HOW PERVASIVE IS CORPORATE FRAUD?

Alexander Dyck
University of Toronto

Adair Morse
University of Chicago, University of California at Berkeley, & NBER

Luigi Zingales*
University of Chicago, NBER, & CEPR

February 2013

ABSTRACT

We estimate what percentage of firms engage in fraud and the economic cost of fraud. Our estimates are based on detected frauds, and frauds that we infer are started but are not caught. To identify the 'iceberg' of undetected fraud we take advantage of an exogenous shock to the incentives for fraud detection: Arthur Andersen's demise, which forces companies to change auditors. By assuming that the new auditor will clean house, and examining the change in fraud detection by new auditors, we infer that the probability of a company engaging in a fraud in any given year is 14.5%. We validate the magnitude of this estimate using alternative methods. We estimate that on average corporate fraud costs investors 22 percent of enterprise value in fraud-committing firms and 3 percent of enterprise value across all firms.

---

* We thank Patricia Dechow, Katherine Guthrie, Phil McCollough, Joseph P. Weber, Michael Weisbach as well as participants at the University of Illinois Symposium on Auditing Research, the European Finance Association, the University of Chicago, the Rotman School at the University of Toronto, Berkeley-Haas Accounting, 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.

Until recently, the United States was deemed the corporate governance standard towards which other countries aspired. The major wave of corporate scandals that emerged at the beginning of the millennium deeply shook this confidence. How was it possible for companies like HealthSouth to falsify its financial statements for 11 years without notice, or WorldCom to transform 3.8 billion of expenses into capital investments, or Enron to allow managers to enrich themselves while hiding billions of liabilities? Do these examples just reflect a few rotten apples, or are they instead the tip of the proverbial iceberg?

The answer to this question is not just intrinsically interesting, but it is extremely important. If we knew the frequency and cost of frauds this would help investors and boards to tailor resources to mitigate the scope of the problem. It would also provide a fact base for reforms, such as the legislative reforms in Sarbanes Oxley and Dodd-Frank. If there are just a few rotten apples, large-scale intervention might be a waste of energy and resources. As the old saying goes, "If it ain't broke, don't fix it". But if these examples are the tip of the iceberg, then further interventions to fix the problems might be warranted.

Frauds we are interested in have a few key characteristics. As is clear in the examples above, frauds involve misrepresentations, concealment, or nondisclosure. They also need to be important rather than minor. This is obviously more difficult to define. It is captured in part by the accountants' term 'materiality', or that "knowledge of the matter would be likely to influence the user of the financial or other statements under consideration." They could involve lying about the past, with financial misrepresentations, or lying about the future by concealing or not disclosing the status of projects designed to deliver future growth. Less important for us is whether those who are involved have an intent to deceive, a criteria at the heart of most legal cases.

Prior research provides some indicators that could be used to size up the pervasiveness of fraud, but this research has at least two important limitations. The first problem is that the major databases that are used to explore fraud have features leading to meaningful biases. Karpoff, Koester, Lee and Martin (2012) (hereafter KKLM) focus on this issue, identifying the four datasources that researchers have relied upon, the major papers that use each of the databases, and the limitations of each database. At the risk of oversimplification, two central limitations are over-restrictiveness, or over-inclusion.

Focusing only on firms where there is a Securities and Exchange Commission Accounting and Auditing Release (AAER), for example, is likely to lead to a sample that is restricted and excludes some cases that we see as frauds. [1] AAER firms are serious cases brought by the SEC after investigation. Befitting their seriousness, AAER firms are more likely than firms in all other databases to be associated with legal charges of financial fraud. But this comes at the cost of omitting other cases that likely are material. The SEC after all has a limited budget, and they don't have incentives to go after all frauds, rather those that are visible and less costly to detect. Moreover, by definition enforcement actions involve accountants,[2] and focus on financial misrepresentations. Non-disclosure of the status of projects delivering future growth, for example, has clear value consequences, but would not be captured by AAERs.

Alternatively, focusing on firms where there has been a financial restatement, using either the sample produced by the General Accounting Office, or the sample from Audit Analytics, is likely over inclusive of cases of financial misrepresentation in including many cases that are unlikely to be material. Hennes et al. (2008), for example categorize 73.6% of GAO restatements as unintentional misapplications of GAAP accounting.

A second limitation with existing studies for our purposes is that they do not focus on the iceberg of undetected fraud. Sizing up the full extent of the fraud problem in US corporations requires an effort to identify frauds being committed in corporations that remain undetected. It is likely that researchers have avoided this, in part as a result of a need for strong identifying assumptions to go from observed data to infer the extent of unobserved. A notable exception to this is Wang (2011) who tackles this problem but does not produce an estimate of the scope of detected and undetected fraud.

[1] Examples that have focused on Accounting, Auditing and Enforcement Releases, include Dechow, Sloan and Sweeney (1996), Miller (2006), and

[2] See AAER-1, 1982 SEC LEXIS 2565, May 17, 1982 and discussion in KKLM.

In this paper we provide an answer to the question of the pervasiveness of fraud. We use the dataset of frauds from Dyck, Morse and Zingales (2010), (hereafter DMZ) who developed a comprehensive sample of frauds from Security Class Action cases. Their frauds include those involving financial misrepresentations, but importantly are not limited to those. The Securities Class Action data, when combined with DMZ extra steps, performs well according to the criteria of KKLM. This sample relies on the fact that the security class action system provides strong incentives (for attorneys and shareholders) to file suit whenever a fraud that is likely to have a material impact is revealed. As a result the sample is unlikely to suffer from problems of over restrictiveness. For large companies, it is highly unlikely that detected frauds exist without a corresponding class action suit. Of more concern is that the sample will be over inclusive. DMZ apply rigorous filters to eliminate likely frivolous suits. [3]

With this dataset we tackle the question of unobservable fraud. We appeal to basic probability rules for guidance of going from observed data of the joint event of engaging in fraud and being caught, to our actual variable of interest, the probability of engaging in fraud regardless of whether they are caught or not. The idea is really quite simple. What is observed is the probability that managers engage in a fraud and that they get caught, Pr(*engage,caught*). The unconditional probability of engaging in a fraud which is what we are ultimately interested in is Pr(*engage*). This is the product of the detection likelihood Pr(*caught*|*engage*) and the observed probability of engaging and getting caught. Thus, if we knew the detection likelihood we could easily calculate the probability of engaging in fraud. Our identification strategy exploits circumstances in which the likelihood of being caught increases significantly. By comparing the differences in detection in this special circumstance, and in normal circumstances, we produce an estimate of the iceberg (i.e. the normal detection likelihood) and then it is a short step to estimate the unconditional pervasiveness of engaging in fraud.

[3] As with all databases, the SCAC also has limitations that KKLM discuss. These include the fact that the SCAC database omits 9.4% of cases that prompted SEC enforcement and had security class action filings. This makes the results we arrive at more conservative. The SCAC omission rate is the lowest across the four databases.

Our primary test takes advantage of the natural experiment created by the demise of Arthur Andersen that forced all firms that previously had Andersen as their auditor to seek another auditor. This forced auditor turnover enhances the incentives of new auditors to be active. When we restrict our attention to firms that had Arthur Andersen as an auditor and were forced to change auditors, we find that the incidence of fraud detection by auditors goes up by a multiple of close to four. This gives a sense of how much undetected fraud exists more generally, with the iceberg being 3 times bigger under the water than above the water. Taking this estimate, and applying it with some additional assumptions, we arrive at our best estimate that 14.5% of large publicly traded corporations engage in fraud. We also use this experiment to produce a very conservative lower bound estimate, which we find to be 5.6%.

To validate these results, we introduce two additional tests and compare our results with others in the literature. Our second test uses a similar approach but applied to a larger sample of fraud detectors taking advantage of details on each fraud revelation from DMZ. For example, they find that 15% of the frauds are brought to light by analysts. Hence, conditional on a fraud being committed, it is reasonable to conjecture that the probability a fraud is revealed is a positive function of the number of analysts following a company. This conjecture is supported by the data. Hence, in companies that have more analysts following them, if a fraud is perpetrated, it is more likely to emerge. We exploit a number of other cross sectional and time series variations in our data that heighten incentives for fraud detections, and produce an estimate of the unconditional probability of engaging in fraud. These results lie within the range suggested by the Arthur Andersen experiment. Our third test employs a different approach, looking at survey results on whether corporate actors were asked to engage in illegal activity, in a setting where there is little incentive not to reveal the truth, producing an estimate of 14.8 percent.

We then take the results one step further and offer an estimate of the social costs of fraud. Even in circumstances where there is no prior information leakage, the change in market value on the day the fraud comes to public attention is triggered is not a comprehensive measure of the social costs of fraud. Frauds are often committed to cover up negative news, which would have been revealed to the market earlier in absence of fraud. In other words, the stock price, and thus the stock drop, at the time of revelation are both too large relative to fundamentals. The amount of damages alleged in legal suits is also not a good measure of such costs, because many of the dollar losses are transfers rather than social losses.

We construct a new measure of the cost of fraud, which we define to be equal to the difference between the enterprise value after the fraud is revealed and what the enterprise value of the company would have been in the absence of fraud. We construct this hypothetical value by making projections from the pre-fraud period, assuming the trajectory would have followed that of other firms in the same industry. Using this approach, we estimate that the median loss is 20.4 percent of the enterprise value of our fraud companies, using firms' enterprise value prior to the beginning of fraud as the benchmark.
We also take advantage of a conceptually similar approach to measure the cost of fraud by Karpoff, Lee and Martin (2008) that is well suited to situations with financial frauds, which leads to an estimate of the social costs of fraud of 21.8% of enterprise value.

Putting the estimate of the extent of fraud with this estimate of the cost per firm of fraud, we produce an estimate of the social cost of fraud for these firms as a percentage of their enterprise value. This price tag is 3% of enterprise value of all large corporations.

The rest of the paper proceeds as follows. Section I describes the data and provides a baseline conservative estimate of fraud pervasiveness based on frauds that are caught. Section II describes our main identification methodology and information about the relevance of the Arthur Andersen demise for our estimation procedure. Section III provides results on fraud pervasiveness from the Arthur Andersen experiment. Section IV introduces and describes two tests that validate these estimates, based on a larger sample of fraud detectors and survey evidence, as well as related literature. Section V provides costs estimates and we conclude in section VI.

I. Data on Caught Fraud Incidence

To establish a baseline of the pervasiveness of corporate fraud in U.S. publicly-traded firms, we start with the DMZ sample of caught frauds. DMZ identify frauds as firms subject to securities class action lawsuits, as compiled in the from the Stanford Securities Class Action Clearinghouse (SCAC).

DMZ argue that this sample is close to the population of caught fraud for large (over $750 million in assets) publicly-traded companies because of the incentive structure for law firms. Class action law firms have automated the mechanism of filing class action suits such that specialist attorneys start searching for a cause to file a suit every time a large negative shock to share prices occurs in large corporation. Since stock prices drop following revelation of most serious corporate frauds, it is highly unlikely that a corporate fraud of any magnitude would emerge without a subsequent class action suit being filed (Coffee, 1986).

