Preliminary
Do Not Quote Without Permission

HOW PERVASIVE IS CORPORATE FRAUD?

Alexander Dyck
University of Toronto

Adair Morse
University of Chicago

Luigi Zingales*
University of Chicago, NBER, & CEPR

August 2011

ABSTRACT

After building a dataset of all corporate frauds in large corporations that impact shareholder value and are caught, we 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. Using the quasi natural experiment of the forced turnover of auditors following Arthur Andersen's demise we find between 4.7 and 9.7% of firms initiate a fraud. Using a broader set of circumstances that enhance activism by other fraud detectors, we arrive at an alternative estimate of 7%. By using industry multiples, we estimate the median cost of a fraud is 40.7 percent of the pre-fraud enterprise value of the company. Hence, taking into account the overall incidence of fraud, we estimate that in publicly-traded companies with more than 750M in assets, corporate fraud costs 2.85 percent of enterprise value.

---

*Alexander Dyck thanks the Connaught Fund of the University of Toronto 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 U.S. was considered to set the corporate governance standard towards which other countries aspired. The major wave of corporate scandals that emerged at the beginning of the new millennium deeply shook this confidence. How is it possible for a company like HealthSouth to falsify its financial statements for 11 years in a row without anybody noticing? How is it possible for WorldCom to transform 3.8 billion of expenses into capital investments? How could Andrew Fastow enrich himself while hiding billions of liabilities from the eyes of investors? More importantly, 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 in directing policy such as the legislative reforms in Sarbanes Oxley and more recent efforts in Dodd-Frank. If there are just a few rotten apples, then there is no need to intervene massively (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 massive intervention to fix the problem might be needed.

Prior research provides some indicators that could be used to size up the pervasiveness of governance problems, but each of these has significant limitations. The extent of financial restatements across US publicly traded firms[1], for example, provides one measure of governance problems but the raw data overstates the problem by including many examples that would not be considered governance weaknesses. More importantly, by design restatements will not capture the many governance violations that do not require manipulating numbers in the financial statements (e.g. lying about the future or about the challenges of integrating recent acquisitions), and no effort is made to estimate the potential extent of financial manipulations that go undetected. Other efforts to size up the problem that look at clear examples of governance

[1] A common source for such restatements is those companies identified by the GAO (2002) as used in, for example, Palmrose and Scholz (2004).

violations (e.g. those companies sanctioned by the Securities and Exchange Commission)[2] will likewise miss many violations given the incentives and limited budget of the SEC, and again do not allow one to identify the full extent of manipulations that go undetected.[3]

In this paper we provide an answer to the question of the pervasiveness of fraud by building on the analysis of Dyck, Morse and Zinales (2010) who developed a comprehensive sample of frauds that involves restatements as well as those that do not involve restatements. To identify all the possible frauds - which we define as any *material* violation of the law - we rely on the fact that the security class action system provides strong incentives to file suits whenever a fraud is revealed. Large publicly traded companies are primary targets of these suits. Hence, for large companies the Stanford Security Class Action database provides a very comprehensive set of material violations of the law. Our approach goes well beyond the annual counts of security class actions though by including a set of screens involving objective and subjective criteria designed to eliminate frivolous suits and suits that do not appear to be related to any illegal action.

This approach, however, does not eliminate the fact that any dataset, no matter how comprehensive it is, will include only the frauds that have been caught. As such, any list is unable to identify the potential submerged part of the iceberg. To address this problem we build on the fact that basic probability rules suggest how we can go from observed data of the joint event of a fraud starting and being caught, to our actual variable of interest, those frauds that are started, regardless of whether they are caught or not. Specifically, by Bayes rule, the unconditional probability of starting a fraud, $Pr(start)=Pr(start,caught) / Pr(caught|start)$. Thus,

---

[2] Specifically, studies have focused on Accounting, Auditing and Enforcement Releases, such as Dechow, Sloan and Sweeney (1996) and Miller (2006). A further description of such studies is provided in the Data Appendix.

[3] This is a decidedly incomplete listing of a voluminous literature, that we will describe in more depth in the next version of the paper.

if we knew the probability that a fraud is caught, given that it is started, Pr(*caught*|*start*), we could easily calculate the Pr(*start*). Our overall identification approach is to exploit circumstances in which the likelihood of being caught increases, impose on these situations that the Pr(*caught*|*start*) approaches one, and thus estimate a revised lower bound on the unconditional pervasiveness of fraud.

Our first test focuses on the quasi-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 enhanced the incentives of auditors to be active. Dyck, Morse, and Zingales (2010) show that the probability of detecting a fraud increases after a turnover of the external auditors. We find that following this forced turnover the probability a fraud is detected in these firms raises from 1% to 1.85%. Given that auditors are able to identify only a fraction of the existing frauds, we estimate that the actual amount of fraud started varies between 4.74% and 9.74%.

Our second test utilizes the dataset introduced by Dyck, Morse, and Zingales (2010) that shows frauds and the mechanisms responsible for their revelation. 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. By using these differences in the probability a fraud is revealed, we can infer the probability that a fraud is committed. By using this method we arrive at an estimate that for a large publicly traded firm, the probability of starting a fraud is equal to 7%, close to the mid-point of the range suggested by the Arthur Andersen experiment.

Using 7 percent as our point estimate and making a few additional assumptions, and exploiting the knowledge that the average duration of frauds is between 1.6 and 1.9 years, we estimate that the fraction of publicly traded firms in which fraud is taking place at a given time varies from 11.2 to 13.2 percent of firms.

To size up the extent of the governance problem in the US it is helpful to combine this estimate of the pervasiveness of corporate fraud with an estimate of the cost this imposes on society. The amount of damages alleged in legal suits does not represent a good measure of such costs, because many of the dollar losses are transfers rather than social losses. If a group of shareholders buy at an inflated price, they suffer a loss, but the shareholders selling make a corresponding gain. The change in market value at the announcement of a fraud is not a comprehensive measure of the social costs of fraud either, since fraud are often committed to cover up negative news, which would have been revealed to the market earlier in absence of fraud. For these reasons we construct a new measure of the cost of fraud, which we define to be equal to the difference between what the enterprise value of the company would have been in the absence of fraud, and the enterprise value after the fraud is revealed. We construct this hypothetical value of what the enterprise value would have been in the absence of fraud by making projections from the pre-fraud period, assuming the trajectory would have followed that of other firms in the same industry.

