
AAERs	2.69%	[n=141]	2.01%	3.66%	2.69%	2.38%
Restatements (filtered)	14.19%	[n=793]	7.40%	13.21%	18.47%	17.67%
SCAC Securities Fraud	3.52%	[n=283]	2.89%	4.97%	2.93%	3.29%

Panel B: Detection Likelihood Estimates from Table 3

	Detection Likelihood
Auditor-Detected Fraud	0.294
AAERs	0.520
Restatements	0.337
SCAC Securities Fraud	0.470
Best Estimate (average of auditor-detected and restatements)	0.332

Panel C: Baseline Estimates Pervasiveness of Corporate Fraud

	Fraud Pervasiveness = Observed Fraud / Detection Likelihood				
	All	*1997-1999*	*2000-2001*	*2002-2003*	*2004 -2005*
Auditor-Detected Fraud	3.4%	3.7%	4.1%	2.4%	--
AAERs	5.2%	3.9%	7.0%	5.2%	4.6%
Restatements	42.1%	22.0%	39.2%	54.8%	52.4%
SCAC Securities Fraud	7.5%	6.1%	10.6%	6.2%	7.0%

Panel D: Best Estimate of Pervasiveness of Corporate Fraud

	Fraud Pervasiveness = Observed Fraud / Detection Likelihood Best Estimate				
	All	*1997-1999*	*2000-2001*	*2002-2003*	*2004 -2005*
Auditor-Detected Fraud	3.0%	3.2%	3.7%	2.1%	--
AAERs	8.1%	6.1%	11.0%	8.1%	7.2%
Restatements	42.8%	22.3%	39.8%	55.7%	53.3%
SCAC Securities Fraud	10.6%	8.7%	15.0%	8.8%	9.9%

Appendix Table A1: Propensity Score Matched, Collapsed Tests: Were AA Clients Committing Corporate More Fraud in the Pre-Period?

The sample is a cross section of firms with $750 million in assets and a Big 5 auditor during 1998-2000. Estimation in Panel A is limited to only Texas headquartered firms. Panel B replicates panel A for the entire United States. Estimation is a propensity-matched estimation. For each AA client firm, we find a non-AA matched firm in the same two digit SIC code, matched on the propensity of assets, sales/assets, EBITDA/sales and leverage to predict being an AA client. Once matched, the dependent variables are measures of uncaught fraud (columns 1-4) or caught fraud (columns 5-12) as follows: the probability of manipulation score (prob-M score) of Beneish (1999) (columns 1-2), fraud score (Fscore) of Dechow, et al (2011) (columns 3-4), auditor-detected fraud suits (DMZ) (columns 5-6), AAERs (columns 7-8), AuditAnalytics accounting restatements (columns 9-10), and SCAC Securities Fraud from Kempf-Splat (2019). All regressions include industry fixed effects (SIC2) and year fixed effects. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively. Robust (for OLS) standard errors are in brackets.

Panel A: Texas Only Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Manipulation (ProbM)		Fraud Score (Fscore)		Auditor-Detected		AAERs		Restatements: Acctng.		SCAC Securities Fraud	
Arthur Andersen	0.206	0.183	0.302	0.769	0.0179	0.0144	0.0446	0.0393	0.0446	0.0463	0.0179	0.0158
	[0.242]	[0.230]	[1.908]	[1.702]	[0.0177]	[0.0140]	[0.0399]	[0.0352]	[0.0605]	[0.0601]	[0.0215]	[0.0190]
Log Assets		-0.0697		-0.492		0.0178		0.0223		0.017		0.0092
		[0.0796]		[0.737]		[0.0168]		[0.0272]		[0.0318]		[0.0114]
Sales / Assets		0.163		-0.952		-0.00122		-0.00826		-0.0251*		-0.00538
		[0.0997]		[1.501]		[0.00238]		[0.0101]		[0.0138]		[0.00447]
EBITDA / Sales		-4.224***		-37.74*		0.0815		-0.454**		-0.186		-0.119
		[1.311]		[19.61]		[0.0794]		[0.223]		[0.300]		[0.104]
LT Debt/Assets		-0.644		-22.19		0.0616		-0.200**		0.0756		-0.0944
		[0.783]		[13.58]		[0.0577]		[0.0918]		[0.142]		[0.0819]
Observations	161	161	164	164	188	188	188	188	188	188	188	188
R-squared	0.005	0.089	0.000	0.101	0.009	0.053	0.011	0.077	0.005	0.021	0.005	0.032