The biggest potential problem with using class action data is not that it misses important frauds, but rather that it might be over-inclusive in including frivolous allegations. DMZ use a filtering process, summarized in the appendix of this paper, to remove this concern. The gist of the screening is to restrict attention to data after 1996, after a law change made the courts became more stringent about evidence for certification. They then limit the data to cases that are not dismissed, presuming this filters out some frivolous cases. And they limit non-dismissed cases by excluding cases with low settlements, based on guidance from the legal literature as to what settlement amounts constitute nominal payments to make the suit go away.

It is worth noting that while we use the term frauds they are better thought of as 'alleged frauds'. Security class action cases are almost always settled (to protect executives from personal liability), and settlements almost always involve no admittance of wrongdoing. For simplicity, in the rest of the paper we nonetheless use the term fraud, and do not append the adjective "alleged".

In total, the sample includes 212 frauds detected in the 1996-2004 period.[4] These frauds include all of the high profile frauds such as Enron, Worldcom, Adelphia and Healthsouth, as well as many others. The class action database provides start and end dates for the frauds.[5] The frauds in the sample

[4] We drop 4 frauds from the DMZ sample because they are not over the $750 million threshold at the beginning of the fraud.

[5] Because these dates can be, and often are, revised as suits progress, we use the most recent definition of the suit window from the legal filings. This definition of duration may be conservative in that the statute of limitations on class actions under Section 10(b) of the Exchange Act dictates that cases must be brought within one year after

have an average duration (from the class action suit period) of approximately 1 year and 7 months (590 days). To gauge the pervasiveness of fraud, we also have to identify the possible population of firms that could have produced frauds. The relevant population for our purposes is, like our fraud sample, the set of U.S. publicly-traded companies with $750 million in assets. In Compustat, 2,976 companies on average per year meet this criterion.

With this information on the start dates of frauds that are caught, the information on their duration, and on the underlying population of firms, we can calculate the percentage of firms that are engaging in frauds that are caught for every year in our sample period. Figure 1 illustrates the incidence of caught frauds. We plot the percentage of large U.S. publicly traded companies that start fraud in each year (the grey bars) and the percentage of firms engaging in fraud (the black bars). This evidence suggests a non-trivial level of fraud taking place, with an average of 1.3 percent of firms starting fraud each year and 3.3 percent of firms being engaged in fraud. We think it is particularly important to use an average of fraud over a time period that involves booms and busts as this period does, as there appear to be time patterns in fraud activity.[6] Note the significant time series variation in these numbers, with the incidence of firms starting fraud peaking in 2000 (2.4 percent of large corporations), and the fraction of firms engaging in fraud peaking in 2001 (5.9 percent of large corporations).

Rather than solely using this data, we make modifications to reflect a clear bias. Figure 2 introduces a correction for the fact that there are some additional frauds that will be caught after we ended our sample collection in 2004, that were taking place during our sample period. To extrapolate these missing frauds we use the distribution of fraud duration for those cases which begin prior to the year 2000 to forecast how many cases are yet to be caught for frauds starting through 2003. Using the duration distribution, we then roll the distribution forward to forecast how many additional cases that began after

---

discovery of the alleged violation, and no more than three years after the violation occurred. This limit was loosened in 2002 as Sarbanes-Oxley legislation changed this to 2 years after discovery, and no more than 5 years after the violation occurred.

[6] Galbraith () was one of the first to conjecture such a relationship. This idea has been explored more recently in theoretical models in Povel, Singh and Winton (2007) and an empirical application in Wang, Winton and Yu (2010).

1999 will yet be caught. This correction raises our estimate of the overall incidence of firms starting fraud to 1.4 percent per year and the overall fraction of firms engaging in fraud in any year based on the 96-2004 period to 4.0 percent of firms. Again, we focus on the average over a period with cycles and boom as the data show significant time series variation with a much higher incidence of frauds starting prior to the demise of Enron and Arthur Andersen in 2001 and the passage of SOX in 2002.

II. Methodology & Statistics: Arthur Andersen Natural Experiment

The figures provide an incomplete picture, as they ignore the fact that some frauds are never caught. Without exploring the likelihood of fraud taking place without being caught, we do not know if these observed estimates are the whole iceberg, or just the tip of the iceberg.

As mentioned in the introduction, our identification strategy for inferring unobserved fraud relies on a basic probability rule. What we observe is the joint event of firm engaging in fraud and being caught, **Pr(*engage, caught*)**. (We will use the convention of bolding the variables we observe.) Our actual variable of interest is the probability of a firm engaging in fraud, regardless of whether it is caught or not, Pr(*engage*). By the law of conditional probability, the unconditional probability of engaging in a fraud can be written as:

$$\Pr(engage) = \frac{\mathbf{Pr}(\boldsymbol{engage, caught})}{\Pr(caught \mid engage)} \tag{0}$$

Thus, if we knew the detection likelihood, that is the probability that a fraud is caught, given that it is ongoing, Pr(*caught*|*engage*), we could easily calculate the Pr(*engage*). In the circumstance where Pr(*caught*|*engage*) is equal to one, then the unobserved Pr(*engage*) would simply be equal to the observed **Pr(*engage, caught*).**

Our strategy is to identify and exploit circumstances in which the likelihood of being caught by a particular fraud detector increases to close to one. In fact we assume the Pr(*caught*|*engage*) is one, providing a conservative bias to our estimates. We compare the caught fraud rate in this ‘full detection’ circumstance to the caught fraud rate in the normal circumstance, and infer the normal detection rate.

II.1. Experiment design for identifying probability of engaging in financial fraud

Our experiment uses the sudden demise of the auditor Arthur Andersen (AA) as a situation in which the likelihood of being caught for financial fraud approaches one. In the fall of 2001, accusations began to emerge about Arthur Andersen as a result of Enron's collapse. In March of 2002, AA was indicted, and in June, AA was convicted. Over the period 2001-2002, all of Arthur Andersen's clients had to change their external auditor. Because new auditors do not want to face litigation risk or reputation risk for actions (or non-actions) taken by prior auditors, new auditors have a strong incentive to "clean house". Cleaning house implies that new auditors address any potentially misleading financial reporting, ranging from gross errors to overly aggressive financial reporting. An important advantage of this is that the line of causality is clear: the turnover leads to the fraud revelation, rather than the fraud leading to the turnover. A significant literature concerning litigation risk finds evidence that more conservative accounting reporting emerged in former AA clients (see Cahan and Zhang (2006), Krishnan (2007)).

Strictly speaking this experiment speaks to the ability to identify only a particular type of frauds, the largely financial frauds auditors are positioned to detect and where they have some incentives to uncover. Below we distinguish financial frauds detectable by an auditor from all frauds by using an asterisk as a superscript. Thus, the observation of a financial fraud being caught by an auditor in conditional probabilities is:

$$\mathbf{Pr(}\boldsymbol{engage^{*},\ caught^{*}_{Auditor}}\mathbf{)} = \Pr(caught^{*}_{Auditor} \mid engage^{*})\,\Pr(engage^{*}). \tag{1}$$

Our main identifying assumption is that post-AA, the probability of auditors detecting an ongoing financial misreporting fraud increases to one for post AA clients. To keep notation simple, we do not include time subscripts. But for all the equations, the AA marker means that the firm was an Arthur Andersen client coming into the demise of 2001-2002, and the detection is immediately thereafter, 2002-2004.

$$\textit{Assumption 1}: \qquad Pr\left(caught^{*}_{Auditor} \mid engage^{*}, AA\right) = \mathit{1}. \tag{2}$$

Assumption 1 is conservative. Auditors may not be privy to information as to the impropriety, thus limiting their capacity to find all fraud.

To lay out specifics of what we need to go from Assumption 1 to identification for financial frauds, we write a ratio equation of the conditional probability equation for AA firms and all firms:

$$\frac{\mathbf{Pr}\left(\boldsymbol{engage^*, caught^*_{Auditor}}\right)}{\mathbf{Pr}\left(\boldsymbol{engage^*, caught^*_{Auditor} \mid AA}\right)} = \frac{\Pr\left(caught^*_{Auditor} \mid engage^*\right)}{\Pr\left(caught^*_{Auditor} \mid engage^*, AA\right)} \; \frac{\Pr\left(engage^*\right)}{\Pr\left(engage^* \mid AA\right)}. \tag{3}$$

We need a couple of additional assumptions. Assumption 2 says that financial fraud was equally likely in AA and non-AA firms prior to 2001. In the next section, we present literature and evidence consistent with this assumption.

$$Assumption\ 2: \qquad \frac{\Pr(engage^*)}{\Pr(engage^* \mid AA)} = 1. \tag{4}$$

With Assumptions 1 and 2, we see that the ratio of observed frauds in all firms to observed frauds in AA firms provides an estimator for the probability of detection, given that a firm has engaged in financial misreporting fraud. This estimator is now based only on observable data:

$$\Pr^{BestEstimate}\left(caught^*_{Auditor} \mid engage^*\right) = \frac{\mathbf{Pr}\left(\boldsymbol{engage^*, caught^*_{Auditor}}\right)}{\mathbf{Pr}\left(\boldsymbol{engage^*, caught^*_{Auditor} \mid AA}\right)}. \tag{5}$$

We call this estimate our Best Estimate of the detection likelihood.

II.2 Were Arthur Andersen Clients More or Less Likely to Engage in Fraud?

Before turning to results, we evaluate Assumption (2), which asserted that the probability of engaging in financial fraud was the same for AA and other auditors. Other studies support this claim. In particular, Agrawal & Chada (2005) find in a matched sample that the existence of Arthur Andersen as the auditor does not associate with firms having more restatements. Likewise, controlling for client size, region, time and industry, Eisenberg & Macey (2004) find that Arthur Andersen clients do not perform any better or worse than other firms.

Because our sample of only large U.S. corporations is different from the aforementioned studies, we construct an additional test to ensure the evidence in prior work holds with our sample of larger firms. We build on the prior research that focuses on restatements as the measure of manipulation. Specifically we test whether pre-indictment (1998-2000) Arthur Andersen clients differ from non-AA clients using the earnings manipulation score “ProbM-Score” of Beneish (1999) and Beneish and Nichols (2007). The idea in these papers is that specific indices made from financial statements can be indicative of fraud taking place or conditions for fraud to take place. The components in the ProbM Score are: days sales in receivables, gross margin, asset quality index, sales growth index, depreciation index, SGA index, leverage, and the ratio of accruals to assets. Beneish motivates how each of these subindices captures an aspect of manipulation, and, thus, we refer the interested reader to Beneish (1999) for a description of each variable. To construct the ProbM Score, we download the appropriate financial statement variables from Compustat, construct all the components directly following the data definitions in Beneish (1999), and use Beneish’s estimated coefficients to construct the ProbM Score. Appendix II details the equation calculation.

We report the result of this analysis in Table 1. The table first provides univariate differences between AA clients and all non-AA clients that meet our size criteria or between AA clients and all non-AA clients that use one of the other Big Five audit firms, a perhaps more appropriate reference group. Panel A shows no significant differences between AA and non-AA clients across all sub-components of the prob-M score and for the prob-M score variable itself. Panel B introduces the possibility that AA and non-AA clients differ on other dimensions. Indeed, AA firms are (trivially) smaller, have more debt, higher sales to assets and higher profitability measured using EBITDA to sales ratio.