Using this method, we estimate that the median loss is 40.7 percent of the enterprise value of our fraud companies, using their enterprise value prior to the beginning of fraud. This implies that 2.85% of the enterprise value of companies is lost to fraud each year (7%*0.407).

The rest of the paper proceeds as follows. Section 1 describes the main data and provides a conservative estimate of fraud pervasiveness based on frauds that are caught. Section 2 presents revised estimates of the incidence of fraud including estimates of 'non-caught' frauds based on the natural experiment of the demise of Arthur Andersen and other cross sectional variation in incentives for other fraud detectors. Section 3 uses the estimates from section 2 and the duration data described in section 1 to produce estimates of fraud pervasiveness. In section 4 we describe the method we used to estimate the costs of fraud and provide our estimation of damages in our sample and in the population of firms with more than $750 million in assets. Section 5 concludes.

I. A Conservative Estimate of Fraud Pervasiveness Based on Caught Frauds

To establish a conservative base-line estimate of the pervasiveness of corporate fraud in US publicly-traded firms, we start with a sample of observed frauds that we believe is close to the population; namely, those identified by securities fraud class action suits in the study of Dyck, Morse and Zingales (2010) described in Table 1. What makes this sub-sample so complete is the incentive for law firms to pursue suits. Class action law firms have automated the mechanism of filing class action suits so that they start searching for a cause to file a suit every time there is a negative shock to share prices. Since stock prices drop following revelation of most serious corporate frauds, it is highly unlikely that a corporate fraud would emerge without a subsequent class action suit being filed (Coffee, 1986). We employ the Stanford Securities Class Action Clearinghouse (SSCAC) data, which are the most comprehensive database of class action securities fraud suits.

The biggest potential problem with using class action data is not that it misses important frauds, but rather that such an approach might be over-inclusive (i.e., containing some allegations that are frivolous). To address this concern Dyck, Morse and Zingales (2010) employ six filters. 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. Operationally, they restrict our attention to 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.[4] 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.[5] Fifth, they exclude from our analysis those security frauds that Stanford classifies as non-standard, including mutual funds, analyst, and IPO allocation.[6] 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[7].

While we use the term fraud in considering these cases, strictly speaking these are only examples of alleged frauds. Settlements almost always involve no admittance of wrongdoing.

[4] 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.
[5] 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.
[6] 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.
[7] 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.

As a result, it is impossible for us to establish whether there was real fraud (which in legal terms implies the intent to deceive) or just gross negligence. For the purpose of this paper, however, this difference is not so relevant. We are interested in understanding the level of incidence of bad forms of governance, not in establishing intent. For simplicity, in the rest of the paper we do not use the adjective "alleged". The appendix relates this sample to other samples of fraud used in the literature.

In total the Dyck, Morse and Zingales (2010) sample includes 230 frauds that are detected in the 1996-2004 period that satisfy their selection criteria. These frauds include all of the high profile frauds such as Enron, Worldcom, Adelphia and Healthsouth, as well as many others. These firms have median enterprise value in the pre-fraud period of $5.6 billion and a median equity value of $4.48 billion.

The class action database provides start and end dates for the frauds. Because these dates can and often are revised as suits progress, we follow Dyck, Morse and Zingales (2010) and 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 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. We later correct the bias for all firms by estimating the average length of attenuation of fraud period for the part of the sample for which restatements were necessary. The frauds in our sample have an average duration (from the class action suit period) of approximately 1 year and 7 months (587 days).

Finally, to address questions of 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 US publicly-traded companies with $750 million in assets in the prior year. In Compustat, 2,149 companies, on average, per year meet this criterion. These Compustat firms are smaller than our fraud firms, with a median asset value of 2.5 billion.

Figure 1 illustrates the incidence of caught frauds. In the figure we plot the percentage of US large publicly traded companies that start fraud in each year (black line) and the percentage of firms where fraud is taking place in that year including new starts and ongoing frauds (gray line). This evidence suggests a non-trivial level of fraud taking place, with an average of 1.1 percent of firms starting fraud each year and 2.9 percent of firms involved in fraud at any one point in time. Note the significant time series variation in these numbers, with the incidence of firms starting fraud peaking in 2000, when 2.1% of firms started frauds that year, and the fraction of firms exposed to fraud peaking in 2001 when 5.3% of firms experienced fraud.

Figure 2 provides an even better indication of the incidence of frauds in our sample as it introduces a correction for the fact that some frauds that will be caught after 2004 were taking place during our sample period. Specifically, we use the distribution of fraud durations for those cases which begin prior to the year 2000 to forecast how many cases are yet to be caught. That is, for the pre-2000 cases we calculate what percent of cases are caught within 1 quarter, 2 quarters, and so on up to a maximum of 20 quarters, assuming that all frauds are caught within five years. Using the duration distribution, we then roll the distribution forward to forecast how many additional cases that began after 1999 will yet be caught.[8]

[8] The percent of frauds caught for each of the twenty quarters, starting in the quarter of fraud commencement and ending in the quarter five years since the fraud started is: {0.093, 0.178, 0.186, 0.081, 0.081, 0.006, 0.011, 0.059, 0.102, 0.034, 0.034, 0.034, 0.028, 0.023, 0.051}

This correction raises our estimate of the overall incidence of frauds being started to 1.3% of firms per year and the overall fraction of firms experiencing fraud in any one year to 3.2% of firms. As expected, the correction has little to no effect on frauds in 2000 and before, but affects our reporting of frauds since then. The data continue to show significant time series variation with a much higher incidence of frauds starting prior to the passage of SOX in 2002.

