Predicting Material Accounting Manipulations*

Patricia M. Dechow
The Haas School of Business
University of California, Berkeley
Berkeley, CA 94705
patricia_dechow@haas.berkeley.edu

Weili Ge
University of Washington Business School
University of Washington
Mackenzie Hall, Box 353200
Seattle, WA 98195
geweili@u.washington.edu

Chad R. Larson
The Stephen Ross School of Business
University of Michigan
Ann Arbor, MI 48109
larsonch@umich.edu

Richard G. Sloan
Barclays Global Investors
45 Fremont St
San Francisco, CA 94105

First version: January 8, 2007
This version: July 12, 2007

* We appreciate the comments of the workshop participants at University of Michigan, the UBCOW Conference at the University of Washington, New York University 2007 Summer Camp, and University of California, Irvine. We thank Sudhakar Balachandran, Ilia Dichev, Bjorn Jorgensen, Bill Kinney, Mort Pincus, and Charles Shi, for their comments. We would like to thank the Research Advisory Board established by Deloitte & Touche USA LLP, Ernst & Young LLP, KPMG LLP and PricewaterhouseCoopers LLP for the funding for this project. However, the views expressed in this article and its content are those of the authors alone and not those of Deloitte & Touche USA LLP, Ernst & Young LLP, KPMG LLP or PricewaterhouseCoopers LLP. Special thanks go to Roslyn Hooten for administering the funding relationship. This paper is dedicated to the memory of our colleague, friend and research team member Nader Hafzalla, who was a joy to all who knew him.

Executive Summary

The objective of this research project is to identify factors predicting accounting manipulations. We construct a comprehensive database of firms that are known to have manipulated their financial statements. Since 1982, the Securities and Exchange Commission (SEC) has issued *Accounting and Auditing Enforcement Releases* (AAERs) documenting enforcement actions against companies, auditors and officers for alleged accounting misconduct. These releases provide varying degrees of detail on the nature of the alleged misconduct. We examine 2,191 *Accounting and Auditing Enforcement Releases* that occurred between 1982 and 2005 and identify all reporting periods in which earnings are alleged to have been manipulated. Our final sample consists of 680 firms with alleged manipulations in their quarterly or annual financial statements.

Below we summarize our analysis and key findings:

1. General characteristics of manipulating firms.

- Most firms manipulated more than one income statement line item. Revenue is by far the most commonly affected item, with alleged manipulations in 55 percent of sample firms. Manipulations of inventory and cost of goods sold occurred in 25 percent of sample firms. Manipulations of allowances, including the allowance for doubtful debts, are also common, occurring in 10 percent of sample firms.

- The most common industries in which manipulations occurred are computers and computer services, retail, and general services (such as telecommunications and healthcare).

- Alleged manipulations are common in large firms. We find that 15.3 percent of the manipulations occur in the largest 10 percent of firms. This is likely due to the SEC's incentive to identify only the most material and visible manipulations involving large losses to numerous investors.

2. Our first set of tests compare the financial statements of manipulating firms in manipulating years versus other years.

- A consistent theme among manipulating firms is that they have shown strong performance prior to the manipulations. Manipulations appear to be motivated by managements' desire to disguise a moderating financial performance. We find that manipulating firms' stock returns outperform the broader market in the years leading up to the manipulation and begin to underperform in the years following the manipulation. In manipulation years, cash profit margins and earnings growth decline, while accruals increase. In addition, order backlog and employee headcount decline, indicating a drop in demand for the firms' products.

- We find abnormal increases in leasing activity during manipulation periods. This result is consistent with managements' increased use of the flexibility granted by lease accounting rules to manipulate their firms' financial statements.

- Investors have abnormally high expectations of future profitability for manipulating firms during manipulation periods. We find manipulating firms have abnormally high price-to-earnings and market-to-book ratios during manipulation years. In addition, issuances of debt and equity are both unusually high during manipulation years. These results suggest that manipulations are intended to avoid disappointing investors' high expectations and to raise capital on favorable terms while expectations are still high.

- Manipulating firms tend to have abnormally low free cash flows. Many of these firms are actively seeking new financing to cover negative operating and investing cash flows. The manipulations help to hide their deteriorating financial performance thus enabling them to obtain financing on more favorable terms.

3. **Our second set of tests compares the financial statements of manipulating firms during manipulation periods to all publicly listed firms. The results are similar to those reported above.**

   - Manipulating firms have abnormally high accruals, deteriorating cash margins and deteriorating earnings growth. They have declining abnormal order backlog, declining abnormal employee headcount and abnormally high financing activities.

4. **Our final set of tests develops a prediction model to assess the probability of manipulation. The model produces a *Fraud-Score* (*F-Score*). An *F-Score* greater than 1.00 indicates a higher probability of manipulation.**

   - The model is built in stages. **Model 1** includes variables that are obtained from the primary financial statements. These variables measure earnings quality and firm performance. **Model 2** adds off-balance sheet and non-financial measures such as leasing activity and abnormal changes in employee headcount. **Model 3** adds market-related variables such as prior stock price performance and the book-to-market ratio. We find that the bulk of the predictive power of the models is obtained from **Model 1** using financial statement variables. **Models 2** and **3** provide modest incremental improvements.

   - We rank all firm-years on Compustat over our sample period (over 100,000 firm-year observations) and calculate *F-Scores* for each firm-year. We find that about half of the manipulation firms have *F-Scores* in the highest quintile of all *F-Scores*. In other words, approximately fifty percent of manipulation firms have *F-Scores* in the top 20 percent of all firm-years. Therefore, our models are relatively powerful at correctly identifying manipulations.

   - There is a trade off between *Type I* and *Type II* errors when using our prediction models. *Type I* errors occur when the model predicts manipulation when no manipulation occurs. *Type II* errors occur when the model predicts no manipulation when a manipulation does occur. *Type II* errors are particularly costly to auditors, because they involve a high probability of lawsuits (though Type 1 errors may involve costs in the form of lost clients). We illustrate how the *F-Score* cut-off will depend on individual users' assessments of the costs of *Type I* and *Type II* errors.

5. Implications

- Our analysis provides guidelines for auditors, financial analysts, corporate executives and investors who are interested in understanding the determinants of manipulations.

- Our *F-Score* model can be incorporated into the audit process as a "first-pass" test in evaluating the likelihood of manipulation for client firms. A high *F-Score* does not guarantee a manipulation, but it does serve as a red flag, signaling the need for further analysis. A high *F-Score* can be explained to clients and used to justify more thorough audit investigations and higher audit fees. In addition, accounting firms can assess individual audit partners' relative *F-Score* risks to determine whether particular partners are signing off on more risky clients. As an example application, we provide the *F-Score* of manipulating firms for each of the major audit firms.

- Our *F-Score* model can be computed by financial analysts and investors to provide a preliminary assessment of 'earnings quality.' A high F-Score indicates that additional analysis is warranted before relying on the reported financial statements.

- Our *F-Score* model can be used by quantitative investment managers as a stock selection signal and risk descriptor.

Abstract

We provide a comprehensive analysis of accounting manipulations disclosed between 1982 and 2005. We create our database through a detailed examination of 2,191 SEC Accounting and Auditing Enforcement Releases (AAERs). Our database contains 680 firms that are alleged to have manipulated their quarterly or annual financial statements. We examine the characteristics of manipulating firms and analyze the ability of (i) financial statement variables; (ii) off-balance sheet and non-financial variables; and (iii) market-related variables, to explain and predict manipulations. The financial statement variables that we find useful include measures of accrual quality and firm performance. The useful off-balance sheet and non-financial variables include the existence and use of operating leases, abnormal changes in employees and order backlog. The market-related variables that are useful include book-to-market, earnings-to-price, prior annual stock price performance, and amount of new financing. Our results suggest that manipulations are most common in growth companies experiencing deteriorating operating performance. We compare manipulating firms to the broader population of public firms and develop a model to predict accounting manipulations. The output of this model is a scaled logistic probability that we term a Fraud-Score (*F-Score*), where values greater than 1.00 indicate higher likelihood of manipulation. We show that over 60 percent of manipulating firms have *F-Scores* greater than 1.00 and that the selection of an *F-Score* cut-off is based on the relative costs of *Type I* versus *Type II* errors. As an example application of the *F-Score*, we provide the median *F-Score* of manipulating clients for each of the major audit firms.

1. INTRODUCTION

What causes managers to manipulate their financial statements? How best can investors, auditors, financial analysts and regulators detect manipulations? Addressing these questions is of critical importance to the efficient functioning of capital markets. For an investor it can lead to improved returns, for an auditor it can mean avoiding costly litigation, for an analyst it can mean avoiding a damaged reputation, and for a regulator it can lead to enhanced investor protection and fewer investment debacles. The objective of this research project is two-fold. First, we develop a comprehensive database of financial manipulations. Our objective is to describe this database and make it broadly available to other researchers to promote research on earnings manipulations. Second, we provide the initial groundwork in analyzing the characteristics of manipulating firms and the determinants of manipulations. Based on this analysis, we develop a model to predict manipulations and provide an associated Fraud-Score (*F-Score*) that can be used to assess the likelihood of manipulations.

We compile our database through a detailed examination of firms that have been subject to enforcement actions by the Securities and Exchange Commission (SEC) for allegedly manipulating their financial statements. Since 1982, the SEC has issued *Accounting and Auditing Enforcement Releases* (AAERs) during or at the conclusion of an investigation against a company, an auditor, or an officer for alleged accounting and/or auditing misconduct. These releases provide varying degrees of detail on the nature of the misconduct, the individuals and entities involved and the effect on the financial statements. We examine the 2,191 *Accounting and Auditing Enforcement Releases* released between 1982 and 2005. Our examination identifies 680 unique firms that have misstated at least one of their quarterly or annual financial statements. Our database includes firm identifiers, reporting periods affected by manipulations, and details on the nature of the manipulations.

Our main findings are summarized as follows. We find that most firms manipulate more than one account. Revenue, which is overstated in 55 percent of the sample firms, is by far the most commonly misstated account. Manipulations of reserves, including the allowance for doubtful debts is also common, occurring in 10 percent of the sample. Manipulations of inventory and cost of goods sold occurred in 25 percent of the sample. Manipulations are also clustered in certain industries, most commonly, computers and computer services, retail, and general services (such as telecommunications and healthcare). They are also common in large firms with 15.3 percent of the manipulations occurring in the largest 10 percent of firms. This is likely due to the SEC's incentive to identify only the most material and visible manipulations involving large losses to numerous investors.

We examine the characteristics of firms in our database along five dimensions. These are (i) accrual quality; (ii) financial performance; (iii) non-financial measures; (iv) off-balance sheet activities; and (v) market-based measures. We provide time-series analysis of these variables for manipulating firms and cross-sectional analysis comparing manipulating firms to the broader population of firms.

We provide several approaches to measuring accrual quality. We use composite measures of working capital accruals (as reported in Sloan (1996)) and broader measures of accruals that incorporate long-term operating assets and liabilities (as reported in Richardson, Sloan, Soliman and Tuna (2005)). We also examine various models of discretionary accruals developed in prior accounting research. We measure discretionary accruals using the cross-sectional modified Jones model (see Dechow, Sloan, and Sweeney (1995) and DeFond and Jiambalvo (1994)) and the performance matched discretionary accruals model promoted by Kothari, Leone, and Wasley (2005)). In addition, we examine measures of earnings quality developed in Dechow and Dichev (2002). Finally, we provide an analysis of two specific accruals: change in receivables and inventory.

We find that all measures of accruals are unusually high during manipulation periods. The broad measure of total accruals developed by Richardson et al. (2005) has the highest statistical association with manipulations. We also find that including periods after the manipulations in these tests provides additional explanatory power with respect to predicting manipulations. This result arises because the subsequent reversal of overstated accruals makes the overstated accruals more obvious. Note that while subsequent accruals are obviously not available for those interested in predicting manipulations, they are available to researchers and regulators who seek to determine whether a manipulation existed after the fact.

We next examine various measures of performance using information reported in the financial statements. We find that accounting rates of return are generally declining at the time that firms misstate their earnings. Additionally, earnings growth rates are negative and cash margins are declining. Contrary to our initial expectations, we find that cash sales are increasing during manipulation periods. We failed to anticipate this result because we expected firms to boost sales through the manipulation of credit sales. There are, however, two explanations for this result. First, manipulating firms tend to be growing their capital bases and increasing the scale of their business operations. The greater scale of operations should lead to increases in both cash and credit sales. Second, an inspection of the AAERs reveals that many firms manipulate sales through transaction management - for example, by encouraging sales to customers with return provisions that violate the definition of a "sale," selling goods to related parties, or forcing goods onto customers at the end of the quarter. All of these manipulation techniques can boost *cash* sales and so accrual-based measures of earnings quality are unlikely to detect such manipulations. A useful area for future research would be to develop measures of earnings quality that capture cash-based earnings manipulation.[1]

We find that two non-financial measures are useful in detecting manipulations. The first is a decline in order backlog. A decline in order backlog suggests a weakening demand for the firm's product and deteriorating operating performance (Lev and Thiagarajan 1993). This decline could lead managers to overstate earnings in order to hide deteriorating performance from investors. The second non-financial measure is new to the literature and is abnormal reductions in the number of employees. Reductions in the number of employees are also likely to occur when there is declining demand for the firm's product. In addition, cutting employees directly improves short-run earnings performance (lowers wage expense).

Our examination of off-balance sheet information focuses on the existence and use of operating leases and the expected return assumption on plan assets for defined benefit pension plans. Operating leases can be used to front-load earnings and reduce reported debt. Therefore, operating leases can be used as 'legal' earnings management and balance sheet management tools. We find that the use of operating leases is unusually high during manipulation years. We also find that manipulating firms have higher expected returns on their pension plan assets than other firms. Higher expected return assumptions reduce reported pension expense. The results for leases and pensions suggest that manipulating firms might exhaust 'legal' earnings management options before resorting to potentially illegal financial manipulations.

[1] Roychowdhury (2006) analyzes the management of cash from operations and production expenses and provides a preliminary step in this direction.

Our final set of variables relate to stock and debt market incentives. Dechow, Sloan, and Sweeney (1995) suggest that market incentives are an important reason for engaging in earnings management. Teoh, Welch, and Wong (1998) and Rangan (1998) provide corroborating evidence that accruals appear to be unusually high during equity issuances. We find that manipulating firms tend to be running short of cash. In addition, we find that manipulating firms are actively seeking additional financing from capital markets. These findings suggest that manipulating firms are attempting to inflate their stock prices in order to raise capital on more favorable terms.

We examine the growth expectations embedded in manipulating firms' stock market valuations. We find that price-earnings ratios and market-to-book ratios are unusually high prior to manipulations, suggesting that investors are optimistic about the future growth opportunities of these firms. We also find that the manipulating firms have had unusually strong stock price performance in the years prior to manipulation. Thus managers may engage in manipulations because they want to avoid disappointing investors and losing their high stock prices. They may do this because they own stock options, or because they plan to raise new financing. Either way, strong prior operating performance is likely to create incentives for managers to continue to report strong results to the market, even if it means manipulating earnings. A consistent theme among manipulating firms appears to be that they have shown strong performance prior to the manipulations and that the manipulations are made to hide deteriorating performance.

Our final tests develop a prediction model for manipulations. The output of this model is a scaled logistic probability that we term a Fraud-Score (*F-Score*). Manipulations identified by the SEC tend to be the most egregious and high profile types of manipulations. This is likely to be a result of limited SEC resources devoted to fraud prevention. Therefore, the percentage of manipulation firm-years is low relative to the total available number of firm-years in the population (less than one percent of firm-years). The model is built in stages. **Model 1** includes variables that are obtained from the primary financial statements. These variables include accrual quality and firm performance. **Model 2** adds off-balance sheet and non-financial measures such as operating leases and abnormal changes in employees. **Model 3** adds market-related variables such as prior stock price performance and the book-to-market ratio. We find that the bulk of the power of the prediction model is obtained using our simple **Model 1**. **Models 2** and **3** provide modest incremental improvements.

We calculate an *F-Score* for both manipulating and non-manipulating firms and rank all available firm-years by *F-Score*. We find that about half of the manipulating firms have *F-Scores* in the highest quintile of all *F-Scores*. In other words, approximately fifty percent of manipulation firms have *F-Scores* in the top 20 percent of all firm-years. In addition, *F-Scores* greater than 1.00 indicate probabilities of manipulation higher than the unconditional expectation. We find that over 60 percent of manipulating firms have *F-Scores* greater than 1.00.