Thus, we implement multivariate tests to control for covariates. Panel C reports regressions of the prob-M score on an indicator variable indicating AA and a series of control variables. Across all specifications, being an AA client has no significant impact on the prob M score. In column 1, we use an OLS specification and include all firms. In column 2, we restrict our attention to the sample with a top-5 auditor. Columns 3 and 4 repeat the analysis using a median specification. We conclude that, as was found in prior studies, Arthur Andersen clients are not statistically or economically different from other auditor's clients prior to 2002.

III. Results: Arthur Andersen Natural Experiment

III.1. Best Estimates for Fraud

The main results from the natural experiment follow directly from the Best Estimate equations. The first input comes from looking at the large corporations who existed in 2002 and had an auditor identified in Compustat in either 2001 or 2002. We code the firm as an AA client if the auditor was AA in either 2002 or 2001. We put no restriction on survival post 2002. To capture the staggering of the change in auditors and to allow the new auditors time to process all of the new client accounts, we measure a firm as having fraud if the fraud started prior to and including 2002, and the fraud was revealed in 2002 or later.

We provide calculations in Table 2. Firms with Arthur Andersen as their auditor in 2001 or 2002 had a 1.379 percent chance of having an auditor reveal existing fraud in the 2002-2004 period. This is compared to all firms in this same time period, which only had a 0.378 percent chance of having an auditor reveal fraud. (Note that in this period, auditors reveal 11 percent of the fraud cases in DMZ.) Comparing these two numbers, we infer the detection likelihood for auditors of auditor-perceptible frauds in normal times (outside the extraordinary AA circumstance) to be just 0.275. This estimate suggests that during normal times there truly is an iceberg with nearly four times the size under the water as what is observed. This estimate is conservative to the extent that the probability of revelation conditional on having been an AA client is less than one. Research suggests this is likely. The incentive to critically review past audit decisions and 'clean house' is dulled by the fact that the same individuals continued to audit former AA firms as other auditing firms hired AA former auditors and they brought their clients with them (Blouin, Grein and Rountree (2005)).

To speak to all frauds, rather than just auditor-perceptible financial frauds, we need to make an additional assumption about the relationship between detection likelihood in cases of auditor-detectible financial fraud and in other types of fraud. We explore two possibilities, one in this section that is in our opinion the best estimate case, and a second in the next section that provides a very conservative lower bound.

First, we assume that detection likelihood is the same for all other frauds as for auditor-detectible financial frauds. This will be the case if (a) there is specialization in fraud detectors, so auditors don't catch all types of frauds (which is a natural assumption given the findings in DMZ), and (b) the iceberg the AA experiment revealed is of similar size for other frauds. This experiment provides no data to support or refute this assumption. We feel that it is likely a conservative assumption. We think detection likelihood if anything is higher for financial frauds as insiders need to produce financial statements, and likely misrepresentations can in many cases be identified by comparing this with past history and competitor information as is common in the accounting literature. It is more challenging for fraud detectors in cases of concealment or non-disclosure as 'they don't know what they don't know,' and such types of frauds are more likely to lie at the heart of frauds involving lying about the future.

Formally, we assume the following:

$$\textit{Assumption } 3(a)\text{:} \qquad \Pr(\textit{caught} \mid \textit{engage}) = \Pr(\textit{caught}^{*}_{\textit{Auditor}} \mid \textit{engage}^{*}). \tag{6}$$

Now we can use the conditional probability definition (0), with the baseline observed frauds that are engaged and caught. Specifically, the unconditional probability of a firm committing fraud is now:

$$\Pr^{\textit{BestEstimate}}(\textit{engage}) = \frac{\mathbf{Pr(\textit{engage}, \textit{caught})}}{\Pr^{\textit{BestEstimate}}(\textit{caught}^{*}_{\textit{Auditor}} \mid \textit{engage}^{*})}. \tag{7}$$

We present results both when we exclude from the numerator the post AA period for AA firms, where we find a probability of engaging and being detected of xx%, and when we don't which as noted in section I is 4.0%. With the detection rate of .275, we arrive at an estimate of the probability of engaging in fraud of 14.5% (=4.0%/.275). This is our best estimate based on the AA experiment.

III.2. Lower Bound Estimates for Fraud

To produce this best estimate we needed to make a number of strong assumptions. We now provide an alternative set of assumptions that are designed to be conservative to provide a lower bound estimate. Specifically, one could look at the forced auditor switch for AA clients as a shock that transformed the incentives not only of auditors, but made all fraud detectors more diligent with respect to these firms. For this to be a valid assumption, the act of changing auditors must clean house of all frauds, financial and other.

Formally, we consider the following where we no longer use an asterisk as we are no longer assuming just auditor detectible frauds are caught:

$$\textit{Assumption 3(b)}: \qquad Pr\left(caught \mid engage, AA\right) = 1. \tag{8}$$

With this assumption we can provide two lower bound estimates as we also show in Table 2. First, we can very easily go to a lower bound estimate for the post AA period for former AA clients. By the basic probability rule given in equation (0) and assumption 3(b), the underlying probability of engaging in fraud for these firms in this time period is exactly the same as the observed probability of engaging and getting caught. In this period for these firms we therefore produce an estimate of engaging in fraud of 6.21% [pr (engage, AA)= **pr (engage, caught, AA)=** 6.21%].

Alternatively, we can follow the same process as in III.1, by comparing the difference between fraud detection for these treated firms and untreated firms to produce a measure of detection likelihood (equivalent to equation 5), and then applying that detection likelihood to the whole sample (equivalent to equation 7). The detection likelihood under these assumption is .715 (4.44%/6.21% = .715). This implies the iceberg under the water is just 40% as big as the portion above the water. Not surprisingly, this detection likelihood is decidedly higher as it is based on much more conservative assumptions. First, it is unlikely that the incentives of the forced AA change lead to the detection of all frauds. Second, while auditors incentives are heightened to detect fraud following the forced turnover producing a difference between AA firms and non AA firms, there is no compelling reason why there would be a similar difference in the detection incentives for other fraud detectors such as analysts, media, employees. If these other detectors incentives rose also for non AA firms, this would bias upwards the estimated detection likelihood. With these flaws in mind, we turn to equation (0) again, and now use this detection likelihood producing a lower bound estimate of fraud likelihood of 5.6% [4.0% /.715]

IV. Validation Methods to Identify Frauds that are not Caught

The Arthur Andersen experiment provides a natural experiment setting to identify hidden auditor-detectible frauds because it provides a clear circumstance in which the likelihood of being caught increases to close to one. This strength in identification comes at a cost however. It is reasonable to wonder if these results depend on one event (albeit one that affected a significant percentage of traded corporations), and to have concerns about the strong assumptions needed to go from cleanly identified detection likelihoods for auditor-detectable frauds to estimates of overall frauds. To explore the validity of our findings based on the demise of AA, this section introduces two additional approaches to estimate fraud and compare the results with those from the Arthur Andersen experiment.

IV.1. Validation by Incentives and Opportunities Estimation

Auditors are important fraud detectors, but only account for eleven percent of detections in DMZ. Instead a village of detectors contributes to fraud revelation, led by analysts, employees, media, short sellers, and non-financial regulators. For each of these actors, one can think of circumstances when the actor can be more effective (e.g. as a result of the information leading to uncovering fraud being more accessible or the incentive to reveal it is higher). For example, if media or analyst coverage of a firm is high, fraud detection may be more likely as journalists and analysts are more likely to understand the firm dynamics and patterns in other firms. Employees may reveal fraud more if a monetary payoff results or, at a minimum, their company is not likely to punish them for whistleblowing.

The concept behind our strategy in this section is the same as in the Arthur Andersen experiment. We look for a situation to assert that the probability of detection of an ongoing fraud approaches one. We do not have such a natural experiment, so we try to make one synthetically by seeing how the probability of detection increases when firms potentially committing fraud score high on all variables capturing the heightened incentive or opportunity for detection. It is slightly more complicated (because of endogeneities), but the intuition is just estimating a probit as a function of a set of indicators for high incentive and opportunities variables. We then calculate an inference as to what a true natural experiment would reveal by adding up the estimate impact on the probability of detection by setting all of those indicators to unity.

We begin by explaining the nature of and data for the heightened incentives and opportunities variables to convey fluidly the methodology.

a. Analyst detection heightened: *Analyst following*. We assume that the likelihood that analysts will capture fraud increases with the extent of analyst following of a firm. We use the number of analysts issuing forecasts, with data taken from I/B/E/S and setting the number of analysts to zero if the data are missing.
b. Media detection heightened: *Media coverage.* We assume that the likelihood that the media will uncover fraud increases in the extent of media coverage of the firm. We manually create a media coverage variable. For each firm in Compustat whose 1995 assets is greater $750 million, we search the Wall Street Journal print edition (via Factiva) and record the number of media hits for the year.[7]
c. Short seller detection heightened: *Institutional ownership*. We assume that short selling incentives to capture fraud are higher in firms with lower expected costs in assembling short positions. We follow the literature and use institutional ownership as a proxy for the shortability of firm stock. We collect data on institutional ownership from Compact D.
d. Regulator detection heightened: *Regulated.* We assume that regulatory scrutiny of firms that results in fraud detection is more likely in the non-financial firms in our sample when the industry the firm is in is regulated. The regulated industries in our sample include: electricity, gas supply, telecommunications, water, and healthcare.
e. Employee detection heightened: *Fortune Best 100.* To capture situations in which employee incentives and opportunities for fraud revelation are high, we introduce two measures. The first one is whether the company was a *Fortune Best 100* firm. Our assumption is that the Fortune Best 100 firms are environments that would less likely penalize whistleblowing.
f. Employee detection heightened: *Qui Tam.* Second, we use the statute regarding *qui* tam lawsuits to consider employees' monetary incentives for whistleblowing. Q*ui tam* lawsuits allow whistleblowing employees to receive payment for bringing forward information about frauds, (so long as part of the fraud is committed against the government and the government recovers money in damages). To identify which industries generally qualify, we searched the data on qui tam lawsuits available from the Department of Justice Civil Division. The overwhelming majority of qui tam suits and settlements

[7] We eliminate lists which are automatically generated (e.g., lists of large stock movers for a day), and we manually check each firm whose company name contains common language words (e.g., Apple). The range of media coverage is from zero (36% of the sample) to 237. The top three media hits in 1995 are Microsoft (237 hits), IBM (235 hits) and AT&T (228 hits).

occur in the healthcare and defense contractor industries. Thus, our *Qui Tam* variable codes whether firms are in these industries.

g. Management insiders detection heightened: *Post-SOX*. Finally, as a measure of management insiders' incentives to reveal fraud, we use a dummy to identify if the infraction took place pre-SOX or post-SOX. The assumption is that post SOX there was additional oversight of management insiders by the board resulting in added scrutiny by insiders.