II. Revised Estimates Including 'Non-Caught' Frauds

This picture is incomplete as it ignores the fact that some frauds are never caught. In short, without exploring the likelihood of frauds taking place without being caught, we don't know if our prior estimates are the whole iceberg, or just the tip of the iceberg. Basic probability rules suggest how we can go from observed data of the joint event of a fraud starting and being caught, Pr(*caught*|*start*), to our actual variable of interest, those frauds that are started, regardless of whether they are caught or not. Specifically, by Bayes rule, the unconditional probability of starting a fraud, Pr(*start*)=Pr(*start,caught*) / Pr(*caught*|*start*). Thus, if we knew the probability that a fraud is caught, given that it is started, Pr(*caught*|*start*), we could easily calculate the Pr(*start*).

Our overall identification approach is to exploit circumstances in which the likelihood of being caught increases, impose on these situations that the Pr(*caught*|*start*) approaches one, and thus estimate a revised lower bound on the unconditional pervasiveness of fraud. In the Appendix, we supplement this evidence based on firms with some survey evidence from MBAs, producing remarkably similar incidence using a quite different technique.

II.1 The Natural Experiment of Arthur Andersen

The first circumstance we exploit is the quasi natural experiment provided by the sudden demise of Arthur Andersen.   Specifically, we look at the increase in the probability that a fraud is revealed after the forced turnover of the external auditor triggered by the demise of Arthur Andersen.   Dyck, Morse, and Zingales (2010) show that the probability of detecting a fraud increases after a turnover of the external auditors. Following its incrimination, all Arthur Andersen's clients had to change their external auditor. This turnover, thus, represents a natural experiment. In important advantage of this setting, is that the line of causality is clear, as the turnover leads to the fraud revelation, rather than the fraud leading to the turnover.

To examine the impact of the forced change in auditor, we focus on all firms with more than $750 million in assets and identify those that had Andersen as their auditor in 2001. All of these firms have to change their auditor starting in 2002, and we compare their incidence of fraud before Andersen's demise and after when auditors have a strong incentive to bring to light frauds.

In Table 2 panel A, we report that firms with Arthur Andersen as their auditor in 2001 had a 1.85% chance of being caught with a fraud in 2002-2004, statistically different that the 1.05% probability of being caught for non-Arthur Andersen clients. Panel B of Table 4 tells us that in the extreme case of auditor turnover, nearly one quarter (23.8%) of frauds were found by auditors. The 23.8% is a stark (and significant) increase from the relative role of auditors (9.7%) for non-Arthur Andersen firms. We can conclude that, at the maximum, auditors have the power to find one-quarter of frauds committed, but that in normal situations they only reveal ten percent (9.7%).

We then use this panel B finding to help us understand the economic significance of the 1.85% found in panel A.   Consider the auditing firms in 2002 who watched the breakup of

Arthur Andersen. It is unlikely that any of the new auditors would have overlooked any firm mis-doings that were in their power to find. Dirty laundry could easily be blamed on Arthur Andersen. What is the magnitude of the increase in observed fraud? Using the percentiles from panel B, we know that for non-Arthur Andersen cases, 1.05% of firms are caught with fraud, or 0.10% of firms are caught with fraud by auditors. For Arthur Andersen firms, 1.85% of firms are caught, or 0.44% of all firms are caught with fraud by auditors. The figure 0.44% is greater than a four-fold increase over the 0.10%. We can conclude that the exogenous shock caused four-fold more frauds to be discovered that had started that may not have been otherwise caught, or pr(*caught*|*start*) = 0.10%/0.44% = 0.227. Using the overall pr(start, caught) = 1.13% from the entire sample, we can infer that pr(*start*) = pr(*start*, *caught*) / pr(*caught*|*start*) = 1.13% / 0.227 = 4.74%. The shock of Arthur Andersen suggests that fully revealing situations would uncover four-fold more frauds being committed, or that 4.74% of all firms start to commit fraud each year.

The increase of 0.105% to 0.440% is a very conservative increase, however, since auditors may work behind the scenes in bringing fraud to light. In particular, note that of the increase from 1.05% to 1.85%, only about half (0.44%) was attributable to an increased activity by auditors. If auditors acted behind the scenes to bring fraud out for Arthur Andersen transitioning firms, then the most that auditors could have impacted the detection is the difference between 1.85% and 1.05%, or 0.80%. In such a case, shock of Arthur Andersen suggests that fully revealing situations would uncover nine-fold more frauds being committed (from 0.10% to 0.10% + 0.80%), or that 9.74% of all firms start to commit fraud each year.

In sum, we find that following this forced turnover the probability a fraud is detected raises from 1% to 1.85%. Given that auditors are able to identify only a fraction of the existing frauds, we estimate that the actual amount of fraud started varies between 4.74% and 9.74%.

II.2 Regression Estimates to Identify 'Non-Caught' Frauds

Dyck, Morse, and Zingales (2010) shows that 69% of frauds are revealed by a core set ("it takes a village") of fraud detectors -- insiders, analysts, employees, auditors, media, short sellers and non-financial regulators. Our identification strategy in this section is to find situations in which the whistle blowing incentives or opportunities are high for each type of detector. (For example, the media will uncover fraud more easily if the media coverage of that firm is high.) Then, we simply estimate the increase in the likelihood of observing fraud for a firm facing high fraud detection situation for each of our village of fraud detectors.

When are incentives for fraud detection high? The answer is going to be different for each type of fraud detector. We use a series of variables each of which is intended to capture a situation of heightened revealed preference to scrutinize, a larger payoff from fraud detection, or greater outside mandate to access the information for a particular fraud detector type.

Specifically, to capture the situation when the payoffs for analysts are high 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. For the media, we again rely on observing when coverage of particular firms by media is high. Since we lack an equivalent to I/B/E/S, 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 hit for the year.[9] For short selling, we follow the literature and use institutional

[9] We eliminate lists which are automatically generated (e.g., large stock movers), and we manually check each firm whose company name contains common language words (e.g., Apple). The range of media coverage is from zero

ownership as our proxy with data for Compact D. For each of these variables we use the median to identify firms exposed to high (above median) and low (below median) incentives.