We point out that users must trade off *Type I* and *Type II* errors when evaluating the power of our models. *Type I* errors occur when the model predicts a manipulation when no manipulation exists. *Type II* errors occur when the model predicts no manipulation when the firm is actually manipulating its results. *Type II* errors are likely to be more costly to auditors and regulators since investors are likely to sue auditors and criticize regulators who fail to detect or prevent manipulations. We quantify how the relative costs of *Type I* versus *Type II* errors determine the choice of an *F-Score* cut-off. For example, we show that if not detecting a misstating firm is 158 times more costly than accusing a non-manipulating firm of manipulation, then an *F-Score* cut-off of 1.00 should be selected.

We provide a simple example of how the *F-Score* can be used as a tool for raising questions for further analysis. We sort manipulating firms into groups based on the auditor signing off on the misstated financial statements. We determine the maximum *F-Score* during manipulating years for each client (since the maximum is the strongest signal of a potential problem). We then determine the median *F-Score* across manipulating clients for each audit firm. We find that there is considerable dispersion in *F-Scores* across audit firms and discuss potential explanations that could be examined in more detail in future research.

The remainder of the paper is organized as follows. Section 2 reviews previous research on this topic. Section 3 describes database construction and research design. Section 4 presents our analysis of manipulation firms and develops our manipulation-prediction model. We also provide our simple analysis of *F-Score* by audit firm. Section 5 concludes.

2. PREVIOUS LITERATURE

Several prior research studies have used Accounting and Auditing Enforcement Releases (AAERs) and other datasets to identify characteristics of manipulating firms. We briefly discuss some of the key findings.

Dechow, Sloan and Sweeney (1996) analyze 436 AAERS released between April 1982 and December 1992. Their final sample after eliminations consists of 92 firms. Each firm is matched in the year prior to manipulation to a control firm in the same three-digit SIC industry and with similar asset values. The authors provide some evidence that accruals appear to be high at the time of manipulation. However, the paper focuses primarily on showing that various corporate governance factors appear to be correlated with manipulation. For example, they find that manipulating firms have a higher number of insiders on the board and a CEO who is more powerful and entrenched. They provide matched-pairs logit analysis; however, they do not report how effective their model is at fraud prediction.

Beneish (1997) analyzes 49 AAERs and 15 firms whose accounting was questioned by the news media between 1987 and 1993. He creates a separate sample of firms he terms "aggressive accruers" using the modified Jones model to select firms with high accruals. His objective is to distinguish the GAAP violators from firms that have high accruals and appear to be applying GAAP aggressively. Beneish (1997) finds that accruals, day's sales in receivables and prior performance are important for explaining the differences between the two groups.

In concurrent research, Ettredge, Sun, Lee, and Anandarajan (2006) examine 169 AAER firms matched by firm size, industry and whether the firm reported a loss. They find that deferred taxes can be useful for predicting fraud, along with auditor change, market-to-book, revenue growth and whether the firm is an OTC firm. Brazel, Jones, and Zimbelman (2006) examine whether several non-financial measures (e.g., patents and trademarks) can be used to predict fraud in 77 AAER firms. They find that growth rates between financial and non-financial variables are significantly different for AAER firms. Skousen and Wright (2006) analyze 86 manipulation firms matched by industry and sales. Similar to Dechow et al. (1996), they focus on governance variables. They find that manipulation firms tend to have managers with higher stockholdings (greater than five percent), have less effective audit committees, have more powerful CEOs, and are more likely to have recently switched auditors.

Richardson, Tuna, and Wu (2002) examine 255 firms that restate earnings between 1971 and 2000 and compare them to 133,208 non-restating firms. They obtain their sample through a Nexis-Lexis search using variations on the word "restate." They exclude restatements due to changes in FASB accounting rules, stock splits, merger and acquisitions, etc. They tests for differences in means for restating firm-years relative to non-restating firm-years and find that restating firms have lower earnings-to-price and book-to-market ratios, raise more financing, and have larger total accruals. They also find that restating firms have longer consecutive strings of growth in quarterly EPS. Similar to Dechow et al (1996) they suggest that capital market pressures are likely to be a motivating factor for the earnings management that results in restatements. Richardson et al do not report the number of restating firms that end up with SEC Enforcement actions. However there is likely to be some overlap between their sample and our sample. Note also, Richardson et al, do not provide a logistic analysis to assess the relative importance of the variables they examine.

This paper extends the literature on accounting manipulations by making three significant contributions. First, prior literature examining accounting manipulations has relied on either small samples of accounting manipulations obtained from a limited number of SEC AAERs or larger samples of earnings restatements obtained from the Government Accounting Office or keyword news searches. For this paper we rigorously collect, categorize and code detailed information on all 2,191 SEC AAERs from 1982 through June 2005. This new database enables us to provide a comprehensive analysis of over 600 manipulating firms, a sample size far larger than those used in prior research. Availability of this database will encourage research on earnings manipulation and increase knowledge of the determinants and consequences of manipulation.

Second, we systematically examine a comprehensive set of prediction variables that relate to accrual quality, performance, and market-related incentives and establish which variables are relatively more important. In addition, we analyze whether off-balance sheet metrics provide useful information over and above measures reported in the financial statements. Although previous literature has analyzed several of the variables we examine in a univariate framework, we extend this research by introducing new variables and confirming their importance in multivariate framework using a larger, more comprehensive sample.

Finally, we develop a parsimonious prediction model and an associated *F-score* that is readily amenable to practical implementation and future research. Although previous literature has identified variables correlated with accounting manipulations, we are not aware of previous work that has detailed the effectiveness and cost and benefit tradeoffs of using fraud prediction models in a large population of firms. By testing our model in a large population of firms, we are able to provide detailed evidence on the number of *Type I* and *II* errors users of our fraud prediction models will likely encounter. Our analysis of *Type I* and *II* errors provides subsequent researchers with a framework for analyzing the costs and benefits of implementing more extensive models with additional variables.[2]

3. DATA AND SAMPLE FORMATION

3.1 Data Sample

The objective of our data collection efforts is to construct a comprehensive sample of material and economically significant accounting manipulations involving both GAAP violations and the allegation that the manipulation was made with the intent of misleading investors. The SEC's series of published Accounting and Auditing Enforcement Releases provides the ideal starting point for our sample construction. The SEC takes enforcement actions against firms, managers, auditors and other parties involved in violations of SEC and federal rules. During or at the completion of a significant investigation involving accounting and auditing issues, the SEC issues an Accounting and Auditing Enforcement Release (AAER). The SEC reviews about one-

[2] For example, corporate governance variables could be added to our models to determine their incremental explanatory power. We do not investigate corporate governance because a comprehensive and detailed analysis would add considerably to the length and complexity of our paper. Other extensions include applying our prediction models by industry and assessing the relative importance of variables.

third of public companies' financial statements each year and checks for compliance with GAAP. If SEC officials believe that reported numbers are inconsistent with GAAP, then the SEC can initiate informal inquiries and solicit additional information. If the SEC is satisfied after such informal inquiries, then it will drop the case. However, if the SEC believes that one or more parties violated securities laws, then the SEC can take further steps, including enforcement actions requiring the firm to change its accounting methods, restate financial statements, and pay damages.

Dechow, Sloan and Sweeney (1996) identify how and when the news of the manipulation first becomes public to the market. Initial sources for manipulations include the SEC in its routine oversight of registrants, anonymous tips from employees and other insiders, journalists and analysts. Another source is the voluntary restatement of the financial results by the firm itself. Restatements are a red flag that often lead to SEC investigations and subsequent enforcement actions.

There are a number of conceivable alternative sources for identifying accounting manipulations. They are discussed briefly below, along with our reasons for not pursuing these alternatives.

1. The GAO Financial Statement Restatement Database. This database consists of approximately 2309 restatements between January 1997 and September 2005. This database was constructed through a Lexis-Nexis text search of press releases and other media coverage based on variations of the word 'restate.' There is some overlap between the AAER firms and the GAO restatement firms since a) the SEC often requires firms to restate their financials as part of a settlement; and b) restatements often trigger SEC investigations. The GAO database covers a relatively small time period, but consists of a relatively large number of restatements. The reason for the

large number of restatements is that the GAO database includes all restatements relating to accounting irregularities regardless of managerial intent, materiality and economic significance. Consequently, it includes a large number of economically insignificant restatements. The objective of our study is to investigate manipulations involving GAAP violations that are both economically significant and made with the intent of misleading investors. Another shortcoming of the GAO database is that it only identifies the year in which the restatement was identified in the press and not the reporting periods that were required to be restated.[3] Our research design requires the specification of the reporting periods that are manipulated so that we can examine the characteristics of these reporting periods.

2. Stanford Law Database on Shareholder Lawsuits. Shareholder lawsuits typically result from material intentional manipulations. However, shareholder lawsuits can also arise for a number of other reasons that are unrelated to financial manipulations. Shareholder lawsuits alleging manipulations are also very common after a stock has experienced a precipitous stock price decline, even when there is no clear evidence supporting the allegation. In contrast, the SEC only issues an enforcement action when it has established intent or gross negligence on the part of management in making the manipulation.

3.2 Datasets

We catalog all the AAERs from AAER 1 through AAER 2261 spanning May 17th, 1982 through June 10th, 2005. We next identify all firms that are alleged to have violated GAAP by at

[3] For example, while Xerox is included in the GAO database in 2002, the restatements in question relate to Xerox's financial statements for 1997, 1998, 1999, 2000 and 2001.

least one of these AAERs (we describe this procedure in more detail in the next section). We then create three data files: the *Detail, Annual* and *Quarterly files*.

The *Detail* file contains all AAER numbers pertaining to each firm, firm identifiers, a description of the reason the AAER was issued, and indicator variables categorizing which balance sheet and income statement accounts were identified in the AAER as being affected by the violation. There is only one observation per firm in the *Detail* file.

The *Annual* and *Quarterly* files are compiled from the *Detail* file and are formatted by reporting period so that each quarter or year affected by the violation is a separate observation. This is suitable for researchers as each quarter or year-end observation can easily be matched to the corresponding data in financial databases such as CRSP, Compustat and I/B/E/S.

The *Annual* file contains the company name, Cusip (cnum in Compustat), Compustat fiscal year, Compustat fiscal month end, calendar date (obtained from the AAER), and the primary AAER used to collect the information. The *Quarterly* file contains the company name, 6-digit Cusip (cnum in Compustat), Compustat fiscal year, Compustat fiscal month end, quarter (1, 2, 3 or 4), calendar date (obtained from the AAER), and the primary AAER used to collect the information.

All three files are in .sas7bdat format and can therefore be easily read into SAS or converted into any other format for other statistical programs. Appendix 1 lists the variable names and description for each file in the database.

[Appendix 1]

3.3 Data Collection

The original AAERs are the starting point for collecting data. Copies of the AAERs are obtained from the SEC website and the Lexis Nexis database. Each AAER is separately examined to identify whether it involves an alleged GAAP violation. In cases where a GAAP violation is involved, the reporting periods that were alleged to be manipulated are identified. Identifiers (e.g., Cusip, Permno) for these firms are then retrieved using Wharton Research & Data Services (WRDS).

The data coding was completed in three phases. In the first phase, all releases were read in order to obtain the company name and period(s) in which the violation took place. The AAERs are simply listed chronologically based on the progress of SEC investigations. To facilitate our empirical analysis, we record manipulations by firm and link them back to their underlying AAERs in the detail file. Note that multiple AAERs may pertain to a single set of restatements at a single firm. Panel A of Table 1 indicates that of the 2,261 AAERs, we are unable to locate 30 AAERs either because they were missing or not released by the SEC. A further 40 AAERs relate to auditors or other parties and do not mention specific company names. This leaves us with 2,191 AAERs mentioning a company name.

Figure 1 reports that in the 2,191 AAERs, the SEC takes action against 2,592 different parties. In 66 percent (1,447) of the cases, the party was an officer of the company, in 30 percent (663) of cases the party was the firm itself, in a further 17 percent (383) of cases the party was an auditor, in 1 percent (14) the party was an attorney, and in 4 percent (85) cases the party was classified as “other,” which includes consultants and investment bankers.

[Figure 1]

Table 1 Panel B provides the distribution of the 2,191 AAERs across years based on the AAER release date. Relatively few AAERs were released prior to 1990. However, the number of AAERs increased particularly after 2000, when over one hundred AAERs were released per year.

The number of AAERs in 2005 falls to 94 because our sample cutoff date is June 10th, 2005 and so our sample does not include the full year.

Table 1 Panel C reports that in many cases there are multiple AAERs referring to the same firm. This is because the SEC can take action against multiple officers as well as the firm itself. The number of releases ranges from one per firm (376 firms) to a high of 24 per firm (Enron). From our reading of the AAERs we obtain a list of 899 firms mentioned in the 2,191 releases.

In phase two, we created the Annual and Quarterly files. All releases were reread thoroughly in order to identify the year and/or quarter-end when the manipulations occurred. Panel D of Table 1 indicates that of the 899 original firms identified, 219 firms involved either wrongdoings that are unrelated to financial manipulations (such as bribes or disclosure related issues) or financial manipulations that were not linked to specific reporting periods. This leaves us with 680 firms with alleged financial manipulations. We lose a further 168 firms because we are unable to obtain a valid Cusip identifier. This is frequently because the action was against a pre-IPO firm that either never went public or was public for less than a year. In addition, firms often change their names after a scandal, so it can be difficult to identify the new firm. For each firm that is in the Detail file but excluded from both the Annual or Quarterly files, we create four indicator variables in the Detail file to categorize why it was excluded. Panel D of Table 1 indicates that for 512 firms, the manipulation involved one or more quarters. In 101 firms the manipulation only involved quarterly financial statements and was corrected by the end of the year. Therefore the annual file contains manipulations of annual data for 411 firms.

[Table 1]

For each annual/quarterly period that was manipulated, an additional field was added to the Annual/Quarterly file. If an understatement of earnings or revenues occurred during the quarter or year of the violation, we code the *understatement* variable 1. Since most AAERs involve the overstatement of earnings or revenues, this flag is helpful in conducting earnings management and other discretionary accruals tests. The Annual file contains 1,060 firm-year observations, and the Quarterly file contains 4,481 firm-quarter observations.

Phase three involves reading the AAERs a final time in order to obtain additional details on the manipulations. For each firm, we summarize the reason(s) for the enforcement action(s) in one or two sentences in the "*explanation"* column of the Detail file. We then create eleven indicator variables to code the balance sheet and income statement accounts that the AAER identified as being affected by the manipulations.

Figure 2 indicates that 1,143 accounts were affected across the 680 manipulating firms. Most manipulations relate to revenue recognition, which occur in 54.6 percent of firms. Types of revenue manipulations include the following: front-loading sales from future quarters (e.g. Coca Cola, Computer Associates), creating fictitious sales (e.g., ZZZZ Best), incorrect recognition of barter arrangements (e.g., Qwest), shipping goods without customer authorization (e.g., Florafax International). Revenue manipulations also frequently involve a manipulation of the allowance for doubtful debts (Allowance). Other accounts frequently affected by manipulations include costs of goods sold and inventory (11.5 percent and 13.2 percent, respectively). Other types of manipulations include capitalizing expenses or creating fictitious assets (e.g., WorldCom). This occurs in about 26.9% of the firms.

Unfortunately, the AAERs do not provide consistent information on the magnitude of the manipulations. Some releases include details on the effect of manipulations on revenue, but not on income; and some releases provide before-tax numbers, while others provide after-tax numbers. In many releases, magnitudes are not disclosed at all. Therefore, there is insufficient detail to provide a consistent analysis of the magnitude of the manipulations.

[Figure 2]

4. EMPIRICAL RESULTS

Our empirical results are presented in five sections. Section 4.1 provides descriptive statistics on the characteristics of firms that manipulate their financial statements. Section 4.2 describes the variables that we hypothesize to be predictive of manipulations. Section 4.3 examines whether these variables differ systematically for manipulating firms in manipulation versus non-manipulation years. Section 4.4 examines whether these variables differ systematically for manipulation years versus all firm years for publicly traded firms listed on Compustat. Section 4.5 develops our fraud prediction model, while Section 4.6 provides a simple application of the *F-Score*.