The empirical methodology follows directly from Wang (2010). We denote the potential for a firm *i*'s fraud to be detected in year *t* as D_{it}, where the firm is caught if D_{it} is positive:

$$\begin{aligned} D_{it} &= X_{it}^{D}\Gamma_{D} + \nu_{it} \\ caught_{it} &= 1 \; if \, D_{it} > 0. \end{aligned} \tag{9}$$

D_{it} is a function of observables X_{it}^{D}. Our strategy is to include a set of indicators $\left[I_{it}^{HiIncentOpp} \right]$ in X_{it}^{D} that equal one when the circumstance leading to detection by each of fraud detectors is high (splits on the above lists of variables). Also in X_{it}^{D} are general characteristic $\left[\mathrm{x}_{it}^{D} \right]$ that would lead to detection in all firms (e.g., size and stock performance). We denote these two groups: $X_{it}^{D}{}' = \left[I_{it}^{HiIncentOpp} \quad \mathrm{x}_{it}^{D} \right]$.

Of course, factors that lead to detection also may affect (usually deter) starting a fraud. Wang (2010) addresses this challenge by applying the Poirier (1980) bivariate probit model for estimating dichotomous outcomes in partial observability settings to the case of corporate fraud. We follow accordingly, defining E_{it} to be the incentive for firm *i* to engage in fraud at time *t*. Fraud is committed if E_{it} is positive:

$$\begin{aligned} E_{it} &= X_{it}^{E}\Gamma_{E} + \mu_{it} \\ engage_{it} &= 1 \; if \, E_{it} > 0. \end{aligned} \tag{10}$$

E_{it} is a function observables X_{it}^{E}, which also includes the high incentives and opportunities indicators: $X_{it}^{E}{}' = \left[I_{it}^{HiIncentOpp} \quad \mathrm{x}_{it}^{E} \right]$.

Identification in Poirier's model comes from two pieces. First, Poirier assumes that $\left(\mu_{it}, \nu_{it} \right)$ are distributed bivariate standard normal. Second, identification depends on our ability to come up with variables which affect either firms' incentives to engage in fraud or detectors' ability to uncover fraud, but not both. The exclusion restrictions, again fortunately for us, are taken up in rigor in Wang (2010). Under the assumption that some of the X_{it}^{E} and some of the X_{it}^{D} are excluded from each other's set, the parameters in Γ_E and Γ_D can be identified using a bivariate probit model:

$$Pr\left(engage_{it}, caught_{it} \,\middle|\, X_{it}^{E}, X_{it}^{D}\right) = \Phi\left(X_{it}^{E}\Gamma_E,\ X_{it}^{D}\Gamma_D\right), \tag{11}$$

where $\Phi(\cdot,\cdot)$ denotes joint cumulative standard normal distribution over the two arguments.

Following Wang (2010), Table 3 lists the factors that we hypothesize should influence both the incentive to engage in fraud and the detectability of fraud. Panel A lists factors we hypothesize affect both the likelihood of engagement and detection, Panel B lists those variables we hypothesize affect primarily the likelihood of engaging in fraud, Panel C lists variables that primarily affect the likelihood of detection.

We start by assuming that well known firm fundamentals of company size, stock return and R&D influence both the likelihood of engaging, and of detection. For example, a disappointing stock return likely encourages engaging in fraud and likely leads to more detection behavior. We follow Wang (2006) and include R&D as a measure of the opacity of firm fundamentals, under the assumption that it is easier to engage in fraud if financials are more opaque and it is harder to detect under the same circumstances.

We also include variables that capture the likelihood of external monitoring, and because they are easily observed also create incentives to reduce engaging in fraud. DMZ identify a number of important monitors in practice including: financial analysts, the media, short sellers, and industry regulators. Yu (2008) finds that firms with more analyst coverage engage in less earnings management. Dyck, Volchkova and Zingales (2008) document the impact of media coverage on uncovering governance abuses. DMZ argue that industry regulators uncover significant information about firms in performing their main tasks of protecting the rate setting process. Higher levels of monitoring are likely to increase detection, but ex ante knowledge of these monitors may deter the likelihood of engaging in frauds. All of these measures appear as our heighten incentives and opportunities situation, which we describe below.

Finally, in both the engagement and detection equations we also include variables to capture internal monitoring. Employees have access to information about fraudulent behavior. Also, the Board may play a role in deterring and uncovering fraud, to the extent they are active in monitoring. As internal monitoring variables, we include employee incentives as gauged by the ability of employees to file for whistleblowing payments and board power. We also include a dummy for the time period being post-Sarbanes Oxley, to capture the pressure on management to stay clean after the increased personal liability ordained by SOX.

Panel B lists variables that primarily influence fraud by influencing the likelihood of engaging in frauds. We focus here on the monetary incentives for top management, in particular the role of options and more generally compensation based on future rather than current returns. We construct a measure of in-the-money options. Firms may commit less fraud if the managers have more exercisable options exposed to the consequences of fraud. Conversely, managers may commit more fraud if the managers are trying to sustain stock price to cash out of their options over time, or if they want to take a risk with a large non-linear payoff. We also construct a measure of the percentage of the Managers Compensation Paid in Future Returns, defined as the percentage of restricted stock grants divided by total compensation. As is the prior variable, managers may want to take on more risk with nonlinear payoffs of compensation being performance based.

Finally, Panel C lists variables that enter only into the ex post detectability of the fraud already being engaged. The helpful insight here is that fraud has a time dimension, and unanticipated information and events may happen after fraud has already started that change the setting for detection. We focus first on abnormal performance measured by abnormal stock returns using a CAPM model for returns. Detectors may pay more attention to a firm if the firm is performing abnormally relative to the industry and market. We think detectors will be motivated not only by the level of returns but also by their volatility and include measures of own firm volatility as captured in the standard deviation of returns over the last 5 years, and market volatility as captures in the standard deviation of the S&P 500.

Table 4 illustrates the differences in the pervasiveness of fraud across these detector incentives, restricting ourselves to those companies where fraud has been detected. The univariate analyses in Panel A suggest that almost all of these circumstances matter, with significantly higher levels of fraud where there is high analyst coverage, high media coverage, high shortability, a *Fortune* best 100 firm and where qui tam suits are possible, with regulated being the only variable not producing a significant result. Panel B reinforces the importance of the setting to fraud detection, in this case being more demanding of the data in seeing if within fraud sample variation in the settings influences the extent of detection by that particular fraud detector.

IV.1.2. Incentives & Opportunities Estimate: Results

Table 5 presents estimates for three equations. The first two equations are independent probit models of the probability of starting fraud and the probability of fraud being detected. The key equation is model (3), the partial observability bivariate probit model of Poirier. In the Table we bold those variables that are in both the engage and detection equations (listed in Panel A) and leave in italics those variables either in the start equation and not the detect equation (top half of the panel), or in the detect and not the start equation (at the bottom of the panel).

The coefficients in the regression from Panel A that we hypothesize affect both engagement and detection, for the most part, align with intuition often having opposite signs in the engagement equation from the detection equation. For example, the high analyst dummy variable is negative in the start equation (insignificant) and positive and significant in the detect equation as expected, consistent with the hypothesis that in recognition of the high number of analysts insiders are less likely to start a fraud and if they do they are more likely to get caught. Similarly, the high shortability indicator that we use to capture the presence of institutional investors has similar opposite effects, deterring engaging in fraud, and encouraging detection. Or consider the Post-Sox dummy that is negative in the start equation and positive in the detection equation.

The regression evidence also supports our hypotheses on variables that affect engagement only or detection only. We find a strong and significant impact of pay based on future returns on the probability of starting fraud. We find also that abnormally poor returns and abnormally high volatility prompt detection.

The true value to this estimation is that it allows us to estimate the unconditional probability of a fraud starting in our sample, and provides an indication of what that unconditional probability would be if we could move to a world where all fraud detectors had maximial incentives. We report these results in the very bottom of the panel. First, with this methodology we can arrive at an estimate of the undonditional probability of starting a fraud. This is 5.3%, and with the average duration of fraud in our sample of 1 year and 8 months this implies the unconditional probability of starting a fraud is 8.85%. This alone doubles our simple estimate reported in figure 2.

We can go further though and explore the same type of thought experiments as the AA experiment by seeing how fraud engagement changes when detector incentives changes. We find three times the likelihood of starting a fraud pre Sox than post Sox, we find 2 and a half times the likelihood of starting a fraud in firms with low analysts compared to high analysts, we find six times the likelihood of starting a fraud outside of healthcare, and we find close to double the likelihood of starting a fraud in firms with low institutional than high institutional ownership. This splits are interesting, but potentially lead to double counting. So at the very bottom of the table we see what would happen if we were in a setting where engaging was most likely, simultaneously considering a pre Sox setting with firms with low institutional ownership, low analyst coverage and not in healthcare. This produces an estimate of the likelihood of engaging in fraud of 15.8%. This is very close to our estimate from the AA experiment.

IV.2. Validation by Survey

A potential concern with the first validation method is that it relies on the same sample of fraud detections as our main Arthur Andersen natural experiment data. Thus, as a second validation, we conducted a survey with University of Chicago MBAs to assess the frequency of illegal behavior in corporate America. In particular, all first year campus Chicago MBAs are required to attend a program called LEAD, which tries to develop soft skills. In the academic year 2004-2005 LEAD program, we inserted an anonymous survey on illegal and unethical behavior students encountered in their previous jobs. The question asked: "*In your job you are asked to do something that is illegal. Example: Your boss asks you to lie in reporting sales*." We then asked them to provide a short description of the illegal act they were asked to do. We also asked in what industry they were working in and what function they were performing at the time.

This method has its own pluses and minuses. On the plus side, this method is the least likely to be affected by the uncaught fraud selection bias. Given that the students have left their previous employers and operate in an academic environment under guarantee of anonymity, it is unlikely that they will omit reporting any fraud they encountered. On the negative side, we might omit major frauds that are concentrated in the headquarters. Given the low level position most MBAs covered before they joined the program, they are unlikely to be privy of major fraud consummated in the corporate headquarters. In addition, it is likely that these frauds are not material, although if they are similarly widespread across the organizations in which they were employed they might end up being so for the organization.

With these caveats in mind, Table 6 Panel A reports the percentage of MBAs who responded they faced a legal dilemma. On average 14.8 percent of the students were asked to do something illegal in their previous employment, almost identical to our main Arthur Andersen estimate. The actions they were requested to perform vary from falsifying sales numbers to reclassifying a job as redundant to get rid of an employee with very high health-related expenses. In all the cases, however, they appear as truly illegal activities, hence there is no sign of misclassification there.

Surprisingly, the incidence of illegal activities does not seem to differ across industries. The only exception is consumer goods, where the incidence is only 7 percent, less than half the sample average. One possible explanation is that manufacturers of consumer products are more sensitive to their public image, because this has a larger impact on sales. This conjecture is supported by the fact that also the incidence of unethical requests is lower than average (27 percent vs. 37 percent) in the consumer industry. Contrary to expectations, the financial service industry does not experience a higher incidence of illegal activity. The same pattern is present if we divide the incidence by function performed by the student in his/her previous employment. Contrary to expectations, investment bankers are not more likely to be asked to undertake something illegal nor are accountants. Illegal activity is very homogenously diffused across the board.