Regulator attention is a simple dummy for firms with a regulator or not. For employees, we introduce two proxies, first whether the company was a "Fortune Best 100" firm to work for, which captures an environment that would be less likely penalize whistleblowing, and second whether the firm is in an industry where *qui tam* lawsuits are possible that provide the possibility of the employee receiving 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).[10] Finally, for insiders we use a dummy to identify if the infraction took place pre-SOX or post-SOX.

This method ignores firms' reactions to facing higher incentives for whistleblowing. Although it is unlikely that the companies are completely ignorant, companies knowingly under greater scrutiny will be less inclined to commit fraud and will exert more effort to hide fraud when they do commit it. In either case, fewer frauds should be detected, and our estimate of the pervasiveness of fraud will be biased downward.

Table 3 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

(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). In a future version, our measure will be dynamic with 1995, 1997, 1999, and 2001 media hits measured.

[10] To identify these industries we searched the data on qui tam lawsuits available from the Department of Justice Civil Division, and identified those industries that account for the vast majority of qui tam suits and settlements. This is almost exclusively provided by companies in the healthcare and defense contractor industries.

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.

In Table 4 we go beyond the univariate analysis by using the series of indicator variables simultaneously and estimating the logit model: $pr(Caught_{it}) = f(HiAnalyst_{it}, HiMedia_{it}, HiShortability_{it}, BestFortune_{it}, Post\ Sox_{it}, Regulated_{it}, QuiTamAble_{it})$. Column 1 shows the baseline probability of starting a fraud in the sample with available information for all indicator variables is 1.65%. In column 2 we see that when combined, many of the indicator variables remain significant including high analyst coverage, high media coverage, post SOX, regulated and in a qui tam industry.

Most interesting for our purposes however is not the individual coefficient estimates, but rather how these results allow us to estimate the predicted probability of starting fraud. We calculate the estimated probability of fraud starting by putting all variables to the high incentive state and using the estimated coefficients. This produces a significantly increased predicted probability of starting a fraud of 7.54%, more than 4 times the baseline probability.

Of course, the model used in column 2 may be too simple in that the detector incentive variables may be picking up the effect of omitted variables. Detector incentives, which are often serially correlated, may be related to the incentive to start a fraud. To address this concern in column 3 and 4 we include a first-stage estimation of the incentive of firms to start committing fraud and then use the firm-year level predicted probability of starting a fraud as an explanatory variable in the main estimation.

Our first stage equation follows the standard Becker formulation on crime, where the probability of starting a fraud is a function of the expected payoff and penalties from fraud. We include four variables. First, we hypothesize that the incentive to start a fraud is higher the more the executives' compensation contract depends upon creating and maintaining a high stock market price. As our proxy for penalties we again introduce a dummy for the post SOX environment, when governance monitors were more active and executives feared the penalties associated with misgovernance. For similar logic, we use the average settlement value paid to class action shareholders for securities fraud the prior year. Finally, we include the firm's P/E ratio.

Again, most important for our purposes is whether these corrections influence our estimate of the probability of starting a fraud. The estimate of 7.28% provided in column 3 shows that again assuming that all of the detector incentive variables are in their high state and including a predicted start variable, the estimated probability of starting is almost unchanged. Using our measures of detector incentives also in the first stage regressions, as we do in column 4, produces an estimated probability of starting a fraud of 6.87%. All of the regressions taken together suggest a value for frauds starting of 7 percent a year. This number is close to the mid-point of the range of estimates using the Andersen experiment, and we will use this as the point estimate in what follows.

III. From Incidence to Pervasiveness of Ongoing Frauds

Having established the incidence of firms starting fraud each year, we can combine this information with evidence on the duration of fraud to provide an estimate of the percentage of firms with ongoing fraud at a point in time. Using the average duration data from the class action filings, this suggests that 11.2% of firms (0.07*1.6 years duration) are involved in ongoing fraud at any point in time. If we use instead the longer duration provided by the financial restatement data to define our duration, we get an even higher estimate of 13.2% (0.07*1.9 years duration).

IV. How Expensive is Corporate Fraud?

IV.1. A Method for Calculating the Cost of Fraud

The results in Section II suggest that 7% (with a likely range of 4.74% to 15%) of U.S. corporations commit fraud every year. Only 1.13% of corporations are caught. If we are going to conclude that detected and undetected frauds are a point of concern, we first should address whether it matters in an economic sense that fraud is committed in the first place. To address the economic significance of fraud, we turn to the second objective of the paper – assessing how costly corporate fraud is.

There are a number of possible methods to calculate the cost of corporate fraud to stakeholders. The simplest method would just be to add up the settlement amounts paid to shareholders and the fines incurred in SEC or judicial actions. This method is incomplete on a couple of dimensions. The securities settlements are a function of how many [equity] investors were hurt by the artificial pricing of the stock. The cost to long-term shareholders and to debt holders would not be captured. In addition, this method fails to capitalize the cost of the market's revised mistrust of management's use of assets.

An alternative method would be to look at the decline in equity and debt values at the moment of fraud revelation. A problem with this method is in choosing the exact timing of the value calculation and in knowing what loss is attributable to the fraud versus to subsequent asset changes supporting the debt and equity.

We choose to follow Berger and Ofek (1995) using a multiples approach. Multiples also can be effective in capturing the long-term consequence to fraud that is embedded the total value of the firm. We modify the multiples approach, however, such that we are aligning assets and sales to the value calculation (the benefit from multiples) without assigning value to firms solely based on a standard industry multiple (perhaps a weakness of multiples when considering only a sample of firms). Specifically, our calculations are as follows.

Assume that a fraud begins right after time *s* and ends right before time *t*. We consider two gauges of how much value the firm should create – value from a fixed asset multiple and value from a sales multiple.[11] In simple notation, for firm *j* we consider firm value multiples based on Y = {*sales, fixed assets*}, where

$$Firm\ Mulitple_{js} = \frac{Y_{js}}{Long\ Term\ Debt_{js} + Equity_{js}} \equiv m_{js}$$

In addition to firm multiple m_{js}, we define an industry multiple appropriate for firm *j*, M_{js}. To calculate M_{js}, we first take the mean industry multiple for each SIC 3-digit industry. We then gather the sales from the Compustat Business Segment files and identify the set of industries for which each firm has 3-digit SIC sales. We then create a weighted average multiple for each of our fraud committing firms where the weights are done according to sales by segments. We do the procedure for the time period *s* defined as the year prior to the start of the fraud and for time period *t*, defined as the year following the fraud.