4. 1 Characteristics of Manipulating Firms

Table 2 provides information about the size and industry membership of manipulating firms. Panel A presents information on size. To calculate size deciles, we rank firms based on their market capitalization of equity in each fiscal year. We then determine the decile rankings of manipulating firms in manipulation years. The results in bold identify the size deciles that are overrepresented in the manipulation firm population. The results indicate that 15.3 percent of firms that misstate their earnings are from the top size decile (decile 10), whereas only 5.3 percent are from the bottom decile (decile 1). There are several possible explanations for why larger firms appear to be relatively more likely to misstate their earnings. First, large firms have greater investor recognition and are under more scrutiny by the press and analysts; therefore, when an account appears suspicious there is likely to be more commentary that alerts the SEC to a potential problem (analyst and press reports are potential triggers for an SEC investigation). Second, large firms are more complex, consisting of many separate reporting units. It is possible for top managers to influence reporting outcomes using accounting data more easily, with less likelihood of audit detection. In addition, although managers in larger firms own a small percentage of the stock, this can represent a larger percentage of their wealth. Their incentives may be relatively less aligned with current and future investors. Third, the SEC has limited resources and so may see a better cost benefit trade-off to focusing its limited enforcement budget on relatively large firms.

Note that small firms appear to be underrepresented in the manipulation sample (5.3% versus an expected level of 10%). Even if a small firm has intentionally manipulated its financial statements, the dollar magnitudes involved and the aggregate losses to investors are likely to be relatively small. Therefore the SEC may choose not to allocate resources to such firms. Small firms are also less complex, and large auditors have more power and are more selective as to which firms they audit. Therefore, smaller firms could be more constrained by the auditors. Finally, recall that 168 firms are excluded from our analysis because we could not obtain their firm identifier. Many of these firms could be excluded from CRSP or Compustat because they delist after the SEC investigation and so either never went public or were public for a very short time. These excluded firms are likely to be smaller in size.

Panel B of Table 2 reports the industry distribution of both manipulation firm-years and all available firm-years on Compustat. We follow Frankel, Johnson, and Nelson’s (2002) SIC-based industry classification scheme. The bolded results highlight industries that are significantly overrepresented for manipulating firms. Over twenty percent of manipulating firms are in the computer industry, whereas only 11.9 percent of firms in the general population are in this industry. The computer industry includes software and hardware manufacturers. This industry is relatively new and has exhibited substantial growth. It is also characterized by substantial investment in intangible assets. Valuations in this industry are often dependent on continual sales growth. Manipulating firms frequently overstate their sales to meet optimistic business expectations (e.g., Computer Associates), ship goods without authorization (e.g., Information Management Technologies Corp), or create fictitious sales (e.g., Clarent Corporation and AremisSoft Corporation). Retail is also overrepresented among manipulating firms (13% versus 9.7%). For example, Sunbeam Corporation front-loaded sales and manipulated reserves for restructurings. Services are also overrepresented (12.4% versus 10.4%). Service firms include firms such as WorldCom, Qwest, and Waste Management. These firms typically capitalized expenses as assets and manipulated sales.

Panel C of Table 2 provides the distribution of manipulations over calendar time. Our sample covers manipulations in fiscal years beginning in 1971 and ending in 2003. The years 1999 and 2000 have by far the most manipulations (7.98% and 7.33% respectively). This may be because growth in technology stocks slowed around this time, providing incentives for managers to misstate earnings in order to mask declining growth.

[Table 2]

4. 2 Predictive Variables for Manipulations

Our next set of tests examines observable variables that we hypothesize to be associated with manipulations. This analysis provides the underpinnings for our subsequent development of our fraud prediction model. Since all variables are consistently reported on an annual basis, we focus only on the sample of firms with annual manipulations in these tests. The tests compare manipulation years to non-manipulation years. Manipulation years are separately compared to (i) all non-manipulation years; and (ii) only years prior to the manipulation. Using all firm years provides the most powerful tests, while using only prior firm years sheds light on the predictive ability of the variables with respect to manipulations.

We investigate several different variables that we hypothesize to be associated with manipulations. Each variable is briefly discussed below. More detailed definitions are provided in Table 3. The variables that we analyze are not intended to be exhaustive of all variables correlated with accounting manipulations. Previous literature has identified several corporate governance variables and non-financial performance variables correlated with accounting manipulations that we do not consider in our analysis. Our goal in this analysis is not to identify and analyze all variables correlated with accounting manipulations, but rather to explore variables that are available for the largest set of firms and readily accessible to accounting researchers and practioners. Focusing on this more limited set of variables allows us to create fraud prediction models that are more general. We leave it to future research whether these additional variables add significantly to the power of our fraud prediction models. The variables analyzed focus on accrual quality, financial performance, non-financial performance, off-balance sheet variables and stock market performance.

Accrual Quality

Starting with Healy (1985) a large body of literature hypothesizes that earnings are primarily manipulated via the accrual component of earnings. We therefore investigate whether manipulation years are associated with unusually high accruals. The first measure termed *Sloan* *accruals*, focuses on working capital accruals and is described in Sloan (1996). The measure includes changes in current assets (excluding cash) less changes in current liabilities (excluding short-term debt) less depreciation. Our next measure is from Richardson, Sloan, Soliman, and Tuna (2006) that we term *RSST accruals*. This measure extends the definition of *Sloan accruals* to include changes in long-term operating assets and long-term operating liabilities. This measure is equal to the change in non-cash net operating assets. We also look at two accrual components. The first is *change in receivables*. Manipulation of this account improves sales growth, a metric closely followed by investors. The second is *change in inventory.* Manipulation of this account improves gross margin, another metric closely followed by investors.

We also employ several 'discretionary accrual' models commonly used in the accounting literature to isolate accruals that are more likely to be attributable to manipulation. Our comprehensive sample of manipulations provides a unique opportunity to investigate whether these models enhance our ability to detect earnings manipulations. First, we employ the cross-sectional version of the *Modified Jones model discretionary accruals* (see Dechow, Sloan, and Sweeney 1996 for modified Jones model, and Defond and Jiambalvo (1994) for the cross-sectional version). We also investigate the effect of adjusting discretionary accruals for financial performance as suggested in Kothari, Leone, and Wasley (2005). We term this *Performance-matched discretionary accruals*.

Finally, we employ two variations of the accrual quality measure described in Dechow and Dichev (2002). The Dechow and Dichev measure is based on the residuals obtained from industry-level regressions of working capital accruals on past, present, and future operating cash flows. Our first variation on this measure takes the absolute value of each residual and subtracts the average absolute value of the residuals for each industry. We term this the *mean-adjusted* *absolute value of Dechow/Dichev residuals*. Our second variation scales each residual by its standard error from the industry-level regression. This measure leaves the sign of the residual intact and provides information on how many standard deviations the residual is above or below the regression line. We term this variable the *Studentized Dechow/Dichev residuals*. We predict a positive association between all accrual variables and manipulation years.

Performance

A potential reason for managers to misstate their financial statements is to mask deteriorating financial performance. Our next set of variables gauges the firm's financial performance on various dimensions. The first we analyze is *change in cash sales*. This measure excludes accruals-based sales, such as credit sales, and we use it to evaluate whether sales that are not subject to accruals management are declining. We also analyze *change in cash margin*. Cash margin is equal to cash sales less cash cost of goods sold. This performance measure abstracts from receivable and inventory manipulations. We anticipate that when cash margins decline, managers are more likely to make up for the decline by boosting accruals. *Change in earnings* is also analyzed since managers appear to prefer to show positive growth in earnings (e.g., Burgstahler and Dichev, 1997). During the manipulation periods, managers are likely to use accruals to show positive increases in earnings. However, it is also possible that even earnings manipulations are not sufficient to mitigate deteriorating earnings. *Change in free cash flows* is a more fundamental measure than earnings since it abstracts from accruals; however, managers may be less concerned about this measure since it is unlikely to play a role in their performance evaluations or in analysts' forecasts. We predict that managers are more likely to misstate when there is a decrease in free cash flows. We also investigate whether deferred tax expense increases during manipulation periods. Larger accounting income relative to taxable income is reflected in the deferred tax expense and could indicate more manipulation of book income (Phillips, Pincus, and Olhoft-Rego 2003).

The remaining variables attempt to analyze various aspects of firm performance that go beyond the basic financial statements.

Non-financial Measures

Economics teaches us that firms trade-off the marginal cost of labor against the marginal cost of capital to maximize profits. Investments in both labor and capital should lead to increases in future sales and profitability. However, unlike capital expenditures, most expenditure on labor must be expensed as incurred (the primary exception being direct labor that is capitalized in inventory). We therefore conjecture that managers attempting to mask deteriorating financial performance will also reduce employee headcount in order to boost the bottom line. Moreover, if managers are overstating assets, then the difference between the change in the number of employees, which is not likely overstated, and the change in assets, which is overstated, might be a useful measure of the underlying economic reality. We measure *abnormal change in employees* as the percentage change in the number of employees less the percentage change in total assets. We predict a negative association between *abnormal change in employees* and manipulations.

Greater order backlog is indicative of higher future sales. When a firm exhibits a decline in order backlog, this suggests a slowing demand and lower future sales. We measure *abnormal change in order backlog* as the percentage change in order backlog less percentage change in sales. We predict a negative association between *abnormal change in order backlog* and manipulations.

Off-Balance Sheet Activities

The most prevalent source of off-balance sheet financing is operating leases. The accounting for operating leases allows firms to record lower expenses early on in the life of the lease (because the interest charge implicit in capital lease accounting is higher earlier on in the life of the lease). Therefore, the use of operating leases (*existence of operating leases*) and unusual increases in operating lease activity (*change in operating lease activity*) could be indicative of managers who are focused on financial statement window-dressing. We predict that *change in operating lease activity* is positively associated with manipulations. *Change in operating lease activity* is measured as the change in the present value of future non-cancelable operating lease obligations following Ge (2006).

Another off-balance sheet activity is the accounting for pension obligations and related plan assets for defined benefit plans. Firms have considerable flexibility on the assumptions that determine pension expense. The expected return on plan assets is an assumption that is relatively easy for managers to adjust. Management can increase the expected return on plan assets and immediately decrease currently reported pension expense. Comprix and Mueller (2006) provide evidence that such income-increasing adjustments are not filtered out of CEO compensation. Therefore, similar to leases, such legitimate adjustments could be indicative of managers who are focused on financial window-dressing. For the subset of firms that have defined benefit plans we obtain the *expected return on pension plan assets* and calculate *the change in expected return on pension plan assets*. We predict that in manipulation years, firms will assume larger expected returns on their plan assets.

Market-related Incentives

One obvious incentive for manipulating earnings is to maintain a high stock price. We investigate whether managers who misstate their financial statements are particularly dependent on a high stock price. We examine two motivations. First, if the firm needs to raise cash to finance its ongoing operations and growth plans, then a high stock price will reduce the cost of raising new equity. High book value, consistent earnings performance and a high stock price will also reduce the cost of issuing new debt. We use various empirical constructs to capture a firm's need to raise additional capital. First, we use an indicator variable identifying whether the firm has issued new debt or equity during the manipulation period (*actual issuance*). Second, we look at the net amount of new financing raised, deflated by total assets (*CFF*). Third, we construct a measure of *ex ante financing need*. Some firms may have wished to raise new capital, but did not because they were unable to secure favorable terms; our ex ante measure of financing need provides a measure of the incentive to raise new capital. Following Dechow, Sloan, and Sweeney (1996) we report an indicator variable that equals one if the firm is estimated to have negative free cash flows over the next two years that exceed its current asset balance. Finally, we examine leverage. We expect that managers of firms with higher leverage will have incentives to boost financial performance both to satisfy financial covenants in existing debt contracts and to raise new debt on more favorable terms.

A second motivation for why managers may be particularly dependent on a high stock price is because a significant portion of management compensation is typically tied to stock price performance. This can cause managers to become overly concerned with maintaining or increasing their firm's stock price, since it affects their wealth. Such managers can become focused on managing "expectations" rather than managing the business. We expect that managers whose firms have had large run-ups in their stock prices and have high prices relative to fundamentals are more prone to "expectations" management. Managers of such firms are predicted to be more likely to misstate earnings to hide diminishing performance. We identify firms with optimistic expectations built into their stock prices using *market-adjusted stock return*, *earnings-to-price,* and *book-to-market*.

[Table 3]

4.3 Time-series Analysis of Manipulating Firms

Table 4 provides our time-series analysis of manipulating firms. This table reports descriptive statistics on the variables identified in Section 4.2 for manipulation years versus non-manipulation years. Panel A compares manipulation years to all available non-manipulation years. Panel B compares manipulation years to years prior to the first manipulation year. We provide Panel B to identify variables that are most likely to be useful in predicting manipulations. For example, manipulating firms typically report deteriorating future performance. But while deteriorating future performance may be associated with manipulations, it cannot be used to predict manipulations. Thus, Panel A sheds light on the overall characteristics of manipulation years, while Panel B focuses on characteristics that are most useful in predicting manipulations.

Table 4 Panel A begins with our various measures of accrual quality. We predict that accruals will be larger in manipulation years. The results indicate that *RSST accruals* has a slightly larger t-statistic than the *Sloan accruals* measure, suggesting that the more comprehensive RSST measure of accruals is more effective at detecting manipulations. *Change in receivables* has the highest t-statistic of all accrual variables of 6.71, probably because half of the manipulating firms are alleged to have manipulated sales. The next set of accrual variables relate to various models of 'discretionary' accruals. The objective of these models is to provide more powerful measures of earnings management by eliminating 'nondiscretionary' accruals that are required under GAAP. However, such modeling comes at the cost of unintentionally removing some of the 'discretionary' accruals. The t-statistic on the *Jones discretionary accrual* models is lower than that on either the *Sloan* or *RSST* model, suggesting that this model could suffer from this problem. Interestingly, *performance-matching* has little effect on the results and if anything, reduces the power of detecting manipulation since it has a lower t-statistic than that on the regularly calculated *Jones* model. The *Studentized Dechow/Dichev* model that takes into account the sign of the residual appears to be the most powerful discretionary accrual model.

We next examine various measures of financial performance. We predict that manipulations are often made to mask deteriorating financial performance. Our first measure is *change in cash sales*. Contrary to our expectations, cash sales actually increase (rather than decline) during manipulation years. A reading of the AAERs helps to explain why. We find that many firms engage in transactions-based earnings management. That is, they front-load their sales and engage in unusual transactions at the end of the quarter (e.g., Coca Cola, Sunbeam, Computer Associates). Cash sales increase with this type of manipulation, providing an explanation for the finding. Cash margins, however, are declining, suggesting that operating performance is deteriorating at the time of manipulations. Earnings are also declining at the time of manipulation, suggesting that accruals are being used to mask the extent of decline. *Change in free cash flows* is not significantly different across manipulation and non-manipulation years. *Deferred tax expense* is also not significantly different. For a small sample of 27 firms subject to SEC enforcement actions, Erickson, Hanlon, and Maydew (2004) show that firms pay substantial taxes on overstated earnings. For example, manipulating cash sales boost both accounting and tax income. If their findings are generalizable, then this could explain why deferred taxes are not unusually high during manipulation years.

We next turn to the non-financial variables, *abnormal change in employees* and *abnormal change in order backlog*. Both variables show economically significant declines during manipulation years, though the small sample size for *abnormal change in order backlog* fails to achieve statistical significance. For our off-balance sheet variables, we find an increase in both the magnitude of operating lease commitments and the percentage of firms that use operating leases during manipulation years. It appears that manipulating firms are quick to exploit the financial reporting flexibility afforded by operating leases. For defined benefit pension plans we have only a small sample size. However the results are suggestive of earnings management. We find that the *expected return on pension plan assets* is higher in manipulation years and that the *change in expected return on pension plan assets* is significantly greater in manipulation years.

The final set of variables captures market-related incentives. As predicted, we find that *ex ante need for financing* is significantly greater in manipulating years (18%) than in non-manipulating years (10%). More firms are issuing either debt or equity (93% versus 88%) and cash from financing more than doubles during manipulating years (19% versus 7%). *Leverage* is not significantly different. *Market-adjusted stock return* is higher during manipulation years (16.3% versus 7.2%). We analyze this finding in more detail later in the paper. *Book to market* ratios are not significantly different while *earnings to price* ratios are lower in manipulating periods, consistent with our prediction that manipulating firms have optimistic future earnings growth expectations built into their prices.