Aside from intrinsic interest in these data, in summary, the survey result that 14.8 percent of MBAs experienced fraud in action suggests that our main estimate that 14.5 percent of firms are engaging in fraud at any point in time is quite valid.[8]

IV.3 Related Literature

There is a voluminous literature on corporate governance and fraud.[9] As Karpoff, Lee and Martin (2008) argue, there are almost no papers that produce estimates of the probability of detection, given that firms are engaging in fraud. Accordingly, very little of this literature speaks to the overall incidence of fraud including detected and undetected frauds.

Two notable exceptions are Wang (2011) and Wang, Winton and Yu (2010). Neither paper reports as a significant result the predicted probability of engaging in fraud, but this can be inferred from results reported in Wang, Winton and Yu (2010). This paper examines frauds by firms that go through IPOs. They start with 3297 IPOs from 1995-2005, use data on detected frauds (they define frauds as firms that had an AAER and/or where there was a securities class action that was not dismissed, exceeded the $2mn threshold and related to financial reporting), and then generate predicted probabilities of engaging in fraud by using a bivariate probit model. Their predicted probabilities are in line with our estimates, ranging from 10-15%.[10]

[8] Entering MBAs typically work a couple of years before returning to school. We ignore that we are cumulating these entering students' experiences over a few work years rather than speaking to in any given year, as the rest of the paper does. Since we doubt that new college graduates are given any responsibility for a year or two and since this test already biases against finding fraud since lower-tier employees are less likely to be knowing participants, we feel the accumulation of experiences in the previous job is not an issue of any magnitude.

[9] Much of the fraud related literature has been published in accounting journals, with important papers cited in KKLM .

[10] 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 literature on options backdating also provides another estimate of engaging in fraud. This setting has a number of strengths. In a sense this is an experiment in which detection rates go to 1 as researchers *ex post* look at the data to identify firms whose actions are most consistent with backdating. It is complementary to our tests, as our sample includes no back dating cases as they developed after we completed our data collection. It also has some limitations, as researchers' identification of likely backdaters does not mean that these situations would satisfy our definition of fraud, as they may not be material.

Bebchuk, Grinstein and Peyer (2010) look in depth at the options backdating scandal first brought to attention by Lie (2005). They attempt to uncover the percentage of publicly-traded firms from 1996-2005 in which CEOs or directors were 'lucky' directors in that they received option grants on the lowest price day of the month, filtering out those that could have taken place simply due to luck. These lucky option grants increased the value of that grant by 20% and CEO pay in that year by 10%. By their estimate 12.4% of firms have such lucky CEOs and 7% of firms have lucky directors, with the percentage of "lucky grants" of 14.5% prior to SOX and 8.4% after. Note that this reveals a big iceberg, as prio to the research of Lie (2005) none of this was revealed. Note further that the 12.4% is very close to our estimate from the AA experiment.

V. How Expensive is Corporate Fraud?

In section II we show that 4.0 percent of large publicly traded firms are eventually revealed to be engaged in fraud. The AA experiment provides a best estimate of the detection likelihood of 27.5%, leading to an estimate that in 14.5 percent of firms, or one in seven, insiders are engaging in fraud. Two validation experiments and related literature provide similar estimates. Is this level of fraud detection and pervasiveness a point of concern? To address that question, we need to go further and provide an assessment of the economic costs associated with frauds.

V.1 Prior Estimates of the Cost of Fraud

Prior research provides various measures of the cost of financial fraud. One method is to take an event study approach and carefully measure the decline in equity (and sometimes debt) capitalization at the moment of fraud revelation (e.g. Feroz, Park and Pastena (1991), Palmrose, Richardson and Scholz (2004), Grande and Lewis (2009)). This is a good measure under two conditions. First, it must be the first indication to the market of the fraud. If there was prior partial leakage, then the event alone would be an underestimate of the actual costs as Gande and Lewis (2009) show. Second, and perhaps more challenging, the measured decline must be solely attributable to the fraud, rather than to the revelation of bad information about fundamentals which the fraud tried to cover up.

A paper that addresses this second issue is Karpoff, Lee and Martin (2008). They use a sample of firms from 1978 - 2002 subject to SEC enforcement actions and put a price tag on fraud. To assess the extraordinary loss they first collect the abnormal returns from a trigger event that brings the fraud to light and add the abnormal returns from *subsequent* public disclosure of enforcement events. To capture the value loss arising from deterioration in fundamentals (they call this the readjustment effect) they attach a market value to the book value of assets written off in subsequent financial restatements. The outcome of this analysis is that they estimate that the mean (median) fraud losses not attributable to the readjustment are 29% (28%) of equity value. Assuming a 25% D/V ratio that is the median in our fraud sample, and that the value of debt is unaffected by the fraud, their study suggests fraud is associated with destruction of 22% of the firms' enterprise value.

V.2. An Alternative Method for Calculating the Cost of Detected Fraud

We take a conceptually similar approach with two modifications. Reflecting Gande and Lewis's (2009) finding of a partial anticipation effect, we allow for the possibility that some information about the fraud leaks even *before* the trigger event that defines the end of the class period for securities class purposes and disregard market responses after the trigger event. We use an industry-adjusted multiples approach to capture an estimate of the value loss to a deterioration in fundamentals, motivated in part by the fact that our cases include not only cases of financial fraud (where restatements are available) but also non-financial frauds (where there are no restatements we can appeal to).

We start by expressing the value of the fraud firm in the pre-fraud period using a performance multiple (e.g. of EV/EBITDA). We then produce an estimate of expected multiple expansion by a typical firm in the industry over the fraud period. We then apply this industry-adjusted multiple to the performance data for the firm after the fraud has been revealed and reflected in the firm financials producing an estimate of 'non fraud implied enterprise value.' The difference between this estimate and the actual enterprise value at this point in time provides an estimate of the value loss for shareholders from fraud. Appendix IV provides more detail of the calculations involved.

To check the validity of the process consider a fraud firm manipulating the financials and reporting inflated earnings numbers for a while that return to normal earnings after the fraud is revealed. For simplicity, assume the firm is the typical industry firm as measured using multiples before the fraud and that there is no multiple change over the fraud period. In this case, the predicted hypothetical multiple is the same as the starting multiple which equals the industry multiple. Note that the time-limited manipulation has no effect on the implied value, but it very well could on the observed multiple that is likely to be lower. This will be the case if investors, consumers, suppliers and others change the terms under which they interact with the firm as a result of the fraud. Karpoff and Lott (1993) label this the lost reputation effect, and Karpoff, Lee and Martin (2008) see this as the biggest source of costs with corporate fraud. Note also that this approach accounts for industry trends that have nothing to do with the fraud. If, for example, during the fraud period industry fundamentals goes down by 10%, and the fraud firm's fundamentals go down by the same percentage, we would not want to attribute this decline as a cost to the fraud, and this process ensures that we do not.

V.3. From Costs of Detected Fraud to Social Costs of Fraud

To go from this number to an estimate of social costs of fraud in firms with detected fraud and undetected fraud a few more refinements are required. First, one reason the enterprise value may be lower after the fraud is that there are fines and other penalties and the stock price could capture these expected fees that the firm will pay. These are not social costs, as someone else receives them (e.g. the government, plaintiff law firms). They should be excluded in considering social costs for fraud firms, and certainly if we consider social costs for non-detected fraud firms where we do not expect these costs to ever be paid. KLM find that the mean punishment fines and settlements, both official and private, are 3.7 percent of equity value. [11]

Second, we seek to apply our findings of the costs of fraud from firms with detected fraud to the iceberg firms with undetected frauds by making some additional assumptions. There are good reasons to believe that there will be reputational costs in firms with undetected frauds as well. If a firm for example commits a fraud by creating a false perception that developments are going well when they are going poorly this will eventually be learned by employees, buyers and suppliers to the firm reducing their reputation even if there is no official fraud. Gande and Lewis's (2009) results that show partial anticipation of the fraud also shows that the market reflects such information even in the absence of an official fraud statement.

V.4. Disclaimers on our Method for Calculating the Cost of Fraud

Before presenting results, we also want to note that our approach also has limitations. One limitation that may make our estimate overly large is that it does not correct fully for information about fundamentals that might be confounded with revelation about the fraud. As noted above, we attempt to correct for this using changes in industry over time, but there could very well be extra fundamental information about the firm that does not relate to the industry (e.g. in a pharmaceutical firm information about a specific clinical trial) and does not relate to the fraud

V.5. Results: The Cost of Fraud Table 7 presents the results from this analysis. Panel A provides summary statistics for our sample, showing data pre fraud and post fraud. Means always exceed medians showing positive skewness in size, and post fraud measures show the decline in enterprise value, while there is growth in assets and fixed assets over the fraud period. Panel B shows more clearly the value destruction over the fraud period, reporting how measured by multiples the median fraud firm saw a significant deterioration from pre to post fraud, while the median industry firm was relatively unaffected

[11] We recognize that using this estimate is conservative. Insurers, and/or other firm stakeholders such as accounting firms or even directors pay part of these fines, reducing the cost borne by firm shareholders.

over the period. The final column in panel B provides the firm counterfactual using our methodology, that is how the median fraud firm would have done over the fraud period if it had simply followed industry changes in multiples. For example the median fraud firm had an EBITDA multiple of 12.31, and by our measure would have had a multiple at the end of the fraud period of 13.25. We find instead that the post fraud multiple is 9.49.

We think the EV/EBITDA multiple is most easily comparable across firms, so we focus on these results, providing results for sales multiples, and asset multiples for robustness. Unfortunately, we do lose some data points with EBITDA multiples arising from losing firms through bankruptcy and negative earnings.

We present our key findings in panel C where we express the fraud costs as a percentage of enterprise value, with the estimates based off firms with median characteristics. We find that fraud destroys 23.2% of enterprise value using the EBITDA multiples approach and assuming all firms pay punitive costs. If we exclude punitive costs, we find that fraud accounts for 20.4% of enterprise value. Using other multiples suggests that this might be a conservative estimate, as other multiples vary from 34.3% (assets multiple) to 40.4% (sales multiple).

Finally, we can estimate the overall costs associated with fraud for large firms by applying this estimate to the population of publicly-traded firms with more than $750 million in assets. We consider a couple of scenarios providing a range of estimates. With 14.5% of firms estimated to engage in fraud, and those frauds in turn costing 20.4% of enterprise value, this suggests 2.96% of enterprise value is lost to fraud (i.e. 0.145*20.4=2.96).

VI. Conclusion

In this paper we set out to answer the question of the pervasiveness of corporate fraud in the United States and to assess its costs. Using a dataset of corporate frauds in large corporations that impact shareholder value and are caught from DMZ, we take the next step to estimate the iceberg of undetected frauds to infer the unconditional probability that a fraud is committed whether or not it is subsequently caught. Our identification comes from observing situations in which the incentives for fraud detection are high. In particular, capitalizing on the natural experiment provided by the demise of Arthur Andersen, we estimate that approximately 14.5 percent of firms are engaging in fraud, based on the increased probability of a fraud being revealed following the forced turnover of external auditors (with a lower bound estimate of 5.6%). We check for validity of our results using an incentives and opportunities estimation framework and by surveying incoming MBAs, providing broadly similar results.