Rather than using a multiple to calculate value directly, we use the industry multiple as the benchmark for how the firm's multiple should have evolved over the time period. The idea is to compare the firm's value of debt and equity at time *t* with the debt and equity which would be

[11] Berger and Ofek (1995) also use an income multiple. We have looked at the EBITDA multiple but, unfortunately a usable multiple is only available for 114 of our sample firms as a result of bankruptcies and other data issues.

projected by the firm's pre-fraud multiple adjusted to a growth or decline rate in its industry benchmark multiples.

$$Loss_{jt} = Long\ Term\ Debt_{jt} + Equity_{jt} - \frac{1}{m_{js}} \frac{M_s}{M_t} Y_{jt}$$

IV.2. Results: The Cost of Fraud

Table 5 presents the results from this analysis. The table reports results at the median, 25th and 75th percentiles, with our focus on the impact for the median firm involved in fraud. We first report results using sales multiples, then asset multiples and EBITDA multiples. Given the dramatic reduction in the number of firms with EBITDA multiples arising from losing firms through bankruptcy and negative earnings, we focus on the sales and asset multiples.

We estimate the loss associated with fraud for the median fraud firm in our sample with available data on sales at 1.57 billion, using fixed assets as 1.58 billion and using EBITDA at 236 million. These numbers are better understood if expressed as a percentage of the enterprise value of the companies prior to the onset of fraud. By this measure, fraud destroys 38.8 to 42.6 percent of enterprise value using sales and fixed assets, which averages out to 40.7 percent of enterprise value (for the reasons mentioned above we do not focus on EBITDA multiples). This estimate can be applied to the population of publicly-traded firms with more than $750 million in assets, taking into account that only 7% of firms start a fraud each year. Doing so, we calculate that the expected loss arising from fraud in US firms amounts to 2.85% of enterprise value. (i.e. .407*.07). One can also express this as a percentage of the equity value of companies. In the fraud sample, the median equity/ EV ratio is 0.81, resulting in an expected loss as a percentage of equity value of 3.52 percent.

How reasonable are these estimates? Interestingly, this estimate is similar to the estimates of the extent of private benefits derived from looking at control premia. Dyck and

Zingales (2004) estimate the control premia in US firms to vary from 2 percent of the value of equity for the median firm (raw data) to 4.4 percent (including controls), bracketing the estimate we derive in this study.

V. Conclusions

In this paper we set out to answer the question of the pervasiveness of corporate fraud in the United States. To address this question, we first seek to establish the incidence of fraud, and next the cost of these frauds.

To establish the incidence of fraud we build a dataset of all corporate frauds in large corporations that impact shareholder value and are caught. Combining this information with the analysis in Dyck, Morse and Zingales (2010), we 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.

Our main result is that we estimate that approximately 7% of firms commit fraud every year, based on the increased probability of a fraud being revealed following the forced turnover of external auditors and exploiting circumstances that increase fraud detector activism using the Dyck, Morse and Zingales data.

Having established the incidence of fraud, we then explore the cost of fraud . We do so by introducing a methodology that compares the value of the firm post fraud with what it would have been if it had followed industry trends from its pre fraud multiple. Using this technique, we estimate the median cost of a fraud is 40.7 percent of the pre-fraud enterprise value of the company. Hence, taking into account the overall incidence of fraud, we estimate that in publicly-traded companies with more than 750M in assets, corporate fraud costs 2.85 percent of enterprise value.

References

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

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

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.

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 and Luigi Zingales, 2004, "Private Benefits of Control: An International Comparison", Journal of Finance 59: 537 - 600.

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

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.

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.

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., 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.

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*.