Panel B replicates the analysis in Panel A using only years prior to the manipulation as non-manipulation years. The results are generally consistent with those in Panel A, but there are a few points to note. First, the significance of the accrual variables declines. For example, the difference between manipulating years and non-manipulating years declines by more than half for RSST accruals (0.077 to 0.036). This suggests that the inclusion of the subsequent accrual reversal boosts the power of these tests (e.g., the subsequent receivable inventory write-off). Note that the modified Jones Model pools across years to calculate the industry coefficients, and the Dechow and Dichev models use future cash flows, so these models would not be completely implementable for financial statement users. The power of both models is relatively unchanged across the two panels, but they involve an implicit hindsight bias.

The results for the performance variables, off-balance sheet variables and market-related variables are similar across panels A and B. The results for book-to-market are now significant and in the predicted direction. Prior to the manipulation, these firms had relatively high market valuations relative to earnings or book value. Thus one reason managers may have manipulated earnings was to maintain the current stock price at artificially high levels.

[Table 4]

Figure 3 provides a graphical timeline of (a) annual raw stock returns; and (b) annual market-adjusted stock returns for manipulating firms before and after the manipulation years. Both graphs reveal that returns are increasing in the three years leading up to the manipulation. In the manipulation years, on average, the firms are able to maintain positive stock returns.[4] However, in the first year after the manipulation years, the stock prices decline and returns are negative. The negative returns likely result from the revelation of the manipulation (Karpoff, Lee and Martin 2007).

---

[4] For the firms misstating for multiple years, we take the average of their stock returns during the manipulation period.

[Figure 3]

4.4 Cross-sectional Analysis of Manipulating Years

Our next test compares manipulating firm-years to all firms listed on the Compustat Annual File between 1979 and 2002. We limit the sample to these years since the first AAER release occurred in 1982, and very few firms are identified as manipulating prior to 1979. Using the AAER database, we identify 338 firms with 594 firm-year observations for our large cross-sectional sample. These tests identify unusual characteristics of manipulating firms relative to the general population. We make this comparison since it is helpful to auditors and investors to make both time-series and cross-sectional comparisons.

Table 5 replicates the analysis in table 4, but compares manipulating years to all firm-years available on Compustat. We exclude the performance matched discretionary accruals, since this adjustment is redundant when using the entire population.

The results for the accrual quality related variables are very similar to those reported in Table 4. The accruals of manipulating firms are unusually high relative to the population. For example, in manipulating years the *RSST accrual* measure is 11.9 percent of assets; whereas, for the population, this measure is 2.9 percent of assets. Similarly, *change in receivables* is 5.8 percent for manipulating firms, whereas growth for the population is only 2.1 percent. The *studentized Dechow/Dichev* measure indicates that manipulating firms' residuals are on average 0.39 deviations from the regression line in the positive direction.

For the performance variables, *change in cash sales* for manipulating firms is about twice as large as for the population (0.468 versus 0.211). As mentioned earlier, this is probably because manipulating firms are front-loading sales. However, on other dimensions, performance for manipulating firms is poor relative to the population. The *change in cash margins* and the *change* *in earnings* are both significantly lower for manipulating firms. The results for non-financial variables and off-balance sheet variables are all in the predicted direction. One difference from Table 4 is that manipulating firms assume significantly higher expected returns on their plan assets than other firms (7.88% versus 7.17%). However, the change in expected returns is no longer significantly different. Note also, that *abnormal change in order backlog* is not statistically significant.

Finally, for the market-related variables, the results indicate that demand for external financing is higher for manipulating firms than for the average firm in the population. We report market-adjusted stock returns in the manipulation year and the prior year. Compared to the average firm, manipulating firms have significantly greater returns in both years. In addition, manipulating firms have high valuations relative to fundamentals when compared to the Compustat population. Similar to the results in Table 4 Panel B, both book-to-market and earnings-to-price are significantly lower for manipulating firms (i.e., they have high valuations relative to fundamentals). The results in Table 5 confirm that the variables identified as unusual in time-series analysis also tend to be unusual in cross-sectional analysis.

[Table 5]

To provide more intuition for how manipulating firms differ from the population, we select two very well-known fraud firms, Enron and Waste Management, and examine their characteristics. Figure 4 compares various measures for these two firms to the average firm on Compustat. In the figures, Enron is listed first, then Waste Management, and then the average firm. Figure 4a provides various measures of accrual quality. Waste Management appears relatively conservative under the Sloan definition of accruals, while Enron appears aggressive.

However, both firms appear very aggressive using the *RSST accruals* measure. Waste Management has *RSST accruals* nearly ten times greater than the average firm. This is not surprising given that Waste Management primarily manipulated their earnings through adjustments to reduce depreciation expenses and non-GAAP capitalization of certain expenses as long-term assets. Enron and Waste Management's *change in receivables* are twice that of the average firm. In terms of performance, both firms show large *change in cash sales* of over 60 percent for Enron and 140 percent for Waste Management (Figure 4b), yet cash margins are declining, as well as earnings (Figure 4c). In addition, both firms are cutting back on the number of employees relative to their asset base, with the abnormal decline in employees being twice as large for Waste Management, and six times as large for Enron compared to the average firm (Figure 4d). Finally, for the market-related variables we see that both Enron and Waste Management show superior stock price performance prior to the first manipulation year, with Enron outperforming the market by 3 percent in the year before manipulation and 30 percent in the first year of manipulation. Waste Management outperforms by 26 percent in the year prior to manipulation and four percent in the first manipulation year. In addition, both firms have high price-to-earnings ratios relative to the average firm. As we know, managers of these firms are very sensitive to their firms' stock prices.

[Figure 4]

4.5 Prediction Analysis and Development of the *F-Score*

In this section we provide multivariate analysis of variables identified in Tables 4 and 5. Manipulations resulting in SEC Enforcement Actions are rare events. Our manipulation sample represents less than half of one percent of the firm-years available on Compustat. However, manipulations are extremely costly to the auditor (in terms of lawsuits), to investors (in terms of negative stock returns), to regulators like FASB and SEC (in terms of reputation for quality and enforcement of accounting rules), and to capital markets (in terms of lost investor confidence and reduced liquidity). Therefore, even though manipulations are rare, a model that can help identify manipulations is useful.

Table 6 provides a correlation matrix between the variables used in our analysis. To save space, we eliminate variables that are statistically insignificant in Table 5, along with certain indicator variables. The correlations use all available firm-years listed on Compustat, so the number of observations is over 100,000 firm-years. Given this large sample size, all correlations exceeding 0.00 in magnitude are significant at conventional levels. The first thing to note is that the correlations between the *manipulation flag* (*manipflag*, an indicator variable for manipulation firm-years) and our variables are low. This is because *manipulation flag* is zero for over 99 percent of observations, which highlights the difficulty of correctly detecting manipulation firms. The accrual quality variables are positively correlated with each other with the highest correlation being 72 percent between *Sloan accruals* (*Sloan_acc*) and *modified Jones model Discretionary accruals* (*da*). These measures both focus heavily on working capital accruals. Generally, the correlations between accrual variables range between 20 and 40 percent. The one exception is *mean-adjusted absolute value of Dechow/Dichev residuals* (*resid*), which is negatively correlated with all accrual variables. Recall that *resid* is the absolute value of the residual from the Dechow/Dichev regression. Dechow and Dichev (2002) argue that when firms have accruals that do not match to past, present and future cash flows, they are likely to have less persistent earnings. Therefore, the sign of the deviation is not important for their analysis. However, in the correlation matrix all other variables are signed. Since *resid* is likely to be larger for firms with both extreme low and high accruals, it is not highly correlated with the other accrual measures. Note, however, that *studentized Dechow/Dichev residuals* (*sresid*) (which is signed) varies in the predicted direction with other measures of accruals. *Change in earnings* (*ch_earn*) tends to be positively correlated with accruals. *Abnormal change in employees* (*ch_emp*) is generally negatively correlated with accruals, suggesting a substitution effect between capital and employees. External financing (*CFF* (*cff*) and *actual issuance* (*issue*)) tends to be positively correlated with accruals. *Market-adjusted stock return* (ret_t) and *lagged market-adjusted stock return* (ret_{t-1}) are also positively related to accruals.

[Table 6]

Table 7 provides our logistic models for the determinants of manipulations. Our dependent variable is equal to one for firm-years involving a manipulation, and zero otherwise. We estimate logistic regressions to determine whether the variables we have examined in univariate tests are jointly significant in predicting manipulation firm-years. We build three models for fraud prediction. **Model 1** includes only financial statement variables as predictors; **Model 2** adds non-financial statement and off-balance sheet variables; and **Model 3** incorporates market-based measures. We form our models in this way so we can see the incremental benefit from including information beyond the financial statements for predicting manipulation. Since Table 6 indicates that some of the variables are correlated and we seek a parsimonious prediction model, we use a backward elimination technique to arrive at our prediction models. The backward elimination technique begins with all of our selected variables; we then use the computational algorithm of Lawless and Singhal (1978) to compute a first-order approximation of the remaining slope estimates for subsequent variable eliminations. Variables are removed based on these approximations. We set the significance level for elimination at the 20% level.[5]

Model 1 begins with our accruals quality measures, the performance measures, and the market-related measures that are computed from variables in the financial statements (ex ante finance need, actual issuances, cash from financing, and leverage). We did not include discretionary accrual measures (the modified Jones model or the Dechow Dichev model) because we want variables that can be relatively easily calculated from the financial statements. Both models require knowledge of other firms in the industry. In addition, the Dechow/Dichev model includes future cash flows and so is not predictive. After performing backward elimination, we retain the following variables: *RSST accruals, change in receivables, change in inventory, change in cash sales, change in earnings, and actual issuance*. For **Model 2**, we retain the variables from **Model 1** and add the non-financial variables and off-balance sheet variables. After backward elimination, we retain *abnormal change in employees* and *existence of operating leases*. For **Model 3**, we add our market-based variables (our two return measures, earnings-to-price, and book-to-market). From which, *lagged market-adjusted stock return* and *book to market* are retained in the model after backward elimination. Table 7 Panel A provides the resulting coefficient estimates for the models. The coefficients are all in the predicted direction.

To examine the quality of our models, we sort and rank firms into quintiles based on the predicted probabilities that the model assigns to each observation. Predicted values are obtained by plugging each firm's individual characteristics into the model and using the estimated coefficients to determine the predicted value. The predicted probability is derived as:

$$Probability = \frac{e^{(PredictedValue)}}{(1+e^{(PredictedValue)})}$$

[5] We run the logistic procedure in SAS, with the model selection equal to BACKWARD and FAST.

We then divide the probability by the unconditional expectation of manipulation to calculate our *Fraud Score* (*F-Score*). The unconditional expectation is equal to the number of manipulation firms divided by the total number of firms. Below is an example of how this is done for **Model 1** for Enron.

Enron in 2000

Predicted Value:
=-6.753+.773 x (rsst_acc)+3.201 x (ch_rec)+2.465 x (ch_inv)+.108 x (ch_cs)+-0.995 x (ch_earn)+.938 x (issue)
Predicted Value:
=-6.753+.773 x (.01659)+3.201 x (.17641)+2.465 x (.00718)+.108 x (1.3333)+-0.995 x (-.01285)+.938 x (1)
Predicted Value=-5.063
Probability $=e^{(-5.063)} / (1+e^{(-5.063)})$
e= 2.71828183
Probability =0.0063
Unconditional probability = 494/(143,452 + 494) = 0.0034
F-Score = 0.0063/0.0034
F-Score for Enron = 1.85

An *F-Score* of 1.00 indicates that the firm has the same probability of fraud as the unconditional expectation. *F-Scores* less than one indicate a lower probability of fraud. *F-Scores* greater than one indicate higher probabilities of fraud than the unconditional expectation. Enron has an *F-Score* of 1.85. This indicates that Enron has almost twice the probability of being a fraud firm as a randomly selected firm from the population.

Table 7 Panel B ranks firm-years into five portfolios based on the magnitude of their *F-Score*. We report the frequency with which manipulating and non-manipulating firms fall into each quintile and the minimum *F-Score* required to be included in each quintile. If our models do a good job in identifying manipulation firms, then we expect manipulation firms to be clustered in the fifth portfolio. The results for **Model 1** that include only financial statement variables indicate that 46.56 percent of manipulation firms are in Quintile 5, compared to the expected level of 20 percent. The cut-off to be included in Quintile 5 (i.e., the minimum value) is 1.224 and so Enron's score for 2000 of 1.85 easily places it in Quintile 5. **Model 2** that includes non-financial and off-balance sheet variables indicates that 46.56 percent of manipulation firms are in Quintile 5, while for **Model 3** that includes market-related variables 44.48% are included in Quintile 5. Figure 5 provides a graphical representation of the results in Panel B. Quintile 5 clearly contains a larger proportion of manipulating firms than expected by chance.

[Figure 5]

Note that the number of observations declines as we increase the data requirements for the sample. This makes direct comparisons across the models difficult. We reran **Model 1** and **2** using only observations available for **Model 3** and find that the number of manipulating firms in Quintile 5 is 155 (42.8%) for **Model 1** and 155 (42.8%) for **Model 2**. Thus, **Models 3** which classifies 161 (44.48%) firms in Quintile 5 provides a small improvement over **Model 1** and **Model 2**, but has greater data requirements. In addition, we found that substituting the annual market-adjusted return rather than the lagged return increased the number of manipulating observations by 14 percent from 362 to 413 and for the non-manipulating firms by 7,599 (7.9%) from 95,168 to 102,767. The coefficient on the contemporaneous return is 0.0561 (versus 0.063 for the lagged return) with a p-value of 0.162. The number of observations in Quintile 5 when the contemporaneous return is used is 187 (45.3%). Therefore, we recommend substitution of the contemporaneous annual return when the lagged-return is not available since this will still provide a small improvement over **Model 1**.

Another way to consider the predictive ability of the models is to consider *Type I* and *Type II* errors. A *Type I* error occurs when our model incorrectly classifies a non-manipulating firm as a manipulating firm. A *Type II* error occurs when our model incorrectly classifies a manipulating firm as a non-manipulating firm. The cost of these two types of errors is not likely the same and is likely to differ for each type of user. From an auditor's perspective a *Type II* error is by far the more costly. When a manipulation goes undetected (and is later revealed), the auditor is likely to be sued by investors and sanctioned by regulatory bodies such as the SEC and the PCAOB. A *Type I* error (a non-manipulating firm is suspected of manipulation) is not costless and may result in lost fees, as the auditor may choose to drop the client or charge a fee that is so high that the client fires the auditor. Since *Type II* errors are more costly to the auditor, an auditor is likely to prefer a model that makes more *Type I* errors than *Type II* errors. This trade-off will determine the *F-score* cut-off that minimizes the auditor's costs.

In Panel C of Table 7, we set the F-score cut-off to 1.00 so that all firms with a higher probability than expected by chance are assumed to be manipulating. A score of 1.00 captures approximately the top 40% of firms in terms of likelihood of manipulation. The results for **Model 1** indicate that we correctly classify 324 of the 494 firms correctly (65.59%). The *Type I* error rate (false classification of a regular firm) is 35.59%. For **Model 2** and **Model 3** there is a slight decline in the sensitivity ratio (correct classification of manipulating firms) to 64.97% and 62.98%, respectively.

[Table 7]

Figure 6 provides further insights into the trade-off between *Type I* versus *Type II* errors. Figure 6A provides the error rates for **Model 1** in Table 7. At an *F-Score* of 0.000 all firms are classified as manipulating firms, so the *Type I* error rate is 100% and the *Type II* error rate is 0%. As higher *F-Scores* are selected the *Type I* error rate declines, while the *Type II* error rate increases. At an *F-Score* cut-off of 1.00, the *Type I* error rate declines from 100% to 33% while the *Type II* error rate increases from 0% to 36% (170 of the 494 manipulating firms have *F-Scores* less than 1.00).