Having established the incidence of fraud, we then explore the social cost of fraud capturing costs borne by investors over and above the losses they incurred from the deterioration in firm fundamentals that often is the spur for the fraud in the first place. We introduce a new methodology that produces an estimate that the median cost of fraud in our sample is 20.4% of the pre fraud enterprise value. We also use pre-existing estimates from other researchers that arrive at broadly similar results. Finally, we put these two findings together to come up with an estimate of the cost of fraud, which we find to be 3% of enterprise value.

We think our findings have relevance both for investors in firms and for policy. The AA experiment suggests that the detection rate in normal times is just 27.5% of what it is during extraordinary times, suggesting significant scope to increase the engagement of fraud detectors. Consistent with a gap between what could be and what is, we also find significant differences across other fraud detectors when we compare situations with high and low incentives for fraud detection. This evidence establishes that existing fraud detectors can be more active. The social cost calculations we arrive at, along with that of prior researchers, also establishes that these frauds have substantial costs over and above hiding weaknesses in firm fundamentals.

What the evidence does not speak to directly is whether it is cost effective for investors and policy makers to take steps to increase on a permanent basis the detection activity. It is surely true that there are costs with heightened detection that we do not measure as well as benefits. It is also undoubtedly true that policy interventions, to the extent they subject all firms to the same treatment, create additional costs.

Appendix I: Data Appendix

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

First, they restrict attention to alleged frauds that ended 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.[12] 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.[13] Fifth, they exclude those security frauds that Stanford classifies as non-standard, including mutual funds, analyst, and IPO allocation frauds.[14] 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.[15]

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

The probability of manipulation, ProbM Score, of Beneish (1999) and Beneish and Nichols (2007) 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

[12] 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.

[13] Grundfest (1995), Choi (2004) and Choi, Nelson, and Pritchard (2005) 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.

[14] 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.

[15] 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.

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

For a complete description of the motivation for each item as an indicator of potential for manipulation and for the compustat codes leading to the calculation of the indices, please see the papers referenced above. We followed their compustat definitions exactly to construct the ProbM Score yearly for the large corporations in our sample.

According to Beneish (1999), a score greater than -2.22 indicates a strong likelihood of a firm being a manipulator

Appendix III – Process for Calculating Implied Value Loss not Attributable to Changes in Fundamentals

.

Specifically, we follow Berger and Ofek's (1995) multiples approach with modification to exploit firm-specific information. Assume that a fraud begins right after time *s* and ends before time *t*. The pre-fraud enterprise multiple, specific to firm *i,* which resides in industry *j,* is:

$$m_{ijs} = \frac{Long\ Term\ Debt_{is} + \ Market\ Equity_{is}}{Y_{is}}, \tag{11}$$

where we consider several valuation bases, $Y \in \{EBITDA, revenue,\ fixed\ assets\}$. Likewise, we define a pre-fraud industry multiple, M_{js}, as the revenue-weighted average multiple for SIC 3-digit industries, indexed by *j*. We exclude the fraud firm in this calculation. We do the same procedure at time *t*, the year ending *after* the fraud revelation date to get M_{jt}. We use the change in the industry multiple as the benchmark for how the firm's multiple would have evolved over the time period if it was just impacted by factors affecting the industry; i.e.:

$$\hat{m}_{ijt} = m_{ijs} \frac{M_{jt}}{M_{js}}. \tag{12}$$

The idea is to compare the fraud firm's value of debt and equity at time *t* with the debt and equity which would be projected by the firm's pre-fraud multiple adjusted to a growth or decline rate in its industry benchmark multiples. The estimated "but-for" or counterfactual valuation is thus the EBITDA, sales, or fixed assets implied enterprise value at time *t*, calculated as:

$$\begin{gathered} Counterfactual\ Enterprise\ Value = \hat{m}_{ijt} Y_{it}, \\ for\ Y_{it} \in \{revenue_{it},\ fixed\ assets_{it}, EBITDA_{it}\}. \end{gathered} \tag{13}$$

The next step is to compare the counterfactual with the actual enterprise value post fraud revelation to produce a dollar loss per firm arising from the fraud. To ensure comparability across firms we also express this dollar loss relative to the pre-fraud enterprise value to define the fraud loss as a percentage of enterprise value.

$$Cost\ Caught\ Fraud_{it} = Counterfacutal\ Enterprise\ Value_{it} - \left(Long\ Term\ Debt_{\text{it}} + \ Equity_{\text{it}}\right).$$

(13)

References (Incomplete)

Agrawal and Chada, “Corporate Governance and Accounting Scandals.” Journal of Law and Economics, 2005

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, “An analysis of forced auditor rotation: the case of former Arthur Andersen clients.” Accounting Review, 2005

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

Cahan and Zhang, “ After Enron: Auditor Conservatism and Ex-Andersen Clients, The Accounting Review, Vol. 81, No. 1 (Jan., 2006), pp. 49-82

Choi, Stephen J. 2004, "Do the Merits Matter Less after the Private Securities Litigation Reform Act?" Working Paper.

Choi, Stephen J. 2005, "Behavioral Economics and the SEC.” Working Paper.

Choi, Stephen J., Karen K. Nelson and A.C. Pritchard. 2005. “The Screening Effect of the Securities Litigation Reform Act.” Working Paper.

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.

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.

Efendi, Jap, Anup Srivastava, and Edward Swanson, forthcoming, “Why Do Corporate Managers Misstate Financial Statements? The Role of in-the-money Options and Other Incentives,” *Journal of Financial Economics*.

Dyck, Alexander, Adair Morse and Luigi Zingales. 2010. ‘Who Blows the Whistle on Corporate Fraud?” *Journal of Finance*.

Galbraith, J.K, *The Great Crash 1929*, 1961, Pelican.

Gande, Amar and Craig Lewis, 2009,“Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers,” *Journal of Financial and Quantitative Analysis*, 44 (4), 823-850.

General Accounting Office, 2002, “Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses, and Remaining Challenges,” 03-018

Griffin, Paula, Joseph Grundfest and Micael Perino, “Stock Price Response to News of Securities Fraud Litigation: Market Efficiency and the Slow Diffusion of Costly Information,” *Stanford Law and Economics Olin Working Pape*r No. 208.

Grundfest, Joseph A. 1995. “Why Disimply?” *Harvard Law Review*, 108, 740-741.

Hennes, K, Leone, A., B. Miller, 2008, “The Importance of Distinguishing Errors from Irregularities in restatement Reserarch: The Case of Restateements and CEO/CFO Turonover,” The Accounting Review, 83(6): 1487-1519.

Johnson, Marilyn F., Ron Kasznik, and Karen K. Nelson. 2000. “Shareholder Wealth Effects of the Private Securities Litigation Reform Act of 1995.” *Review of Accounting Studies,* 5(3) 217-233.

Johnson, Marilyn F., Karen K. Nelson and A.C. Pritchard. 2003. “Do the Merits Matter More? Class Actions under the Private Securities Litigation Reform Act.” Working Paper.

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

Karpoff, Jonathan, Allison Koester, D. Scott Lee and Gerald Martin, 2012 “ An Analysis of Dababase Challenges in Financial Misconduct Research,”

Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin, 2007. “The Legal Penalties for Financial Misrepresentation.” Working Paper.

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 M., and John R. Lott, Jr. “The reputational penalty firms bear from committing criminal fraud.” Journal of Law and Economics 36:2 (October 1993): 757–802.

Krishnan, GV, “Did earnings conservatism increase for former Andersen clients?” Journal of Accounting, Auditing and, 2007

LaPorta, Lopez-de-Silanes, and Shleifer, forthcoming, “The Law and Economics of Self-Dealing,” *Journal of Financial Economics.*

Miller, Gregory S. 2006, "The Press as a Watchdog for Accounting Fraud." *Journal of Accounting Research* 44, no. 5 (December): 1001-1033.

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) (Spring): 139-180.

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

Romano, Roberta, 1991, "

Shleifer, Andrei and Robert Vishny, 1997, "A Survey of Corporate Governance"

Thompson, Robert and Hillary Sale, 2003, "Securities Fraud as Corporate Governance: Reflections Upon Federalism," *Vanderbilt Law Review*

Thompson, Robert and Randall Thomas, 2004,"The Public and Private Faces of Derivative Lawsuits," *Vanderbilt Law Review*.

Wang, Tracy Yue, 2011, "Corporate Securities Fraud: Insights from a New Empirical Framework", *Journal of Law, Economics and Organization*,

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

Winston, Clifford, 1998, "U.S. Industry Adjustment to Economic Deregulation," *Journal of Economic Perspectives*," 89-110.

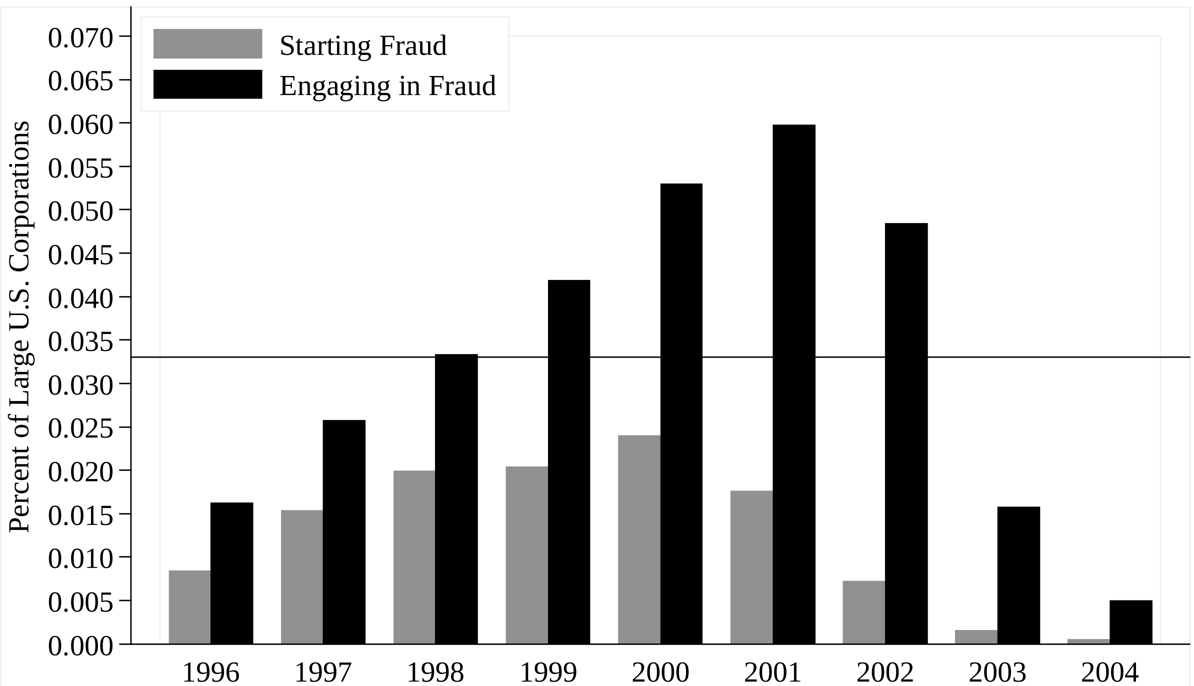

Figure 1: Percentage of Large Corporations Starting and Engaging in Fraud

The grey and black bars respectively report the percent of firms starting and engaging in fraud. The reference line at 0.033 is the overall mean percent of firms engaging in fraud.