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

Figure 1: Pervasiveness of Starting and Ongoing Frauds

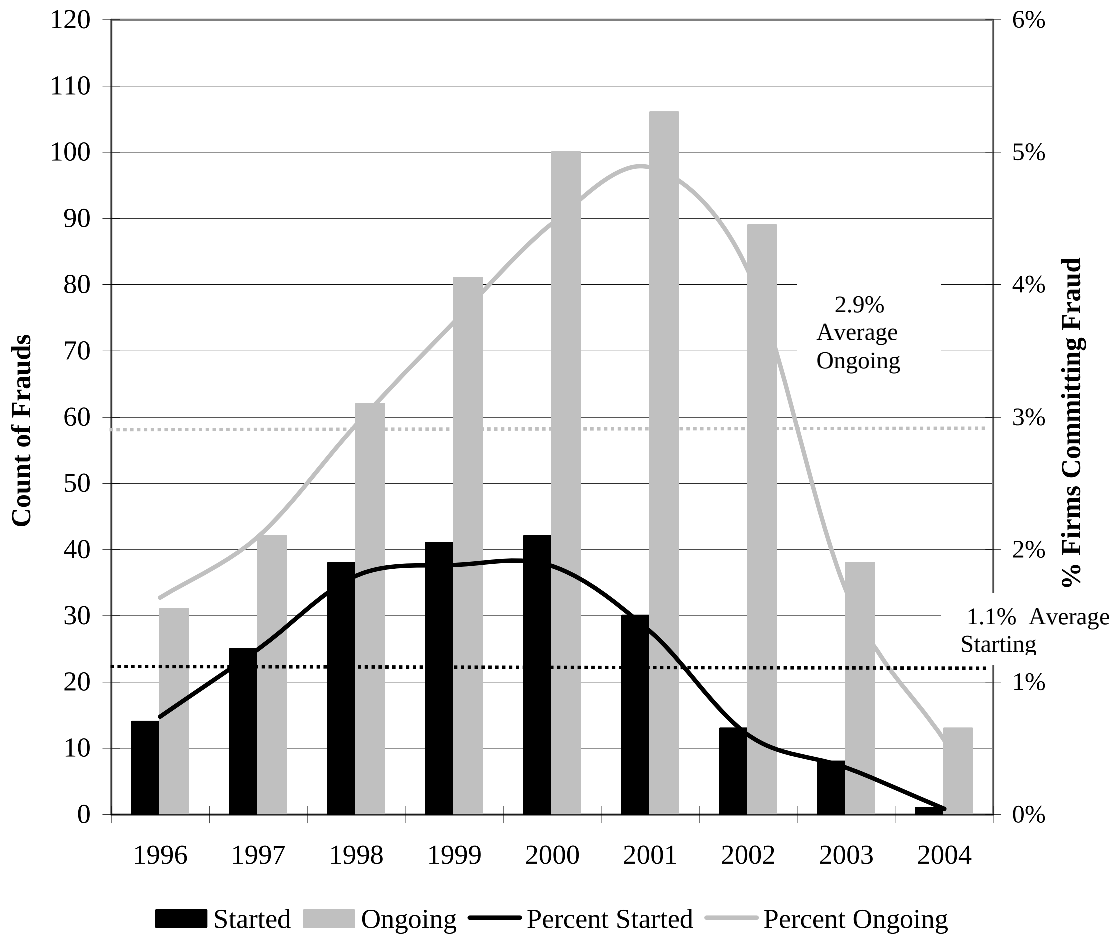

Figure 2: Pervasiveness of Starting and Ongoing Frauds – Adjusted for Truncation

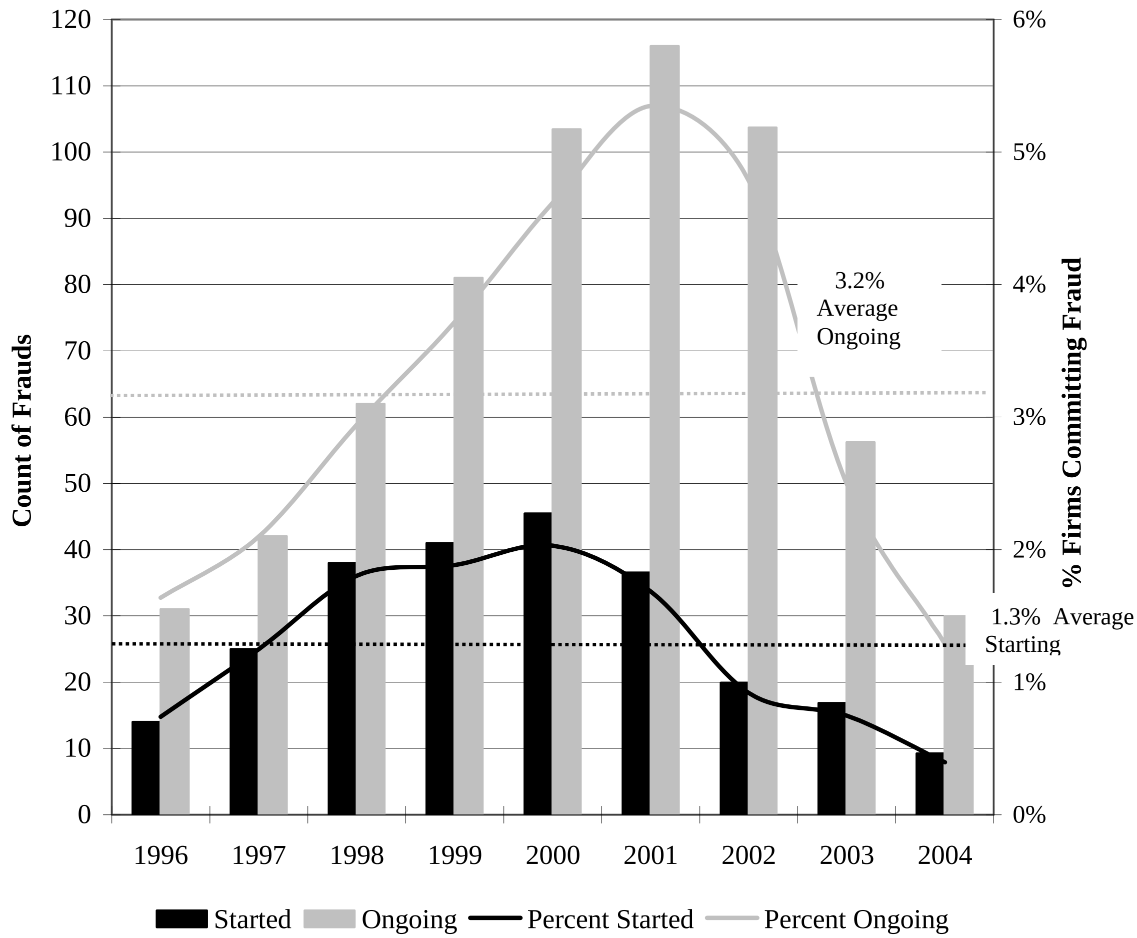

Table 1 – Data Definition and Sources

This table identifies the main variables used in our analysis, defines the variables, and provides the sources.