How should a user think of the trade-off between *Type I* versus *Type II* errors? We provide one approach in Figure 5B. Figure 5B reports the relative cost of errors ratio calculated as the number of *Type I* errors divided by the sensitivity (correctly classified manipulating firms) for each *F-score* cut-off. From an auditor's perspective, assume that the cost of investigating a firm for manipulation is $1. Also assume that the detection rate for firms investigated is 100% (all firms investigated that have manipulated are detected). When the cost of missing a manipulation firm (in terms of lawsuits) is over $290, then Figure 5B indicates that an *F-Score* of 0.000 should be used and all firms are investigated. At the other extreme, if the cost of missing a manipulation firm is less than $50, then Figure 5B indicates that no firms should be investigated (i.e., just pay the lawsuits as they occur). Where the Model is useful is for cases between these extremes. At an *F-Score* cut-off of 1.00, 51,061 of the 143,452 non-manipulating firms have *F-Scores* greater than 1.00, while 324 of the 494 manipulating firms have *F-Scores* greater than 1.00. The relative cost ratio is 158 (51,061/324). No investigation of *F-Scores* greater than 1.00 would result in missing 324 manipulating firms. If the cost of investigating a non-manipulating firm is 158 times the cost of missing a manipulating firm, then the decision rule is to investigate all firms with an *F-Scores* cut-off of 1.00.[6]

[Figure 6]

[6] To see how costs are reduced note that in the *all firms investigated* scenario the cost to investigate is $143,453, while the cost of not investigating is 494 x $300 = $148,200. Therefore, the auditor investigates all firms and does not use the model. At the other extreme, if the cost of a lawsuit is less than $50, then the maximum cost incurred investigating no one is 494 x $50 = $24,700. However, if it is approximately 158 times more costly to have a lawsuit than to investigate a non-misstating firm, then the cut-off the auditor should make is investigate firms with *F-Scores* greater than 1.00. At this point the cost of not investigating is 324 x $158 = $51,192, while the cost of investigating non-misstating firms is 51,062 x $1 = $51,062.

A shortcoming of the analysis in Table 7 is that we develop our prediction model and evaluate its effectiveness using the same sample. We next investigate the robustness of this analysis by estimating model parameters in an early time period and evaluating the predictive ability of the model in a later hold-out period. Specifically, we re-estimate the models presented in Table 7 for the time period 1979 to 1998. We then use the new estimates from these models to predict the probability of manipulation (*F-Scores*) for the hold-out sample of firm-years from 1999 to 2002. Note that we determine the variables for inclusion in the model based on the entire sample. Therefore this approach still has some residual look-ahead bias. However, this analysis evaluates the stability of the coefficients and specification of the model.[7]

The estimated coefficients from the 1979 to 1998 model are presented in Table 8 Panel A. Note that the coefficients on the financial statement variables are very similar to those in Panel A of table 7. The most significant differences are that the coefficient on abnormal change in employees becomes less negative, while the coefficient on book-to-market becomes more negative.

After estimating the models, we then test our models out of sample by using the estimated parameters in Table 8 Panel A to assign probabilities to the hold-out sample. We rank the hold-out sample firms into quintiles and report the frequency and mean probabilities for manipulating and non-manipulating firms by quintiles. The results are reported in Table 8 Panel B. Compared to the results in Table 7 Panel B, all models show a slight improvement in the percent of manipulating firms classified in Quintile 5. For **Model 1** the percentage is 47.37 versus 46.56% in Table 7; for

[7] We also reran the entire analysis using only the 1979 to 1998 time period. **Model 1** adds cash from financing (*CFF*) to the variables already identified; **Model 2** drops change in cash sales but does not add *CFF*. **Model 3** drops *abnormal change in employees* and also does not add *CFF*. Therefore the variables in the model are relatively insensitive to the time period used to estimate the model.

Model 2 the percentage is 50.81 versus 46.56% in Table 7; and for **Model 3** the percentage is 45.10 versus 44.58% in Table 7.

[Figure 6]

In Panel C of Table 8 we provide further analysis using our *F-score* cut-off of 1.00. Note that the minimum cut-off probabilities for Quintiles 4 and 5 have declined in Table 8. For example for **Model 1** the minimum *F-Score* to enter Quintile 5 declines from 1.224 in Table 7 to 1.145 in Table 8. Therefore, an *F-Score* of 1.00 is a relatively higher threshold in Table 8 and so we could end up misclassifying more misstating firms. This is observable for **Model 1**, where 57.89% of manipulating firms are correctly classified (versus 65.59% in Table 7). However, for **Models 2** and 3, the *F-Score* cut off of 1.00 produces very similar correct classification. For **Model 2** 64.52% of manipulating firms are correctly classified in Table 8 (versus 64.97% in Table 7). For **Model 3**, 60.78% of manipulating firms are correctly classified in Table 8 (versus 62.98% in Table 7). Overall an *F-Score* cut-off of 1.00 in Table 8 produces similar results to those reported in Table 7.

[Table 8]

4.6 Simple Application of the *F-Score* to Audit Firms

Our final test provides one simple application of the *F-Score*. For this test we investigate the *F-Scores* of manipulating firms for each audit firm. Having a client firm investigated by the SEC is costly to audit firms in terms of litigation payments and audit partner time. Could knowing the *F-Score* of a client firm in the period when the client was manipulating the financial statements have been a red flag to the auditor? We identify the auditor of 348 manipulating firms in our annual database. We use the name of the auditor at the time of manipulation. We then examine the *F-score* of firms with available data during manipulation years for each auditor. We select the maximum *F-Score* during manipulation years for each client firm. We use the maximum since this represents the strongest signal of a potential manipulation. We then report the median *F-Score* across client firms for each auditor.

Figure 7 indicates that the *F-scores* for the Big-four auditors are 1.186 for PWC, 1.195 for KPMG, 1.433 for Ernst & Young, and 1.679 for Deloitte & Touche. Further analysis is required to understand these cross-sectional differences. For example, does Deloitte & Touche specialize in auditing higher risk industries (so that on average, their clients have higher *F-Scores*), or could their profits improve by greater screening of clients? Historically, it appears that the large audit firms had more risky clients than they do today. For example, Price Waterhouse has an *F-Score* of 2.112 and Peat Marwick, Mitchell has an *F-Score* of 2.053. Does this reflect a general decline in the level of acceptable audit risk due to litigation concerns? More generally, do audit firms charge higher audit fees to compensate for the risk of high *F-Score* firms? How do *F-Scores* vary across audit partners? In addition, it would be interesting to determine whether high *F-Score* firms tend to be ones with internal control problems that require more audit work following the Sarbanes-Oxley Act.

[Figure 7]

5. CONCLUSION

This paper provides a comprehensive sample of firms investigated by the SEC for manipulating earnings. We conduct a detailed analysis of 2,191Accounting and Auditing Enforcement Releases available between 1982 and 2005 and identify 680 firms with manipulated quarterly or annual earnings. We document the most common types of manipulations and find that the overstatement of revenues and reserves are the most frequent types of manipulations. We also identify the industries and time periods in which manipulations are most common.

We investigate the characteristics of manipulating firms on various dimensions, including accrual quality, financial performance, non-financial performance, off-balance sheet activities, and market-related variables. We find that at the time of manipulations, accrual quality is low and both financial and non-financial measures of performance are deteriorating. We also find that financing activities and related off-balance sheet activities are much more likely during manipulation periods. Finally, we find that managers of manipulating firms appear to be very sensitive to their firm's stock price. These firms have experienced strong recent earnings and price performance and trade at high valuations relative to fundamentals. The manipulations appear to be made with the objective of covering up a slowdown in financial performance in order to maintain high stock market valuations.

Based on the above findings, we develop a logistic model to determine the probability of manipulations. The output of this model is an *F-Score* – a scaled probability that a firm has engaged in an earnings manipulation. We show that our models have power to detect manipulations both within sample and using a holdout sample. Using a cut-off *F-Score* of 1.00, we find that our models correctly identify over 60 percent of manipulating firm-years. We suggest that the *F-Score* can be used as a first-pass screening device for detecting possible manipulations.

Our paper provides useful insights into research on earnings management. Prior research has generally focused on measures of discretionary accruals as proxies for incentives to engage in earnings management. Our results suggest that researchers could also consider using an *F-Score* as an alternative proxy. In addition, we find that growth in cash sales is unusually high during manipulation years. An important avenue for future research on earnings management is to develop models of cash flow management. Existing models focus on accruals management, but manipulating firms also accelerate cash sales in order to boost earnings. Our results also raise interesting avenues for future research in better understanding audit client selection and risk. For example, how do *F-Scores* vary across audit firms and what explains the underlying differences?

One unavoidable issue in developing models to detect manipulation is that the revelation of a manipulation is a rare event. Thus, similar to bankruptcy prediction models, our models generate a high frequency of false positives (i.e., many firms that do not have enforcement actions against them are predicted to have manipulated their earnings). An obvious limitation of our analysis is that we can only identify manipulations that were actually identified by the SEC. There are likely many cases where a manipulation goes undetected, or is at least not subject to an SEC enforcement action. An interesting avenue for future research would be to investigate other high *F-Score* firms. For example, do high *F-score* firms engage in earnings management, within the realms of GAAP? Do they experience declines in subsequent financial and stock price performance? Are they more likely to record future asset write-offs or write-downs? In addition, can developing models at an industry level reduce the frequency of false positives?

Finally, our analysis should provide useful insights to auditors, regulators, investors, and other financial statement users about the characteristics of manipulating firms. By better understanding these characteristics, financial statement users should be in a better position to identify and curtail manipulation activity in the future. The efficient functioning of capital markets depends crucially on the quality of the financial information provided to capital market participants. Curtailing manipulation activity should lead to improved financial information and hence improved returns for investors and more efficient allocation of capital.

References

Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance 23: 589-609.

Beneish, Messod D. 1997. Detecting GAAP Violations: Implications for assessing earnings management among firms with extreme financial performance. Journal of Accounting and Public Policy, 16: 271-309.

Brazel, Joseph F., Keith L. Jones, and Mark F. Zimbelman. 2006. What can nonfinancial measures tell us about the likelihood of fraud? Working Paper. North Carolina State University.

Burgstahler, David and Ilia Dichev. 1997. Earnings management to avoid earnings decreases and losses. Journal of Accounting and Economics, 24:99–126.

Comprix Joseph and Karl A. Mueller III. 2006. Asymmetric treatment of reported pension expense and income amounts in CEO cash compensation calculations. Journal of Accounting and Economics, 42: 385–416.

Dechow, Patricia M, and Ilia Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77 (Supplement): 35-59.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 1995. Detecting earnings management. The Accounting Review 70: 193-226.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actins by the SEC. Contemporary Accounting Research 13: 1-36.

Dechow, Patricia M. and Ilia Dichev, 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77 Supplement: 35-59.

DeFond Mark L. and James Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. Journal of Accounting and Economics 17: 145-176.

Erickson, Merle, Michelle Hanlon, and Edward L. Maydew. 2004. How much will firms pay for earnings that do not exist? Evidence of taxes paid on allegedly fraudulent earnings. The Accounting Review 79, 387-408.

Ettredge, Michael, Lily Sun, Pinching Lee, and Asoka Anandarajan. 2006. Do deferred tax data signal earnings fraud? Working Paper. University of Kansas - School of Business.

Frankel, M. Richard, Marilyn F. Johnson and Karen K. Nelson, 2002. The relation between auditors' fees for no audit services and earnings management. The Accounting Review 77: 71-103.

Ge, Weili, 2006. "Off-balance-sheet activities, earnings persistence and stock prices: Evidence from operating leases." Working paper. The University of Washington.

Healy, Paul, 1985. The effect of bonus schemes on accounting decisions. Journal of Accounting and Economics 7: 85-107.

Karpoff, Jon., Scott. Lee and Gerald. Martin, 2007. The cost to firms of cooking the books. Journal of Financial and Quantitative Analysis, forthcoming .

Kothari, S.P, Andrew Leone, and Charles Wasley, 2005. Performance matched discretionary accrual measures. Journal of Accounting and Economics 39: 163-197.

Lawless J. F. and Singhal K. 1978. Efficient screening of nonnormal regression models. Biometrics 34: 318-327.

Lev, Baruch S. Ram Thiagarajan. 1993. Fundamental information analysis. Journal of Accounting Research, 31:190-215.

Phillips, John, Morton Pincus, and Sonja Olhoft Rego. 2003. Earnings Management: New Evidence Based on Deferred Tax Expense. Accounting Review, 78 Issue 2: 491-521.

Rangan Srini. 1998. Earnings management and the performance of seasoned equity offerings. Journal of Financial Economics 50: 101-122.

Richardson, Scott, Irem Tuna, and Min Wu. 2002. Predicting earnings management: The case of earnings restatements. Working Paper, University of Pennsylvania.

Roychowdhury, Sugata. 2006. Earnings management through real activities manipulation. Journal of Accounting and Economics, 42; forthcoming.

Skousen, Christopher J. and Charlotte J. Wright. 2006. Contemporaneous risk factors and the prediction of financial statement fraud. Working Paper. University of Texas at Arlington.

Sloan, Richard G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71: 289-315.

Teoh Siew Hong, Ivo Welch, and T. J. Wong 1998. Earnings management and the underperformance of seasoned equity offerings. Journal of Financial Economics 50: 63-99.

Appendix 1: Variable Definitions of the Enforcement Releases Datasets

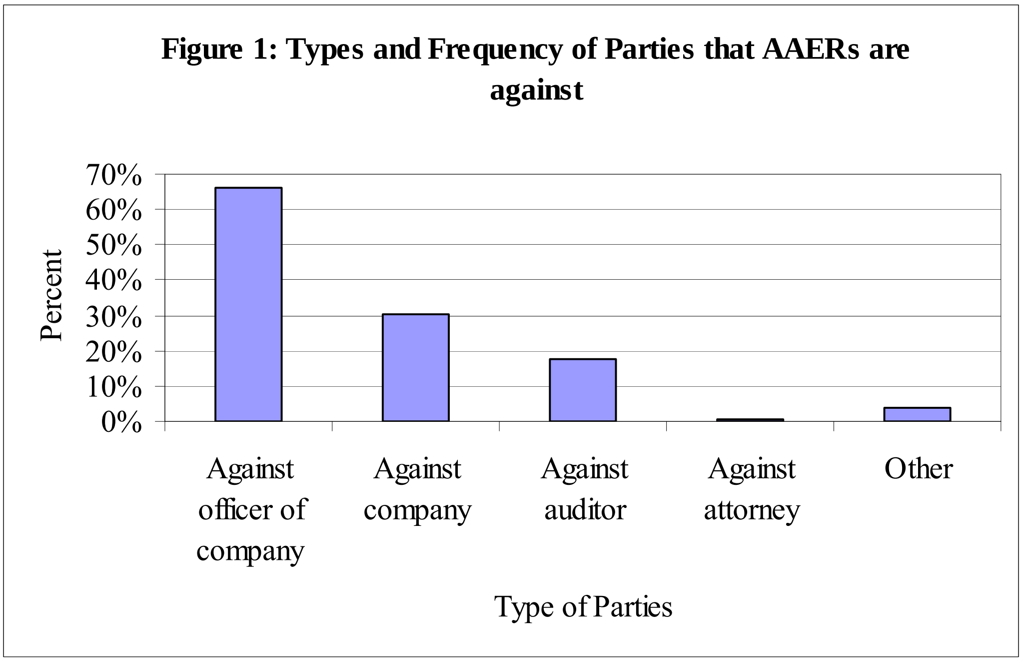
Panel A: **DETAIL FILE**	**detail.sas7bdat**
Variable Name	**Description**
coname	Name from AAER
cnum	6-digit Cusip
ticker	Compustat ticker
gvkey	Compustat Gvkey
permno	CRSP Permno
iticker	IBES Ticker
eticker	Exchange Ticker
explanation	Two sentence explanation of the violation
aaer#	SEC AAER numbers that relate to each firm (aaer1 through aaer24)
Indicator variables (file inclusion):	
annual	Equals 1 if the firm is in the Annual file, 0 otherwise
quarter	Equals 1 if the firm is in the Quarterly file, 0 otherwise
reason	Reason why firm is not included in Annual or Quarterly files
Indicator variables (exclusion from Annual or quarterly files):	
audit	Equals 1 if the AAER was brought against the auditor and there was no manipulation, 0 otherwise
bribes	Equals 1 if the AAER was for bribe charges, 0 otherwise
disclosure	Equals 1 if related to disclosure issue only and not earnings manipulation, 0 otherwise
nodates	Equals 1 if the time period of the financial manipulations cannot be determined from the AAER, 0 otherwise
other	Equals 1 if related to other issues not listed above, 0 otherwise
Indicator variable (Accounts affected):	
rev	Equals 1 if manipulation affected Revenues, 0 otherwise
rec	Equals 1 if manipulation affected Accounts Receivables, 0 otherwise
cogs	Equals 1 if manipulation affected Cost of Goods Sold, 0 otherwise
inv	Equals 1 if manipulation affected Inventory, 0 otherwise
res	Equals 1 if manipulation affected reserves accounts, 0 otherwise
debt	Equals 1 if manipulation affected bad debts, 0 otherwise
mkt_sec	Equals 1 if manipulation affected Marketable Securities, 0 otherwise
pay	Equals 1 if manipulation affected Accounts Payable, 0 otherwise
asset	Equals 1 if manipulation affected an asset account but could not be classified in an asset account above, 0 otherwise
liab	Equals 1 if manipulation affected liabilities, 0 otherwise
inc_exp_se	Equals 1 if manipulation could not be classified in an income, expense or equity account above, 0 otherwise
figure	Equals 1 if the actual amount of the manipulation can potentially be obtained from the AAER, 0 otherwise