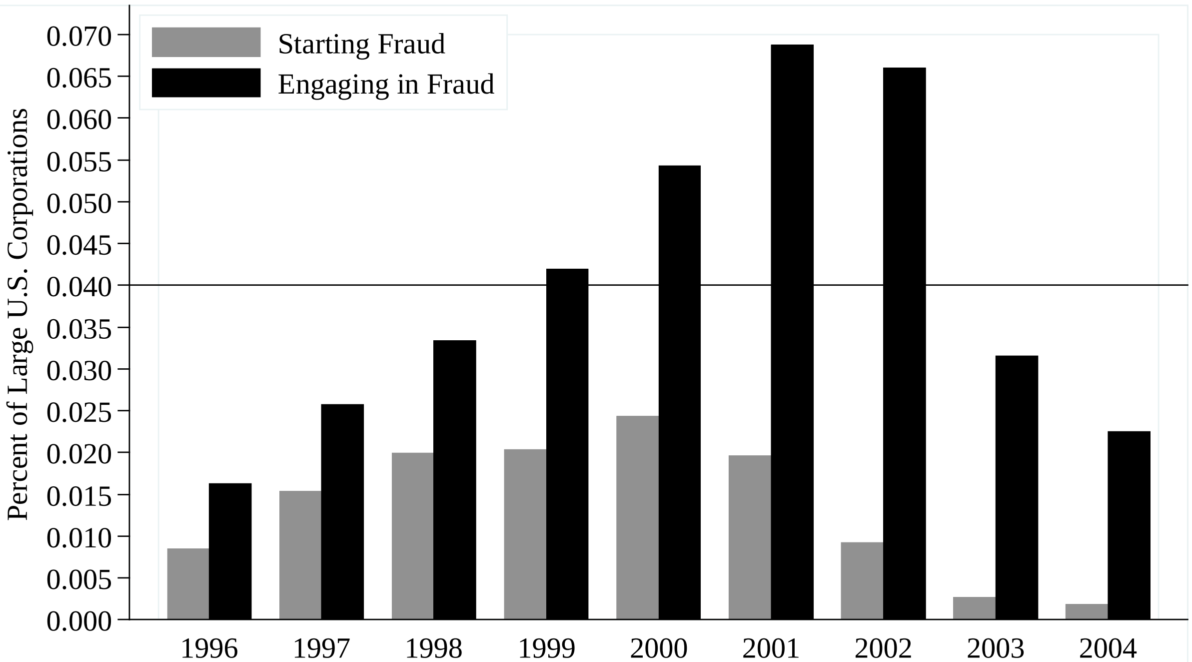

Figure 2: Adjusted Percentage of Large Corporations Starting and Engaging in Fraud

The grey and black bars respectively report the percent of firms starting and engaging in fraud. The reference line at 0.04 is the overall mean percent of firms engaging in fraud. The figure adjusts for the empirical distribution of frauds which in expectation will be caught.

Table 1: Did Arthur Andersen have Clients More Likely to Commit Fraud?

All of the statistics below are for 1998-2000. The columns divide the sample of all Compustat firms with more than $750 in assets into Arther Andersen (AA) clients and otherwise. In the last three columns, otherwise is all non-AA clients which have a Big 5 auditor. Presented are the means and counts, as well as p-values for ttests that the non-AA clients differ from AA clients on each of the statistics. In Panel A, the variables examined are the ProbM Score (probability of manipulation) of Beneish (1999), followed by the eight financial statement components Beneish identifies as making up the scoring of manipulation. The penultimate variables is whether the auditor issued a qualified opinion (which is very rare) or issued a nonqualified opinion with and explanation. The final variable is whether AA firms started fraud more (an indicator for fraud propensity) prior to 2000. In Panel B, the variables are characteristics of the clients; namely, the log of total assets and the ratios of long term debt ot total assets, sales to assets, and EBITDA to sales. Panel C presents OLS and quantile (median) estimations as to whether AA cleints differ on the ProbM Score of Beneish (the dependent variable). ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively. standared errors are in brackets.

Panel A: Univariate Tests for Difference of AA Firms in Manipulation, Fraud, and Auditor Opinion

	AA Firms		All Non-AA			Big 5 Non-AA		
	Mean	Obs.	Mean	Obs.	P-value	Mean	Obs.	P-value
ProbM Score	-2.232	912	-2.230	4,557	0.979	-2.249	3,335	0.704
Days Sales Receivables	1.048	1,008	1.095	4,821	0.597	1.111	3,538	0.544
Gross Margin Index	0.848	1,037	1.007	4,938	0.157	1.011	3,641	0.197
Asset Quality Index	2.452	1,038	2.875	4,943	0.675	2.847	3,643	0.690
Sales Growth Index	1.333	1,037	1.283	4,939	0.240	1.283	3,642	0.295
Depreciation Index	194.6	1,035	286.6	4,940	0.718	247.6	3,641	0.832
SG&A Expense	1.068	940	1.048	4,672	0.694	1.057	3,436	0.854
Accruals/Assets	1.046	1,035	1.112	4,941	0.220	1.126	3,641	0.198
Leverage Index	-0.051	1,095	-0.046	5,224	0.137	-0.047	3,842	0.195
Opinion with Exception	0.188	1,094	0.161	4,558	0.0283**	0.159	3,514	0.0206**
Fraud Started	0.017	1,097	0.016	5,231	0.836	0.017	3,847	0.877

Panel B: Comparison of Size, Use of Debt, and Profitability for pre-2000 AA and non-AA Firms

	AA Firms		All Non-AA			Big 5 Non-AA		
	Mean	Obs.	Mean	Obs.	P-value	Mean	Obs.	P-value
Log Assets	8.078	1,097	8.152	5,227	0.0904*	8.190	3,844	0.013**
LT Debt / Assets	0.324	1,093	0.241	5,226	0.000***	0.248	3,843	0.000***
Sales / Assets	0.811	1,097	0.717	5,225	0.000***	0.754	3,843	0.026**
EBITDA / Sales	0.102	1,095	0.092	5,225	0.001***	0.096	3,843	0.074*

Panel C: Multivariate Tests for Difference of AA Firms in Manipulation (ProbM : dependent variable).

	OLS	OLS	Medians	Medians
Arthur Andersem	0.035	0.024	0.015	0.015
	[0.031]	[0.032]	[0.012]	[0.012]
Log Assets	-0.069***	-0.084***	-0.023***	-0.026***
	[0.009]	[0.010]	[0.003]	[0.004]
Sales / Assets	-0.094***	-0.100***	-0.040***	-0.038***
	[0.017]	[0.018]	[0.006]	[0.007]
EBITDA / Sales	-1.418***	-1.434***	-0.787***	-0.815***
	[0.139]	[0.152]	[0.052]	[0.058]
LT Debt / Assets	-0.240***	-0.318***	-0.111***	-0.155***
	[0.060]	[0.068]	[0.022]	[0.026]
Constant	-1.427***	-1.267***	-2.166***	-2.121***
	[0.081]	[0.094]	[0.030]	[0.036]
Observations	11,033	8,672	11,033	8,672
R-squared	0.022	0.025	0.015	0.015
Sample:	All 1994-2000	Big 5 1994-2000	All 1994-2000	Big 5 1994-2000

Table 2 – Pervasiveness of Fraud Based on AA Natural Experiment

Observed inputs:			Estimate: Detection likelihood		Estimate: Unconditional likelihood of fraud	
From Figure 1:						
Pr (engage, caught) =	4.0%					
Observed frauds are occurring in 4% of large corporations.						
Panel A – Assumption 3a: "Best Estimate"						
Pr (engage*, caught$_{Aud}$*) =	0.38%					
Pr (engage*, caught$_{Aud}$*|AA) =	1.38%					
	Thus…		Pr (caught | engage$_{Aud}$*) =	0.38% / 1.38% = 0.275	Pr (engage)	4.0% / 27.5% = 0.145
			Pr (caught | engage) =	27.5% of fraud are caught	Pr (engage) =	14.5% of firms are committing fraud
Panel B – Assumption 3b: Lower Bound Estimate						
Pr (engage, caught) =	4.44%					
Pr (engage, caught|AA) =	6.21%					
	Thus…		Pr (caught | engage) =	4.44% / 6.21% = 0.715	Pr (engage)	4.0% / 71.5% = 0.056
			Pr (caught | engage) =	71.5% of fraud are caught	Pr (engage) =	5.6% of firms are committing fraud

Table 3: Variables in the Engaging and Detecting Bi-Probit for Incentives and Opportunities Estimation

The table presents the variables, their description, and the source of the data for the Incentives and Opportunities Estimation. The variables are divided into three categories – those included in both the engaging in fraud and detecting fraud equation, those included only in the engaging in fraud equation, and those included in only the detecting fraud equation. High Incentives and Opportunities indicator variables are a subset of the variables which could potentially affect both engaging and detecting fraud.

Panel A: Variables in Both Engaging and Detecting Equations

High Incentives and Opportunities Indicator Variables

Variable	Description	Source
Analyst Coverage Indicator	A dummy variable that takes the value of 1 for companies with higher than the median value of analyst coverage in companies with more than $750 million in assets.	I/B/E/S
Media Coverage Indicator	A dummy variable that takes the value of 1 if the firm has higher than the median value of media coverage in companies with more than $750 million in assets. We manually collect media coverage by searching the Wall Street Journal print edition and recording the number of media hits for the year 1995.	Factiva
Shortability Indicator	A dummy variable that takes the value 1 for companies with a greater than median level of institutional shareholding in the prior year.	Compact-D
Regulated Firm Indicator	A dummy variable that take the value of 1 if the firm is in the following categories: financials, transportation equipment, transportation, communications, electric, gas and sanitary services, drug, drug, proprietaries and druggists sundries, petroleum and petroleum products wholesalers pharmaceuticals, healthcare providers, and healthcare related firms in business services.	Industries identified in Winston (1998) and others
Fortune Best 100 Indicator	A dummy variable that takes the value of 1 if the company is a *Fortune Best 100* firm.	Fortune magazine
Qui-Tam Industry Indicator	A dummy variable that takes the value of 1 if the industry is one in which qui tam lawsuits are possible. Included are healthcare and defense contractor industries.	Civil Division, Department of Justice
Post Sox Indicator	A dummy variable that takes the value of 1 if the time period is post-SOX.	Legislation date

Main Variables in Both Engaging and Detecting Equations

Variable	Description	Source
Company Size	Log of total book assets	Compustat
Stock Return	Total return on stock	CRSP
R&D	R&D expenditures / total assets	Compustat

Panel B: Variables in Engaging Equation Only

Variable	Description	Source
In-Money Exercisable Options	The sum of the in-the-money exercisable options for all executives.	Execucomp
Option & Restricted Stock Grants	The average of the ratio of restricted stock grants divided by total compensation across executives for a firm-year.	Execucomp