Variable	*Description*	*Sources*
Detector of Fraud	The actor who first identifies the fraud based on a combined reading of the legal case documents and an average of 800 articles from Factiva in a window from 3 months before the class action period to settlement. Ten detector categories include: external auditor, financial analyst, investor, shortseller, media, strategic players, financial market regulators, non-financial market regulators, employees and professional service organizations. The detection is attributed to the media only when the story does not indicate another actor as the principal source of information. Strategic players include suppliers, clients and competitors. Financial market regulators are the SEC and stock exchanges. Non-financial market regulators include industry regulators (e.g. FERC, FAA, FDA) and other government agencies. Professional service firms are law and insurance firms.	Security Class actions filings available from Stanford Securities Class Action Database, Articles in Factiva.
Fraud Duration	The class period defined in the security class action, reflecting all adjustments made before settlement. .	Stanford Securities Class Action Database
Financial Restatement Dummy	Observation has value 1 if: the firm filed a 10-Q/A or 10-K/A filing or an 8-K which referred to restatement information [116 cases]; it announced an intention to restate its financials but did not as a result of bankruptcy (e.g. Enron) [7 cases]; it took a one-time accounting-related charge [6 cases]; and, it is an ongoing case where there are accounting-related investigations [3 cases].	SEC filings, General Accounting Office (GAO) report on Financial Statement Restatements.
Analyst coverage indicator variable	A dummy variable that takes the value of 1 if the firm has higher than the median value of analysts in the Compustat sample of companies with more than $750 million in assets.	I/B/E/ S
Media coverage indicator variable	A dummy variable that takes the value of 1 if the firm has higher than the median value of median coverage in the Compustat sample of companies with more than $750 million in assets. We manually collect media coverage by searching the Wall Street Journal print edition (via Factiva) and recording the number of media hits for the year 1995.	Factiva
Qui-Tam Industry	A dummy variable that takes the value of 1 if the industry is one in which qui tam lawsuits are possible. We identify these industries based on the 3 digit SIC codes corresponding to the health care and defense contractor industries which account for the bulk of all qui tam cases. Specifically, these industries include	Civil Division, Department of Justice, Lexis.
In the money exercisable options	The sum of the in-the-money exercisable options for all executives.	Execucomp
P/E ratio	Price to operating earnings before depreciation, winsorized	Compustat, Crisp
Settlements	The sum of settlement dollars including insurance payouts prior year	Stanford Cass actions, Factiva
Regulated Firm Dummy	Firm in following categories: financials (SIC 6000-6999), transportation equipment (SIC 3700-3799), transportation, communications, electric, gas and sanitary services (SIC 4000-4999), drug, drug proprietaries and druggists sundries (SIC 5122), petroleum and petroleum products wholesalers (SIC 5172), pharmaceuticals (SIC 2830-2836), and healthcare providers (8000-8099), and healthcare related firms in Business Services.	Industries identified in Winston (1998) and others.
Shortability dummy	A dummy that takes the value 1 for those companies with a greater than median level of institutional shareholding in the prior year, our proxy for the ease of shorting the stock.	Compact – D

Table 2: Pervasiveness of Fraud 2002-2004 for Former Arthur Andersen Clients

Panel A

	Percentage of Large Firms Committing Fraud 2002-2004		
Increased Incentives for Detector			**p-value for diff.**
	Not Arthur Andersen 2001 (1943 total firms)	Arthur Andersen 2001 (398 total firms)	
Auditor	1.05%	1.85%	0.0225

Panel B

Detector	**Percentage of Frauds Detected by (row) in setting (column) during 2002-2004:**			**p-value for diff.**
	All Firms (83 total cases)	Not Arthur Andersen firm 2001	Arthur Andersen firm 2001	
Auditor	13.3%	9.7%	23.8%	0.0285

Table 3: 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 4: Estimation -- Probability of Being Caught and Detector Incentives

	Main Model:	*Logit(Caught)=f(Detector Incentives, Predicted Started$_{t-i}$)*		
	1	2	3	4
	Observed	*Single Logit*	*Two-Stage Logit*	*Two-Stage Logit*
Hi Analyst Coverage		0.636*	0.610*	0.604
		(0.329)	(0.33)	(0.38)
Hi Media Coverage		0.621***	0.615***	0.593**
		(0.225)	(0.23)	(0.29)
Hi Institutional Own		-0.383	-0.378	-0.378
		(0.311)	(0.31)	(0.38)
Best Fortune100		0.154	0.094	0.125
		(0.319)	(0.32)	(0.39)
End Post Sox		0.572***	0.647***	0.669**
		(0.158)	(0.18)	(0.30)
Regulated		0.302*	0.211	0.208*
		(0.164)	(0.17)	(0.22)
Qui Tam		0.871***	1.101***	0.991**
		(0.185)	(0.20)	(0.34)
Predicted Start$_{t-1}$			5.521	5.873*
			(6.03)	(3.84)
Constant		-5.279***	-5.325***	-5.311***
		(0.357)	(0.37)	(0.44)
Pseudo R-Square		0.030	0.032	0.032
Observations	10,444	10,444	10,444	10,444
Estimate	**1.65%**	**7.54%**	**7.28%**	**6.87%**
	First Stage:	*Logit(Start)$_{t-1}$ = f(E[Penalty], E[Payoff], Detector Incentives)*		
In Money Options			1.109***	1.142***
			(0.21)	(0.21)
P/E Ratio			-1.366**	-1.378**
			(0.60)	(0.62)
Post SOX			-0.812***	-0.827***
			(0.27)	(0.27)
Settlements$_{t-1}$			-0.103**	-0.102**
			(0.05)	(0.05)
Hi Analyst Coverage				0.068
				(0.36)
Hi Media Coverage				0.238
				(0.24)
Hi Institutional Own				-0.007
				(0.35)
Best Fortune100				-0.317
				(0.40)
Regulated				0.074
				(0.18)
Qui Tam				0.850***
				(0.21)
Constant			-3.793***	-4.171***
			(0.12)	(0.37)
Pseudo R-Square			0.039	0.503
Observations			10,444	10,444

Table 5: Cost of Median Fraud

Multiple Factor	*25th Percentile Firm Loss (in $ millions)*	***Median Firm Loss (in $ millions)***	*75th Percentile Firm Loss (in $ millions)*	*# Cases*
Sales	27.9	**- 1,569.5 (38.8% of enterprise value)**	- 8,829.1	186
Fixed Assets	- 90.5	**- 1,575.3 (42.6% of enterprise value)**	- 10,969.4	178
EBITDA	1,623.5	**- 236.4 (8.5% of enterprise value)**	- 4,185.8	114
Multiples Average	520.3	**- 1,127.1 (30.0% of enterprise value)**	- 7,994.8	

Data Appendix

Comparing Our Sample with Other Fraud Samples

Many accounting studies focus on a sample of companies identified by the GAO that restated their financial statements between 1997 and June 2002 (e.g. Palmrose and Scholz (2004)). This 'GAO sample' includes all type of restatements (i.e. major and minor, revenue increasing and decreasing, and as a result of new GAAP, reclassification of accounts, merger/acquisition, restructuring charges or fraud).