Appendix 1: (continued)

Panel B: ANNUAL FILE	**ann.sas7bdat**
Variable Name	**Description**
coname	Name from AAER
cnum	6-digit Cusip
yeara	Compustat year
fyr	Compustat fiscal month end
date	Actual date collected from AAER (DD/MM/YYYY)
p_aaer	Primary AAER used to collect data from
understatement	Equals 1 if earnings/revenues were understated in the year, 0 otherwise

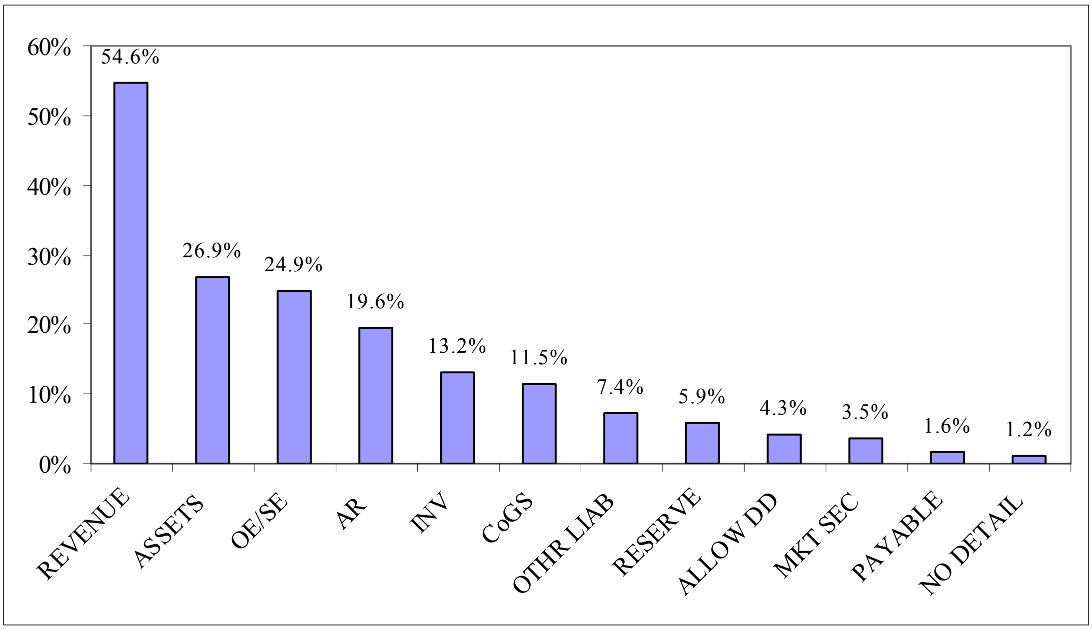
Panel C: QUARTERLY FILE	**qtr.sas7bdat**
Variable Name	**Description**
coname	Name from AAER
cnum	6-digit Cusip
yeara	Compustat year
fyr	Compustat fiscal month end
qtr	Quarter (1, 2, 3 or 4)
date	Actual date collected from AAER (DD/MM/YYYY)
p_aaer	Primary AAER used to collect data from
understatement	Equals 1 if earnings/revenues were understated in the quarter, 0 otherwise

Figure 1
Percent of the 2,191 AAERs that are against various parties.



Notes: One AAER can include multiple parties. The total number of parties is 2,592, so the percentages add up to 118 percent.

Figure 2
Type of manipulations mentioned in the AAERs for 680 firms included in either the quarterly or annual file.

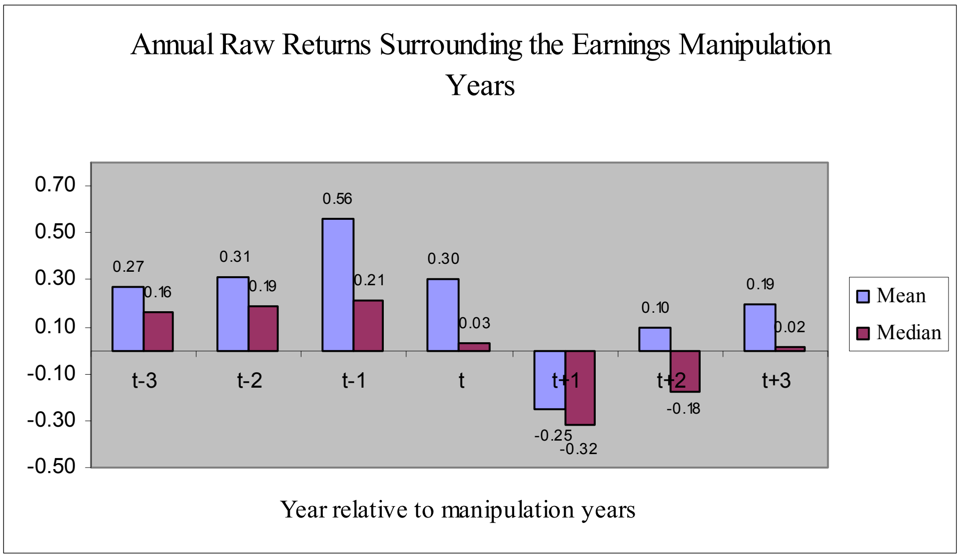


Notes:
There are 1143 manipulations mentioned in the AAERs for 680 firms so percentages add to more than 100 percent.

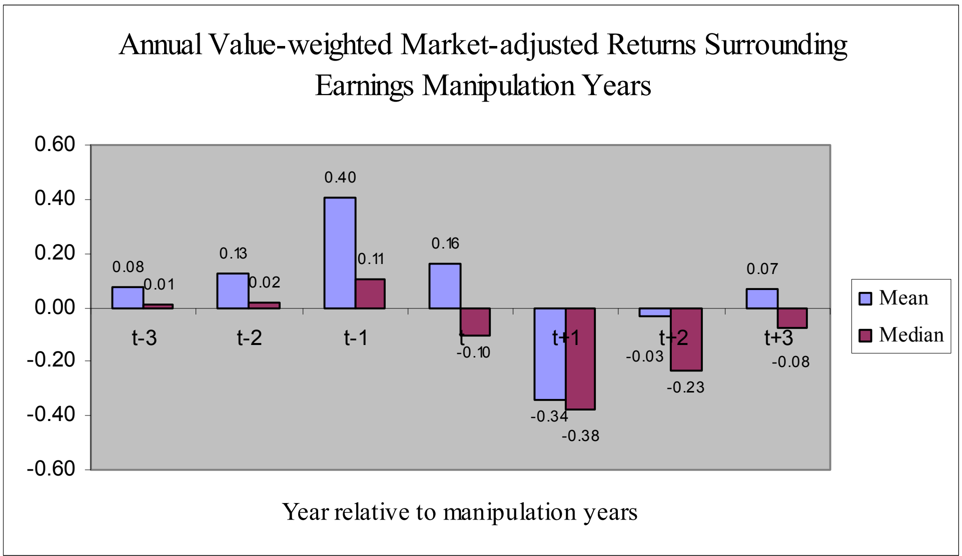
REVENUE	=	Manipulated revenue
ASSETS	=	Manipulated assets
OE/SE	=	Manipulation of other expense/shareholder equity account
AR	=	Manipulated accounts receivable
INV	=	Manipulated inventory
CoGS	=	Manipulated cost of goods sold
OTHR LIAB	=	Manipulated liabilities
RESERVE	=	Manipulated a reserve account
ALLOW DD	=	Manipulated allowance for bad debt
MKT SEC	=	Manipulated marketable securities
PAYABLE	=	Manipulated payables
NO DETAIL	=	No disclosure on how manipulation occurred

Figure 3
Stock price performance surrounding manipulation years

(a) Annual raw stock returns surrounding manipulation years.

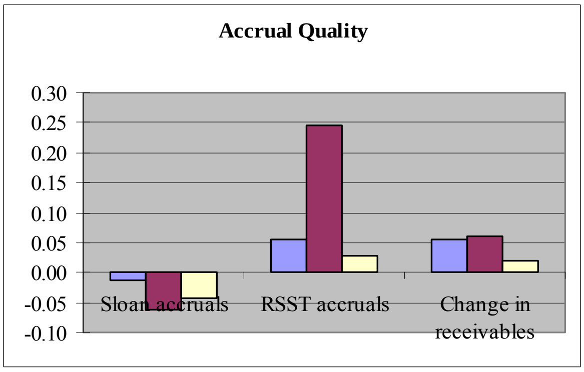


(b) Annual market-adjusted stock returns surrounding manipulation years.

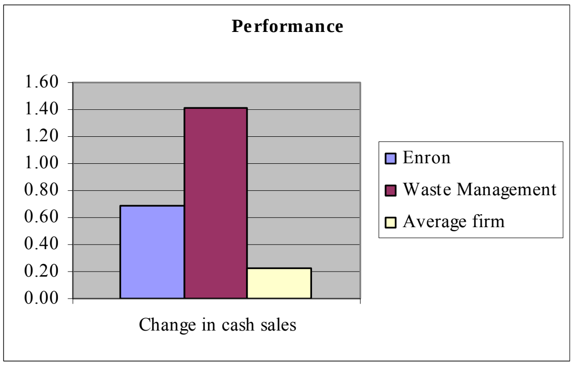


Note: For all firm-years with available returns data on CRSP. Returns include delisting returns. For year t-3 n=155, for year t-2 n=186, for year t-1 n=209, for year t n=505, for year t+1 n=213, for year t+2 n=182, for year t+3 n=146. Year t is the average return for all manipulation firms. Market-adjusted returns are calculated as the difference between annual raw returns and value-weighted market returns.

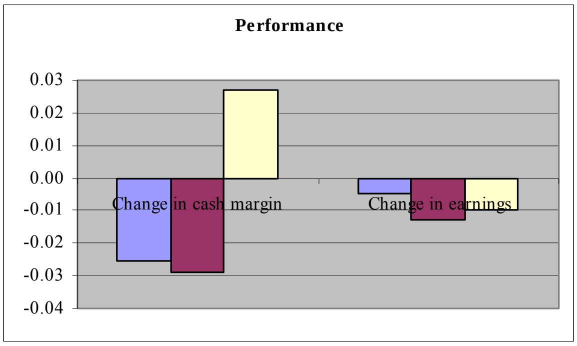
Figure 4
Comparison of various financial, nonfinancial, and market-related measures for *Enron* and *Waste Management* during their manipulation years to the average values of the variables for firms listed on Compustat.



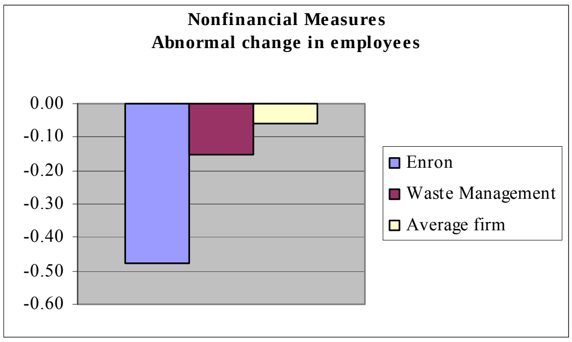
(a)



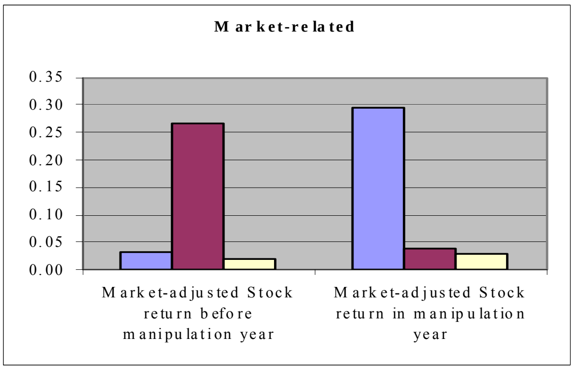
(b)



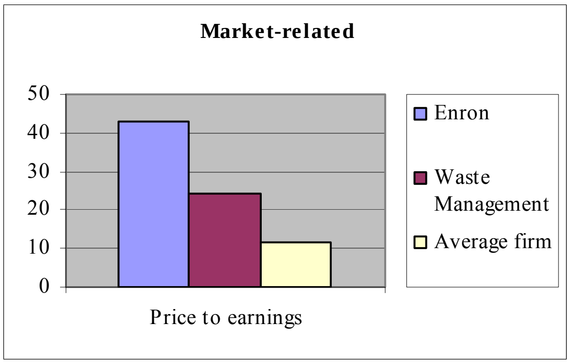
(c)



(d)

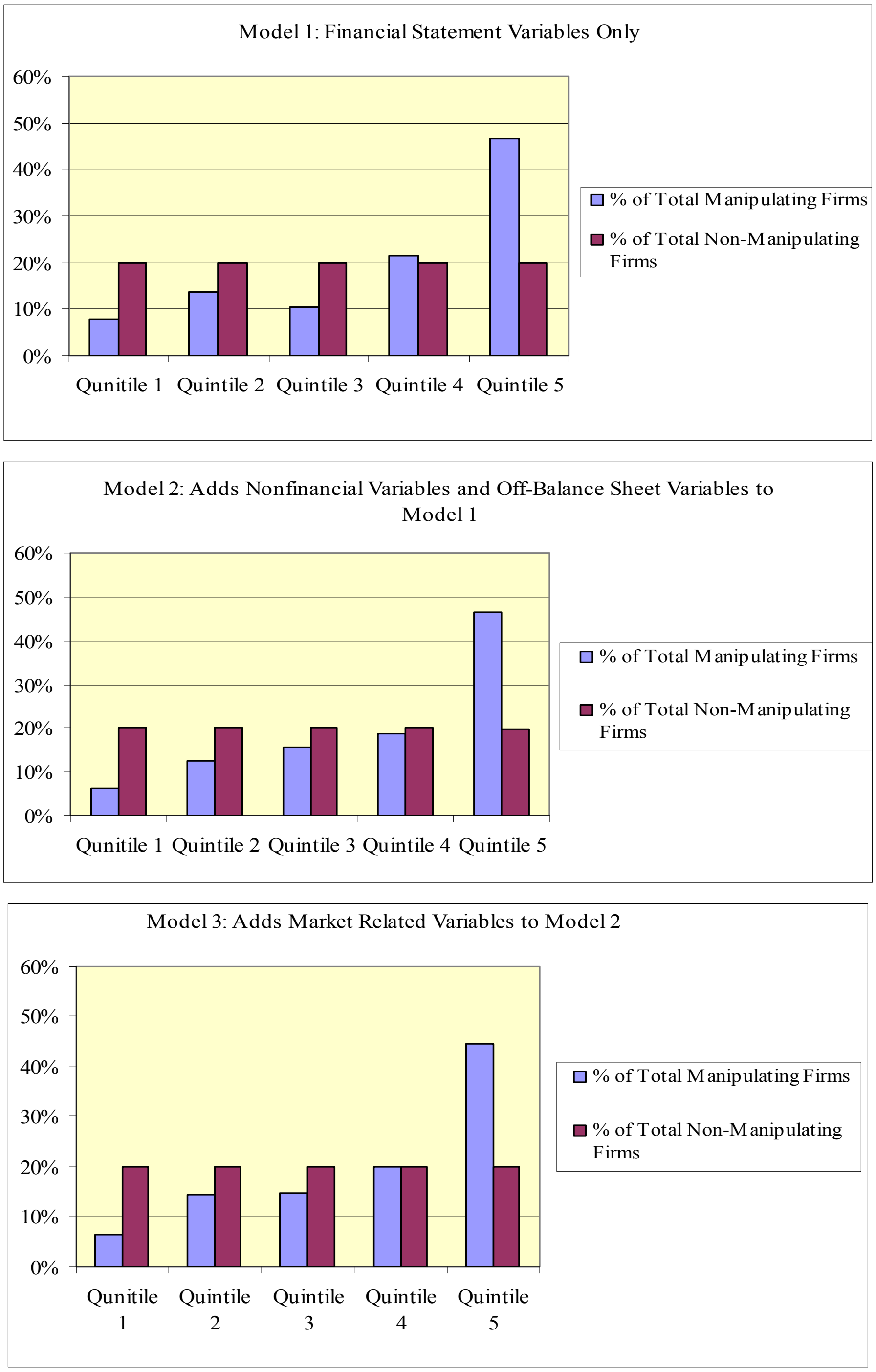


(e)



(f)

Figure 5
Percentage of manipulating firms in each quintile for the prediction models reported in Table 7



Note: Each prediction model is estimated using data from 1979 to 2002, and the F-Scores are then calculated for each firm-year. The firm-year observations are ranked based on the magnitude of F-Score into five quintiles (Quintile 5 has the highest predicted values of manipulations).