Panel C: Variables in Detecting Equation Only		
Abnormal ROA	Residual from regression with *i* denoting company; *j* industry; and *t* time: $ROA_{ijt} = \alpha_0 + \alpha_1 ROA_{ijt-1} + \alpha_2 \overline{ROA_{jt}} + \epsilon_{ROA,ijt}$, where $\overline{ROA_{jt}}$ denotes the industry average. This estimation removes serial correlation and the industry effect.	Compustat
Abnormal Stock Return	Residual from CAPM regression: $r_{it} = r_{ft} + \beta_i(r_{mt} - r_{ft}) + \epsilon_{r,it}$. r_{mt}, r_{it}, and r_{ft} denote the market return, the firm return, and the risk free rate, all in quarter *t*.	CRSP
Abnormal Settlements	Residual from regression with *j* denoting industry, and *t* time: $S_{jt} = \gamma_0 + \gamma_1 S_{j,t-1} + \epsilon_S$. S_{ji} is the sum of settlement dollars including insurance payouts of an industry j in year t.	DMZ
Sarbenes-Oxley Shock	Equals one if the start date is pre-SOX and the period of the potential detection is post-SOX.	Legislation date

Table 4: Pervasiveness of Fraud by Investor Incentives Splits

Panel A

	Percentage of Large Firms Committing Fraud		
	All Firms 1.13%		
Increased Incentives for Detector			**p-value for diff.**
Analyst	Low Analyst Coverage 0.67%	High Analyst Coverage 1.86%	0.000
Media	Low Media Coverage 1.02%	High Media Coverage 1.69%	0.000
Short Sellers	Low Shortability 0.53%	High Shortability 1.24%	0.000
Industry Regulators	Not Regulated 1.39%	Regulated 1.61%	0.287
Employees	Not Fortune Best 100 Firm 1.16%	Fortune Best 100 Firm 2.11%	0.039
Employees	Not Qui Tam 1.35%	Qui Tam 3.01%	0.000

Panel B

Detector	**Percentage of Frauds Detected by (row) in setting (column):**			**p-value for diff.**
Analysts	All Firms 9.1%	Low Analyst Coverage 3.5%	High Analyst Coverage 11.8%	0.000
Media	9.1%	Low Media Coverage 7.8%	High Media Coverage 9.9%	0.326
Short Sellers	4.4%	Low Shortability 0.0%	High Shortability 4.6%	0.000
Industry Regulators	10.1%	Not Regulated 2.9%	Regulated 16.9%	0.000
Employees	13.0%	Not Fortune Best 100 Firm 12.8%	Fortune Best 100 Firm 16.7%	0.167
		Not Qui Tam 11.4%	Qui Tam 18.5%	0.098

Table 5: Partial Observability, Bi-variate Probit Estimates

Reported are estimates from three equations: two independent probit models and the partial observability bivariate probit model of Poirier. Equations 1 and 2 are models of the probability of starting fraud and the probability of fraud being detected, respectively. Equation 3 is the jointly estimated Poirier version of detection given that a fraud is start. Coefficients are reported as well as the marginal effects for each model. The final model shows the marginal effects as well as the conditional marginal effects,. These estimated a interpreted as a one unit change in the independent variable implies that marginal effect on the probability of detection for companies already engaging in fraud.

	(1)		(2)		(3)		
	Probit (Start=1)		Probit (Caught=1)		BiProbit(Caught, Start)		
	Estimates	MFX	Estimates	MFX	Estimates	Pr(Start),P(Caught)	P(Caught|Start)
Start Equation							
Log Assets	0.055	0.0020			0.471***	0.0229	
	[0.033]				[0.116]		
Stock Returns	-0.062	-0.0023			-4.192***	-0.2038	
	[0.068]				[0.763]		
Log R&D Expenditures	-0.009	-0.0003			0.025	0.0012	
	[0.017]				[0.057]		
Log Options Held	-0.001	0.0000			-0.073**	-0.0036	
	[0.013]				[0.037]		
RestrictedStock+ Option	0.596***	0.0224			0.784**	0.0381	
Grants % Comp.	[0.187]				[0.378]		
Annual SEC Fines	0.004	0.0001			-0.001	0.0000	
	[0.002]				[0.006]		
Hi Analysts (Lagged)	0.235**	0.0081			-0.016	-0.0008	
	[0.100]				[0.348]		
High Media (1995)	-0.120	-0.0048			0.808**	0.0392	
	[0.093]				[0.316]		
Hi Shortability (Lagged)	0.106	0.0039			-1.032***	-0.0502	
	[0.089]				[0.313]		
Fortune Best (Lagged)	0.065	0.0026			0.204	0.0099	
	[0.166]				[0.549]		
Healthcare,Qui Tam	0.330***	0.0170			-1.834***	-0.0892	
	[0.127]				[0.574]		
Regulated	-0.182	-0.0065			0.928**	0.0451	
	[0.115]				[0.414]		
PostSOx	-0.269	-0.0092			-1.452***	-0.0706	
	[0.179]				[0.474]		
Caught Equation							
Hi Analysts (Lagged)			0.375***	0.0139	0.885***	0.0797	0.1766
			[0.081]		[0.307]		
High Media (1995)			-0.111	-0.0043	-0.394	-0.0355	-0.1775
			[0.078]		[0.258]		
Hi Shortability			0.144*	0.0055	1.070***	0.0964	0.3386
			[0.075]		[0.276]		
Fortune Best (Lagged)			-0.585**	-0.0128	-0.920	-0.0828	-0.2067
			[0.268]		[0.576]		
Healthcare,Qui Tam			0.374***	0.0207	3.005***	0.2707	0.8195
			[0.120]		[0.596]		
Regulated			-0.011	-0.0004	-0.451	-0.0406	-0.2035
			[0.094]		[0.316]		
PostSOx			0.279***	0.0122	1.091***	0.0982	0.3946
			[0.069]		[0.287]		

Table 5 Continued

Abnormal Return (Lagged)		0.494*	0.0189	-1.518***	-0.1367	-0.2995
		[0.277]		[0.430]		
StDev Return		0.186**	0.0071	4.069*	0.3666	0.8030
		[0.081]		[2.105]		
StDev S&P 500		48.30**	1.8480	174.5**	15.72	34.44
		[19.18]		[68.21]		
Log Assets (Lagged)		0.055**	0.0021	-0.060	-0.0054	-0.0119
		[0.027]		[0.0850]		
Stock Return (Lagged)		-0.211***	-0.0081	0.693***	0.0623	0.1367
		[0.072]		[0.197]		
Log R&D Expenditures (Lagged)		0.013	0.0005	-0.014	-0.0012	-0.0027
		[0.016]		[0.0472]		
Observations	6349	8287		6347		
Pseudo R-squared	0.063	0.066				

Inferences		
Unconditional Probability of Starting a Fraud Estimate	5.3%	5.3%
Average Duration of Frauds (from DMZ)	1 year 8 months	1 year 8 months
Unconditional Probability of a Firm Engaging in Fraud Estimate		8.85%

Unconditional Probability of a Firm Engaging in Fraud Estimate when Turning on Opportunities & Incentives:

PreSOx	11.3%
PostSOx	3.4%
HiAnalysts	5.9%
LowAnalysts	14.2%
Healthcare/Qui Tam	1.3%
Not Healthcare/Qui Tam	9.3%
HiShortability (High Institutional)	6.8%
LowShortability (Low Institutional)	12.0%

Note: These should not sum to the mean overall, because of the two-equation model.

Low Shortability, PreSOx, Low Analyst, Not Healthcare	15.8%

Table 6: Pervasiveness of Fraud in a Survey of MBAs

We asked MBAs entering the University of Chicago whether they faced a legal dilemma in their jobs before they joined the MBA program, where we defined "legal dilemma" as "In your job you are asked to do something that is illegal. Example: Your boss asks you to lie in reporting sales." Panel A reports the percentage of MBAs who responded positively by industry, and Panel B reports the percentage by occupation or function in their jobs.

Panel A:

Industry	***Illegal***	**N**
Consulting	11.76%	51
Consumer goods	6.67%	15
Financial services	15.08%	126
Health/Pharmaceutical	14.29%	14
Other	18.18%	77
Total	14.84%	283

Panel B:

Function	***Illegal***	**N**
Accounting	11.11%	18
Consulting	11.54%	52
Corporate - Finace	15.00%	20
Corporate-Sales	13.33%	15
Corporate - Product Management	12.50%	8
Corporate -Other	33.33%	21
Investment Banking	16.67%	42
Investment Management	11.11%	18
Other	13.48%	89
Total	14.84%	283

Table 7: Cost of Fraud

The table presents the statistics and results from the counterfactual exercise to estimate the cost of corporate fraud. Panel A reports the statistics of the enterprise value of firms, and the equity and long term debt componets, as well as the financial statement line items which enter the multiples analysis. The statistics are reported for fraud firms of DMZ's original sampel of 216 firms which have statistics for pre and post periods. (The pre and post columns represent the same set of firms, but the counts vary slightly by which balance sheet items might be missing.) Median and means are in millions of USD. Panel B presents the median multiples corresponding to the data in panel A. The frequency counts line up to panel A. The industry multiples are at the SIC 2-digit level. The final column in panel B is the counterfactual result we use to calculate the counterfactual enterprise value, using the pre periof firm actual multiple and the growth in the corresponding industry multiple. Finally, Panel C reports the cost of corporate fraud per firm, corresponding to the calculations of counterfactual multiples. The first two columns report unadjusted costs, and the final two columns are the adjustments to subtract out legal costs when firms are never costs, following the estimates in Karpoff, Lee and Martin (2010).

Panel A: Statistics

	Pre-Fraud			Post-Fraud		
	Median	Mean	Frequency	Median	Mean	Frequency
Equity	5,120	15,632	198	2,214	9,458	198
Long Term Debt	812	2,871	196	1,068	6,160	196
Enterprise Value	6,812	18,513	198	4,256	15,556	198
EBITDA	348	1,086	195	328	1,383	195
Sales	2,417	6,690	194	3,378	8,295	194
Assets	3,850	14,683	198	4,350	21,583	198
Fixed Assets	623	2,565	195	813	3,439	195

Panel B: Multiples

	Pre-Fraud			Post-Fraud		
	Median Firm Actual	Median Industry		Median Firm Actual	Median Industry	Firm Counterfactual
EBITDA Multiple	12.31	10.39		9.49	10.68	13.25
Sales Muliple	2.31	1.85		1.43	1.92	2.19
Assets Multiple	1.38	1.12		0.82	1.05	1.27
Fixed Assets Multiple	9.34	5.58		5.02	5.88	9.13

Panel C: Costs

	Cost		KLM Adjusted Cost	
Based on:	$ million	% of Enterprise Value	$ million	% of Enterprise Value
EBITDA Multiple	1,099	23.2%	923	20.4%
Robustness:				
Sales Muliple	2,071	42.9%	1,888	40.4%
Assets Multiple	2,132	35.5%	1,801	34.3%
Fixed Assets Multiple	2,306	38.3%	2,079	37.5%