Our sample differs in two principle ways. First, many of these cases will not make it into our sample. This arises because the GAO sample includes: some non-US firms; the GAO sample includes many smaller firms that do not meet the selection criteria for our sample (the median market cap in the GAO sample (measured at date t-1) is $ 214 million while the market cap of firms in our sample (also measured at t-1) is $ 3525 million); and, because the underlying fraud is not sufficiently serious to trigger a lawsuit that withstands scrutiny and yields a settlement or is ongoing . Second, this approach does not allow for cases of fraud where firms do not issue restatements, a category of frauds that accounts for 38 percent of our observations.

Other accounting studies have focused on a narrower sample of firms where the SEC has sanctioned the firm and released an Accounting, Auditing and Enforcement Release (AAER) (e.g. Dechow, Sloan and Sweeney (1996), Miller (2006)). We will capture these cases if there is a simultaneous suit under federal securities laws that meets our tests for inclusion. The SEC sample also is focused on smaller firms (the median market cap (measured at t-1) for AAER firms is 262 million) and, given its limited budget, on a few high profile and egregious cases of fraud.[12] Our companion paper provides a more complete comparison of these samples and the relationship of our sample to these.

The larger size of firms in our sample likely corresponds with additional scrutiny both before the fraud was brought to light and evaluation of the fraud and how it got uncovered after the fact. This additional scrutiny aids us in identifying the likely source of the information about fraud and in identifying some of the interactions among fraud detectors, including identifying actors who pushed the board to action. These factors help to account for the higher percentage of cases in our sample where indications of fraud arise from actors outside the firm. In our sample, we identify the firm as the source of information in 32% of cases whereas the firm is identified as the source in between 49% and 58% of cases in the GAO sample (1997-2002, and 2002-2005 respectively), and in 71% of cases in the AAER sample used by Miller (2006).[13]

Legal scholars have been the biggest user of the SSCAC database to construct samples of probable frauds (see citations above). A potential concern with this sample is that it is potentially missing additional cases of alleged fraud that are filed as a class action under state laws or as a derivative action. Thompson and Sale (2003) and Thompson and Thomas (2003, 2004) provide analysis and evidence that exploring such suits would not turn up many additional cases as there has been a profound shift in cases from state to federal courts, accentuated by the

---

[12] Dechow, Sloan and Sweeney (1996) write: "because our sample is subject to SEC enforcement actions, it is almost certainly biased toward the inclusion of the more obvious and spectacular cases of earnings manipulation."
[13] Correspondence with Shiva Rajgopal, January 2007.

passage of PSLRA and the Uniform Standards Act (1998). Their comprehensive analysis of these filings in Delaware in 1999 and 2000 shows that almost all such cases that withstand scrutiny are breach of fiduciary duties in merger and acquisitions (and thus not fraud in the general use of this term in that they do not involve misrepresentations).

Finally, others (E.g. Romano (1991)) have constructed their estimate of frauds by taking a random sample of publicly traded companies and then examining all litigation associated with these companies.

Identifying Frauds that Require Restatements

We distinguish between frauds that required financial restatements and frauds that do not. To identify whether the fraud involved restatements we used information from the United States General Accounting Office (GAO) report on Financial Statement Restatements that identifies 918 restatement announcements from 1997 to June 2002, which we matched to those in our sample. We also searched a firm's SEC filings after the revelation of fraud for either (a) a 10-Q/A or 10-K/A filing which indicate amended filings; or (b) an 8-K which referred to restatement information. We identified a fraud as involving misrepresentation if any of the following conditions applied: it restated its financials [116 cases]; it announced an intention to restate its financials but did not as a result of bankruptcy (e.g. Enron) [7 cases]; it took a one-time accounting-related charge [6 cases]; and, it is an ongoing case where there are accounting-related investigations [3 cases].

The residual category of frauds that don't require financial misrepresentation, are primarily composed of "failure to disclose" material information, and a disclosure of misleading forward-looking information, with the case of CVS illustrating the first type and Ascend the second type. In the case of CVS, the alleged fraud was to issue positive statements concerning its business and operations and possibilities for expansion but not to disclose that a national shortage of pharmacists was negatively impacting CVS's business forcing a scale back in expansion plans. Or consider the case of Ascend Communications, where the company followed a competitor's announcement that it would ship a 56K modem, with a near immediate announcement that it too would ship a 56K modem and beat the competitor to market, even though there were strong indications, including the supplier that allegedly would produce the modem, that suggested this was not possible.

Appendix I. Complementary Pervasiveness Estimation: Survey Evidence from Chicago MBA Students

A potential concern with relying on the analysis above is that it is based on the same sample. In this section we present results based on an independent measure of the frequency of illegal behavior in corporate America we derived by conducting a survey with University of Chicago MBAs.

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 we inserted in this program an anonymous survey on illegal and unethical behavior students encountered in their previous jobs.

For the purpose of the survey we defined a "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." 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.

With these caveats in mind, let's look at the data. Table 5 Panel A reports the percentage of MBAs who responded they faced a legal dilemma, divided by the industry they worked for before they joined the MBA program. On average 15% of the students were asked to do something illegal in their previous employment. 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%, less then 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% vs. 37%) 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.

How does this survey-based evidence compare with the one emerging from the legal suits? Since the average student has two years of work experience, if we assume that the average duration of the fraud is one year, we have an incidence of fraud per year equal to 7.5%, which is remarkably in line with the evidence collected from the legal suit.

Appendix **Table 1: Pervasiveness of Detected Fraud in a Survey of MBAs**

For the purpose of the survey we defined a "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 they faced a legal dilemma by industry they worked for before they joined the MBA program. Panel B reports the percentage of MBAs who responded they faced a legal dilemma by function they performed before they joined the MBA program.

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
Grand 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
Grand Total	14.84%	283