Figure 6:
Analysis of error rates for Model 1 reported in Table 7 for F-Scores ranging from 0 to 2.94.

Figure 6A

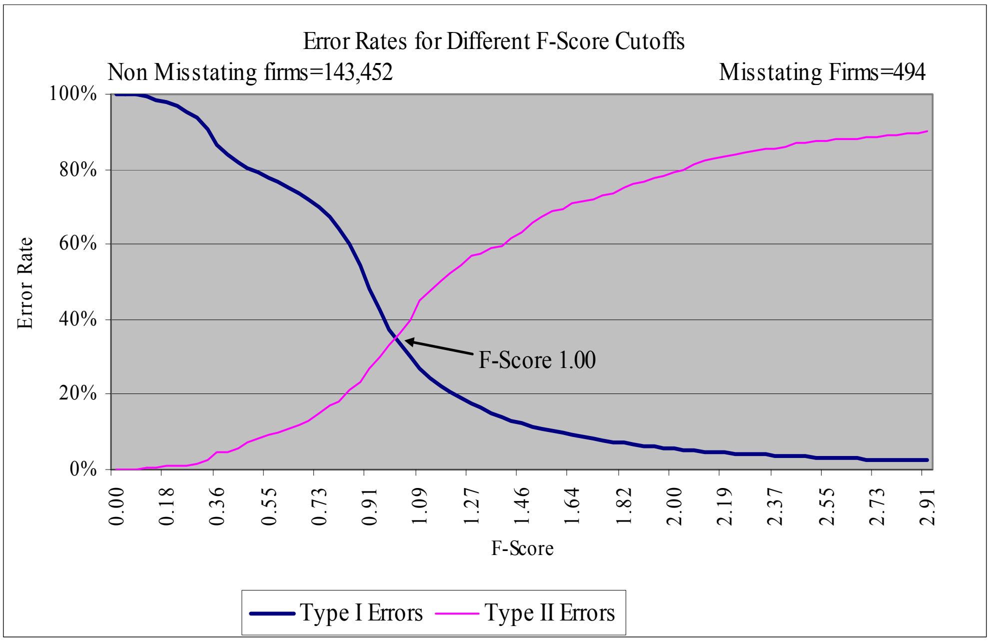
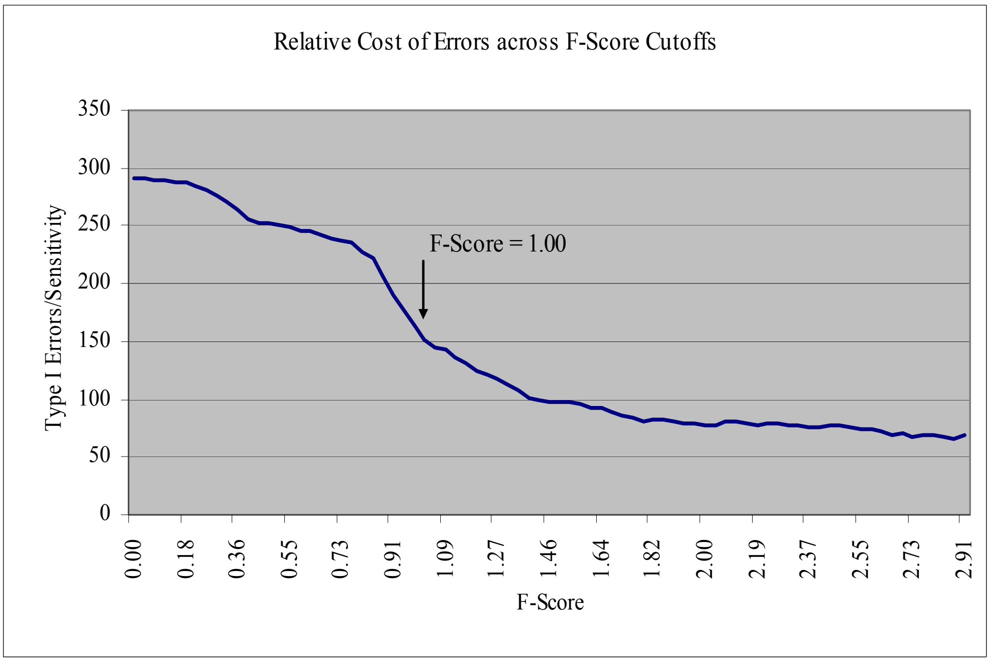
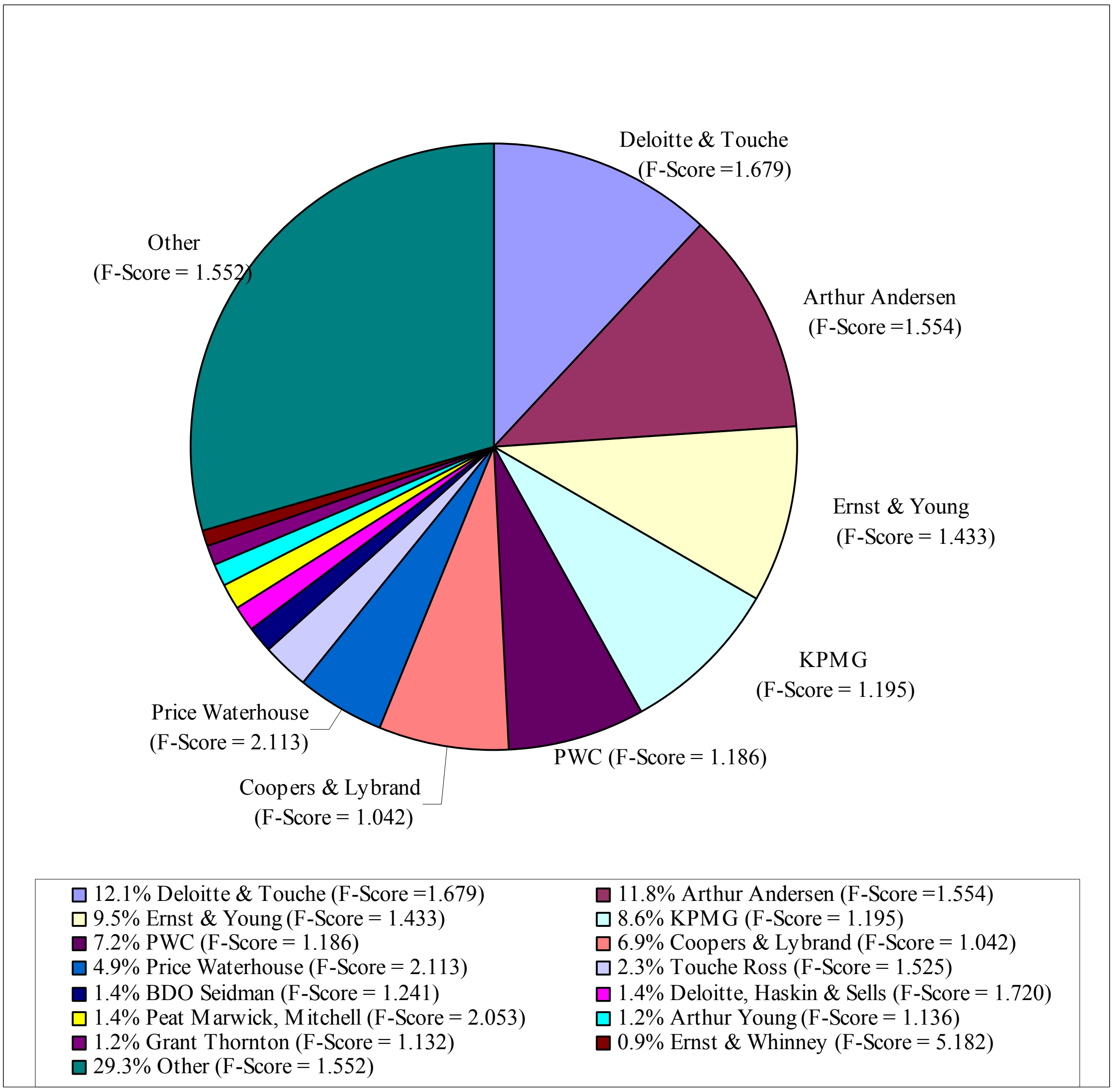


Figure 6B



Note: Figure 5A reports the Type I and Type II error rates for a given F-Score based on Model 1 in Table 7.
Type I errors = misclassified non-manipulating firm; Type II errors = misclassified manipulating firm.
Figure 5B reports the number of Type I errors divided by the sensitivity for each F-score cut-off. For example at an F-Score cut-off of 1.00, the total number of non-manipulating firms is 143,452 of which 92,391 have F-Scores less than 1.00, the remaining 51,061 firms (Type I error firms) have F-Scores greater than 1.00. At an F-Score cut-off of 1.00, 324 of the 494 manipulating firms have F-Scores greater than 1.00 (sensitivity firms or correctly classified manipulating firms is 324), while 170 (type II error firms) have F-scores less than 1.00. At this F-Score cut-off the relative cost ratio is 158 (51,061/324). If the cost of investigating a non-manipulating firm is less than 158 times the cost of missing a manipulating firm, then investigating all firms with F-Scores of 1.00 or higher would reduce overall costs to the audit firm.

Figure 7: Auditors for 348 firms manipulating at least one annual financial statement and median *F-Score* for each auditor.



Note: We report the auditor's name who signed off on the manipulated financial statements. We take the maximum *F-Score* during manipulation years for each firm. We then calculate the median *F-Score* for client firms for each auditor. Percentages are based on the total number of 348 manipulating firms. *F-Scores* are from Model 1 in Table 7.

Table 1
Sample description

Panel A: Sample selection of AAERs

Number of AAERs	**Number**
AAER No. 1- No. 2261 from May 1982 to June 2005	2261
Less: missing AAERs	(30)
Less: AAERs that do not involve specific company names	(40)
Total	**2191**

Note: Among 30 missing AAERs, eleven AAERs are intentionally omitted and nineteen AAERs are missing.

Panel B: Frequency of AAERs by year

AAER release date	**Number of AAERs**	**Percentage**
1982	2	0.1%
1983	16	0.7%
1984	28	1.3%
1985	35	1.6%
1986	39	1.8%
1987	51	2.3%
1988	37	1.7%
1989	38	1.7%
1990	35	1.6%
1991	61	2.8%
1992	79	3.6%
1993	76	3.5%
1994	120	5.5%
1995	107	4.9%
1996	121	5.5%
1997	134	6.1%
1998	85	3.9%
1999	111	5.1%
2000	142	6.5%
2001	125	5.7%
2002	209	9.5%
2003	237	10.8%
2004	209	9.5%
2005	94	4.3%
Total	2191	100.0%

Table 1 (continued)

Panel C: Frequency of the number of AAERs by firm

Number of AAERs for each firm	**Number of firms**	**Percent of firms**	**Total AAERs**
1	376	41.8%	376
2	236	26.3%	472
3	106	11.8%	318
4	67	7.5%	268
5	40	4.4%	200
6	33	3.7%	198
7	15	1.7%	105
8	9	1.0%	72
9	3	0.3%	27
10	6	0.7%	60
11	2	0.2%	22
12	2	0.2%	24
13	1	0.1%	13
15	1	0.1%	15
20	1	0.1%	20
24	1	0.1%	24
Total	**899**	**100.0%**	**2214**

Note: There are 23 (2214 less 2191) AAERs involving multiple companies.

Panel D: Number of distinct firms

Number of distinct companies mentioned in the AAERs	**Number**
AAER No. 1- No. 2261 from May 1982 to June 2005	**899**
Less: Enforcements which are unrelated to earnings manipulation (e.g., bribes, disclosure etc.) or firms with manipulations that cannot be linked to specific reporting periods	219
Earnings manipulation firms	**680**
Less: firms without CUSIP	168
Firms with at least one quarter of manipulated numbers	**512**
Less: firms with quarterly manipulations corrected by the end of the fiscal year	101
Firms with at least one annual manipulated number	**411**

Table 2
Frequency of manipulating firms by size, industry and calendar year
(both annual and quarterly manipulations)

Panel A: Frequency of the manipulating firms by firm size (market capitalization) deciles

Decile rank of market value of Compustat population	**Frequency**	**Percentage**
1	22	5.3%
2	32	7.6%
3	34	8.1%
4	43	10.3%
5	33	7.9%
6	51	**12.2%**
7	45	10.7%
8	55	**13.1%**
9	40	9.6%
10	64	**15.3%**
Total	419	100.0%

Panel B: Frequency of the manipulating firms by industry

Industry	**Manipulating Firms**	**Compustat Population**
Agriculture	0.2%	0.4%
Mining & Construction	2.7%	3.0%
Food & Tobacco	2.7%	2.1%
Textile and Apparel	2.9%	1.4%
Lumber, Furniture, & Printing	2.5%	3.2%
Chemicals	2.5%	2.1%
Refining & Extractive	1.0%	4.5%
Durable Manufacturers	18.6%	20.1%
Computers	**20.3%**	**11.9%**
Transportation	4.1%	5.9%
Utilities	1.9%	4.7%
Retail	**13.0%**	**9.7%**
Services	**12.4%**	**10.4%**
Banks & Insurance	12.2%	16.8%
Pharmaceuticals	2.9%	3.9%
Total	100.0%	100.0%

Note: There are 419 manipulating firms in the annual and quarterly files that have data to calculate market value and 483 manipulating firms that have SIC codes. Industries are based on the following SIC codes: Mining: 1000–1299, 1400–1999; Food: 2000–2199; Textiles: 2200–2799; Drugs: 2830–2839, 3840–3851; Chemicals: 2800–2829, 2840–2899; Refining: 1300–1399, 2900–2999; Rubber: 3000–3499; Industrial: 3500–3569, 3580–3659; Electrical: 3660–3669, 3680–3699; Miscellaneous Equipment: 3800–3839, 3852–3999; Computers: 3570–3579, 3670–3679, 7370–7379; Transportation: 4000–4899; Utilities: 4900–4999; Retail: 5000–5999; Banks: 6000–6999; Services: 7000–7369, 7380–8999.

Table 2 (continued)

Panel C: Distribution of manipulating firm-years

Year	Firm-years	Percentage
1971	1	0.13%
1972	1	0.13%
1973	1	0.13%
1974	2	0.26%
1975	2	0.26%
1976	1	0.13%
1977	1	0.13%
1978	4	0.52%
1979	9	1.18%
1980	13	1.70%
1981	16	2.09%
1982	28	3.66%
1983	21	2.75%
1984	24	3.14%
1985	16	2.09%
1986	28	3.66%
1987	24	3.14%
1988	22	2.88%
1989	37	4.84%
1990	27	3.53%
1991	40	5.24%
1992	43	5.63%
1993	40	5.24%
1994	33	4.32%
1995	36	4.71%
1996	35	4.58%
1997	40	5.24%
1998	49	6.41%
1999	61	7.98%
2000	56	7.33%
2001	38	4.97%
2002	13	1.70%
2003	2	0.26%
Total	**764**	**100.00%**

Note: This table is calculated based on the sample of 411 manipulating firms (as shown in Table 1 Panel D) with at least one manipulated annual financial statement.