Panel B: United States Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Manipulation (ProbM)		Fraud Score (Fscore)		Auditor-Detected		AAERs		Restatements: Acctng.		SCAC Securities Fraud	
Arthur Andersen	-0.00318	-0.0155	-0.191	-0.098	0.00347	0.00458	-0.00867	-0.00732	0.0191	0.0197	-0.00433	-0.00236
	[0.0752]	[0.0711]	[0.717]	[0.715]	[0.00532]	[0.00546]	[0.0107]	[0.0107]	[0.0173]	[0.0173]	[0.0109]	[0.0108]
Log Assets		-0.118***		-0.215		0.0145***		0.0229***		-0.000107		0.0103
		[0.0273]		[0.351]		[0.00536]		[0.00774]		[0.00794]		[0.00643]
Sales / Assets		0.296***		-0.466		0.0055		0.0202**		0.00387		-0.00274
		[0.0567]		[0.373]		[0.00472]		[0.00939]		[0.0116]		[0.00568]
EBITDA / Sales		-3.601***		-25.31***		-0.0697*		-0.0976		-0.253**		-0.360***
		[0.604]		[8.869]		[0.0396]		[0.0615]		[0.123]		[0.0949]
LT Debt/Assets		-0.131		-6.316**		0.0182		0.0173		0.0104		-0.0422
		[0.200]		[2.841]		[0.0151]		[0.0274]		[0.0465]		[0.0320]
Observations	1,628	1,628	1,640	1,640	1,937	1,937	1,937	1,937	1,937	1,937	1,937	1,937
R-squared	0	0.097	0	0.031	0	0.026	0.001	0.025	0.001	0.006	0	0.027

Appendix Table A2: Industry and Region of Corporate Fraud Cases in Experiment

The sample is U.S. publicly traded corporations with more than $750 million in assets. Auditor-detected frauds are frauds in the DMZ (2010) sample of class action frauds which were detected by an auditor, either by an auditor resignation or by the auditor issuing a qualified opinion and either the firm or analysts revealing the fraud. Restatements are from AuditAnalytics and refer to restatements triggered by accounting mis-application. AAERs are the SEC investigation releases used in Dechow, Ge, Larson and Sloan (2011). SCAC are securities fraud class actions from Kempf-Splat (2019). In all cases, the frauds considered are those starting prior to the watershed date of October 31, 2001 and being detected before the end of the 2003 period. Panel A presents cross-sectional correlations. Panel B and C report fraud incidence by industry or geography.

Panel A: Correlations among Fraud Measures

	Auditor Detected	AAER	Restatements	SCAC
	Sample: 2757 firms (1 observation per firm) over the period 11/2001-12/2003			
Severe Financial Reporting Violations				
Auditor-Detected Fraud	1.000			
AAERs	0.152	1.000		
Restatements				
Restatements (filtered)	0.104	0.129	1.000	
Securities Fraud at large				
SCAC Securities Fraud	0.348	0.190	0.120	1.000
	*Note: All Correlations are statistically significant at < 1% level			

Panel B: Corporate Fraud Incidence by Industry

	Auditor Detected	AAER	Restatements	SCAC
	Sample: 2757 firms (1 observation per firm) over the period 11/2001-12/2003			
Agriculture	--	--	0.5%	--
Banks Insurance	33.3%	24.8%	24.0%	24.6%
Chemicals	--	--	2.4%	1.5%
CommunicationTransport	20.8%	12.0%	12.9%	13.1%
Computers	8.3%	15.4%	9.8%	14.6%
Durable Manufacturing	--	9.4%	5.8%	3.1%
Food Tobacco	--	2.6%	3.7%	--
Lumber Furniture Printing	--	1.7%	1.1%	3.8%
Mining Construction	--	--	2.1%	--
Pharmaceuticals	6.3%	1.7%	1.3%	10.0%
Refining & Extractive	--	0.9%	2.9%	3.1%
Retail & Wholesale	4.2%	12.0%	9.8%	2.3%
Services & Health	8.3%	10.3%	5.5%	3.1%
Textile Apparel	--	--	1.8%	--
Utilities	18.8%	9.4%	16.4%	20.8%
Total	100%	100%	100%	100%

Panel C: Corporate Fraud Incidence by Region

	Auditor Detected	AAER	Restatements	SCAC
	Sample: 2757 firms (1 observation per firm) over the period 11/2001-12/2003			
Midwest	17%	26%	25%	25%
Mountain	4%	2%	6%	5%
Northeast	33%	30%	21%	34%
Southeast	29%	20%	15%	15%
Southwest	13%	10%	16%	12%
West	4%	12%	17%	8%
Territories	--	--	1%	--
Total	100%	100%	100%	100%