Table 3: Variable definitions

	Variable	Abbreviation	Pred Sign*	Calculation
	Manipulation flag	*manipflag*	N/A	Indicator variable equal to 1 for manipulation firm-years and 0 otherwise
Accruals quality related variables	*Sloan accruals*	*Sloan_acc*	+	[[ΔCA(DATA 4) – Δcash and STI (DATA 1)]– [ΔCL (DATA 5) – ΔSTD (DATA 34) – ΔTP (DATA 71)] – Dep (DATA 14)]/Average total assets; following Sloan (1996)
	RSST accruals	*rsst_acc*	+	(ΔWC + ΔNCO + ΔFIN)/Average total assets, where WC = [CA (DATA 4) – CASH and STI (DATA 1)] – [CL (DATA 5) – STD (DATA 34)]; NCO = [Assets (DATA 6) – CA (DATA 4) - LTI (DATA 32)] – [total Liabilities (DATA 181) – CL (DATA 5) – LTD (DATA 9)]; FIN = [STI (DATA 193) + LTI (DATA 32)] – [LTD (DATA 9) + STD (DATA 34) + PRE Stock (DATA 130)]; following Richardson et al. (2006)
	Change in receivables	*ch_rec*	+	ΔReceivables (DATA 2)/Average total assets
	Change in inventory	*ch_inv*	+	ΔInventory (DATA 3)/Average total assets
	Modified Jones model discretionary accruals	*da*	+	The modified Jones model discretionary accrual is estimated cross-sectionally each year using all firm-year observations in the same two-digit SIC code: *Sloan Accruals* = α + *β(1/Beginning assets) +γ(ΔSales-ΔRec)/Beginning assets + ρΔPPE/Beginning assets + ε*. The residuals are used as the modified Jones model discretionary accruals.
	Performance-matched discretionary accruals	*dadif*	+	The difference between the modified Jones discretionary accruals for firm i in year t and the modified Jones discretionary accruals for the matched firm in year t, following Kothari et al (2005); each firm-year observation is matched with another firm from the same two-digit SIC code and year with the closest return on assets.
	Mean-adjusted absolute value of Dechow/Dichev residuals	*resid*	+	The following regression is estimated for each two-digit SIC industry: $\Delta WC = b_0 + b_1{*}CFO_{t-1} + b_2{*}CFO_t + b_3{*}CFO_{t+1} + \varepsilon$. The mean absolute value of the residual is calculated for each industry and is then subtracted from the absolute value of each firm's observed residual.
	Studentized Dechow/Dichev residuals	*sresid*	+	The scaled residuals are calculated as $\frac{\hat{e}_i}{\hat{\sigma}\sqrt{1-h_{ii}}}$ where h_{ii} is the *ii* element of the hat matrix, $X(X^TX)^{-1}X^T$ and $\hat{\sigma}=\sqrt{\frac{1}{n-m}\sum_{j=1}^{m}\hat{\varepsilon}^2_j}$ where m is the number of parameters in the model and n is the number of observations. SAS can output the scaled residuals using the following code: proc reg data= dataset; model Y=X; output data=temp student=studentresidual;

Performance variables	*Change in cash sales*	*ch_cs*	-	Percentage change in cash sales [Sales(DATA 12)-ΔAR(DATA 2)]
	Change in cash margin	*ch_cm*	-	Percentage change in cash margin [1-(CoGs(DATA 41)+(Change in inventory(DATA 3)))/(Sales(DATA 12)-(Change in AR(DATA 2)))]
	Change in earnings	*ch_earn*	?	Earnings_t (DATA 18)/Average total assets_t - Earnings_{t-1}/Average total assets_{t-1}
	Change in free cash flows	*ch_fcf*	-	ΔFree cash flows (income (DATA 18)-RSST accruals) /average total assets (DATA 6)
	Deferred tax expense	*tax*	+	Deferred tax expense for year t (DATA 50) / total assets for year t-1 (DATA 6)
Non-financial variables	*Abnormal change in employees*	*ch_emp*	-	Percentage change in the no. of employees (DATA 29) - percentage change in assets (DATA 6)
	Abnormal change in order backlog	*ch_backlog*	-	Percentage change in order backlog (DATA 98) - percentage change in sales(DATA 12)
Off-balance-sheet variables	*Existence of operating leases*	*leasedum*	+	An indicator variable coded 1 if future operating lease obligations are greater than zero
	Change in operating lease activity	*oplease*	+	The change in the present value of future non-cancelable operating lease obligations (DATA 96, 164, 165, 166 and 167) deflated by average assets following Ge (2006)
	Expected return on pension plan assets (%)	*pension*	+	Expected return on pension plan assets (DATA 336)
	Change in Expected return on pension plan assets (%)	*ch_pension*	+	ΔExpected return on pension plan assets (DATA 336 at t) - (DATA 336 at t-1)
Market Incentives	*Ex ante financing need*	*exfin*	+	An indicator variable coded 1 if [(CFO-past three year average capital expenditures)/Current assets]<-0.5
	Actual issuance	*issue*	+	An indicator variable coded 1 if the firm issued securities during the manipulation year (an indicator variable coded 1 if DATA 108>0 or DATA111>0)
	CFF	*cff*	+	Level of finance raised (DATA 313/average assets (DATA 6))
	Leverage	*leverage*	+	Long-term debt (DATA 9)/ Total assets (DATA 6)
	Market-adjusted Stock return	ret_t	+	Annual buy-and-hold return inclusive of delisting returns minus the annual buy-and-hold value-weighted market return
	Lagged market-adjusted Stock return	ret_{t-1}	+	Previous years annual buy-and-hold return inclusive of delisting returns minus the annual buy-and-hold value-weighted market return
	Book to market	*bm*	-	Equity (DATA 60)/ Market value (DATA 25 x DATA 199)
	Earnings to price	*ep*	-	Earnings (DATA 18)/ Market Value (DATA 25 x DATA 199)

*Predicted Sign shows the expected direction of the relations between various firm-year characteristics and manipulations.

Table 4 Panel A
Descriptive statistics of manipulation years versus non-manipulation years for AAER firms.

	Manipulation years			Non-manipulation years				Manipulate - Non-manipulate		
Variable	**N**	**Mean**	**Median**	**N**	**Mean**	**Median**	**Predicted sign**	**Diff. in Mean**	***One tailed P-value***	**t-statistics**
Accruals quality variables										
Sloan accruals	567	**0.018**	-0.004	4220	**-0.025**	-0.026	+	**0.043**	*0.001*	5.12
RSST accruals	574	**0.117**	0.061	4284	**0.040**	0.032	+	**0.077**	*0.001*	5.53
Change in receivables	586	**0.059**	0.031	4509	**0.025**	0.015	+	**0.035**	*0.001*	6.71
Change in inventory	574	**0.037**	0.005	4374	**0.020**	0.004	+	**0.018**	*0.001*	4.16
Modified Jones model discretionary accruals	539	**0.054**	0.022	3656	**0.002**	0.001	+	**0.051**	*0.001*	3.85
Performance-matched discretionary accruals	538	**0.051**	0.028	3656	**0.000**	0.001	+	**0.052**	*0.001*	3.75
Mean-adjusted absolute value of Dechow/Dichev residuals	342	**0.015**	-0.011	1972	**-0.002**	-0.022	+	**0.017**	*0.001*	3.12
Studentized Dechow/Dichev residuals	342	**0.389**	0.274	1972	**0.051**	0.034	+	**0.338**	*0.001*	5.39
Performance variables										
Change in cash sales	518	**0.445**	0.188	4114	**0.198**	0.101	-	**0.247**	*0.001*	5.91
Change in cash margin	502	**-0.021**	0.002	3884	**0.007**	0.001	-	**-0.029**	*0.080*	-1.41
Change in earnings	541	**-0.022**	-0.011	4362	**-0.008**	0.000	?	**-0.015**	*0.043*	-1.72
Change in free cash flows	523	**0.030**	0.006	3954	**0.013**	0.004	-	**0.016**	*0.154*	1.02
Deferred tax expense	603	**0.0015**	0.000	4647	**0.0023**	0.000	+	**-0.0008**	*0.443*	-0.77
Non-financial variables										
Abnormal change in employees	519	**-0.225**	-0.095	4195	**-0.095**	-0.055	-	**-0.130**	*0.001*	-3.54
Abnormal change in order backlog	149	**-0.029**	-0.069	1036	**0.051**	-0.028	-	**-0.080**	*0.115*	-1.21
Off-balance sheet variables										
Change in operating lease activity	592	**0.015**	0.002	4693	**0.008**	0.000	+	**0.008**	*0.001*	4.23
Existence of operating leases	621	**0.783**	1.000	5019	**0.624**	1.000	+	**0.159**	*0.001*	8.87
Expected return on pension plan assets (%)	80	**7.90**	9.00	639	**7.63**	8.75	+	**0.27**	*0.441*	0.77
Change in expected return on plan assets (%)	66	**-0.002**	0.00	549	**-0.094**	0	+	**0.092**	*0.051*	1.96
Market-related variables										
Ex ante financing need	424	**0.184**	0.000	2367	**0.106**	0.000	+	**0.078**	*0.001*	3.94
Actual issuance	600	**0.930**	1.000	4152	**0.884**	1.000	+	**0.046**	*0.001*	3.97
CFF	438	**0.187**	0.095	2489	**0.070**	0.003	+	**0.117**	*0.001*	7.73
Leverage	621	**0.196**	0.162	5019	**0.189**	0.140	+	**0.006**	*0.216*	0.79
Market-adjusted stock return	499	**0.163**	-0.102	3561	**0.072**	-0.022	+	**0.091**	*0.086*	1.34
Book to market	573	**0.578**	0.380	4150	**0.554**	0.476	-	**0.024**	*0.247*	0.69
Earnings to price	366	**0.069**	0.046	3229	**0.083**	0.065	-	**-0.014**	*0.001*	-3.47

All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers.

Table 4 Panel B
Descriptive statistics on manipulation years versus YEARS PRIOR TO MANIPULATION YEARS for AAER firms

	Manipulation years			Early years				Manipulate – Early Years		
Variable	**N**	**Mean**	**Median**	**N**	**Mean**	**Median**	**Predicted Sign**	**Diff. in Mean**	***One tailed P-value***	**t-statisics**
Accruals quality variables										
Sloan accruals	567	**0.018**	-0.004	2638	**0.000**	-0.014	+	**0.018**	*0.019*	2.08
RSST accruals	574	**0.117**	0.061	2673	**0.081**	0.045	+	**0.036**	*0.005*	2.59
Change in receivables	586	**0.059**	0.031	2842	**0.042**	0.024	+	**0.017**	*0.001*	3.32
Change in inventory	574	**0.037**	0.005	2742	**0.033**	0.012	+	**0.004**	*0.180*	0.92
Modified Jones model discretionary accruals	539	**0.054**	0.022	2197	**0.013**	0.003	+	**0.041**	*0.001*	3.07
Performance-matched discretionary accruals	538	**0.051**	0.028	2197	**0.012**	0.006	+	**0.039**	*0.002*	2.84
Mean-adjusted absolute value of Dechow/Dichev residuals	342	**0.015**	-0.011	783	**-0.003**	-0.019	+	**0.018**	*0.001*	3.07
Studentized Dechow/Dichev residuals	342	**0.389**	0.274	783	**0.202**	0.140	+	**0.187**	*0.003*	2.74
Performance variables										
Change in cash sales	518	**0.445**	0.188	2489	**0.245**	0.132	-	**0.200**	*0.001*	4.70
Change in cash margin	502	**-0.021**	0.002	2331	**0.006**	0.000	-	**-0.028**	*0.091*	-1.34
Change in earnings	541	**-0.022**	-0.011	2688	**0.001**	0.000	?	**-0.024**	*0.002*	-2.84
Change in free cash flows	523	**0.030**	0.006	2378	**0.010**	0.003	-	**0.020**	*0.110*	1.23
Deferred tax expense	603	**0.002**	0.000	2723	**0.004**	0.000	+	**-0.002**	*0.013*	-2.13
Non-financial variables										
Abnormal change in employees	519	**-0.225**	-0.095	2633	**-0.123**	-0.070	-	**-0.101**	*0.003*	-2.72
Abnormal change in order backlog	149	**-0.029**	-0.069	559	**0.047**	-0.039	-	**-0.075**	*0.143*	-1.07
Off-balance-sheet variables										
Change in operating lease activity	592	**0.015**	0.002	2991	**0.010**	0.000	+	**0.005**	*0.004*	2.64
Existence of operating leases	621	**0.783**	1.000	3312	**0.516**	1.000	+	**0.266**	*0.001*	14.24
Expected return on pension plan assets (%)	80	**7.90**	9.00	190	**8.59**	8.75	+	**-0.690**	*0.036*	2.10
Change in expected return on plan assets (%)	66	**-0.002**	0.00	148	**-0.054**	0	+	**0.052**	*0.268*	1.11
Market-related variables										
Ex ante financing need	424	**0.184**	0.000	960	**0.113**	0.000	+	**0.071**	*0.001*	3.34
Actual issuance	600	**0.930**	1.000	2505	**0.906**	1.000	+	**0.024**	*0.022*	2.03
CFF	438	**0.187**	0.095	1005	**0.116**	0.023	+	**0.071**	*0.001*	4.19
Leverage	621	**0.196**	0.162	3312	**0.188**	0.148	+	**0.007**	*0.186*	0.89
Market-adjusted stock return	499	**0.163**	-0.102	2297	**0.111**	0.012	+	**0.052**	*0.001*	7.44
Book to market	573	**0.578**	0.380	2522	**0.660**	0.513	-	**-0.082**	*0.019*	-2.35
Earnings to price	366	**0.069**	0.046	2315	**0.083**	0.068	-	**-0.014**	*0.001*	-3.45

All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers.

Table 5
Descriptive statistics on manipulation firm-years versus Compustat firm-years for the sample from 1979 to 2002.

	Manipulation firm-years			Compustat firm-years				Manipulate – Compustat		
Variable	**N**	**Mean**	**Median**	**N**	**Mean**	**Median**	**Predicted sign**	**Diff. in Mean**	***One-tailed P-value***	**t-statistics**
Accruals quality variables										
Sloan accruals	569	**0.016**	-0.002	169,183	**-0.042**	-0.037	+	**0.059**	*0.001*	7.23
RSST accruals	576	**0.119**	0.063	173,881	**0.029**	0.020	+	**0.091**	*0.001*	6.40
Change in receivables	587	**0.058**	0.032	177,043	**0.021**	0.009	+	**0.038**	*0.001*	7.77
Change in inventory	577	**0.037**	0.006	178,519	**0.010**	0.000	+	**0.027**	*0.001*	7.14
Modified Jones model discretionary accruals	516	**0.060**	0.028	150,557	**0.000**	0.001	+	**0.061**	*0.001*	4.68
Performance-matched discretionary accruals	343	**0.015**	-0.011	91,249	**0.000**	-0.020	+	**0.016**	*0.001*	3.01
Mean-adjusted absolute value of Dechow/Dichev residuals	343	**0.391**	0.277	91,249	**0.003**	0.014	+	**0.388**	*0.001*	6.57
Performance variables										
Change in cash sales	511	**0.468**	0.195	153,090	**0.211**	0.077	-	**0.256**	*0.001*	5.51
Change in cash margin	495	**-0.016**	0.001	146,477	**0.026**	0.003	-	**-0.042**	*0.080*	-1.41
Change in earnings	533	**-0.023**	-0.011	166,303	**-0.009**	-0.001	?	**-0.014**	*0.053*	-1.62
Change in free cash flows	516	**0.032**	0.007	156,410	**0.019**	0.004	-	**0.013**	*0.234*	0.73
Deferred tax expense	593	**0.0012**	0.000	183,579	**0.0011**	0.209	+	**0.0001**	*0.466*	0.08
Non-financial variables										
Abnormal change in employees	489	**-0.163**	-0.095	145,773	**-0.063**	-0.048	-	**-0.100**	*0.001*	-3.41
Abnormal change in order backlog	142	**-0.004**	-0.057	36,496	**0.087**	-0.041	-	**-0.091**	*0.102*	-1.28
Off-balance-sheet variables										
Change in operating lease activity	594	**0.015**	0.002	183,701	**0.008**	0.000	+	**0.007**	*0.001*	4.07
Existence of operating leases	594	**0.800**	1.000	183,701	**0.658**	1.000	+	**0.142**	*0.001*	8.60
Expected return on pension plan assets	78	**7.88**	9.00	26,263	**7.17**	8.50	+	**0.71**	*0.070*	1.81
Change in expected return on plan assets	66	**-0.002**	0.0	22,242	**-0.040**	0.00	+	**0.04**	*0.352*	0.92
Market-related variables										
Ex ante finance need	433	**0.189**	0.000	110,828	**0.163**	0.000	+	**0.026**	*0.084*	1.38
Actual issuance	577	**0.931**	1.000	171,118	**0.816**	1.000	+	**0.115**	*0.001*	10.80
CFF	448	**0.193**	0.094	116,006	**0.134**	0.006	+	**0.059**	*0.001*	3.82
Leverage	594	**0.198**	0.155	183,559	**0.191**	0.128	+	**0.006**	*0.222*	0.77
Mkt-adj return	501	**0.162**	-0.102	168,074	**0.046**	-0.059	+	**0.116**	*0.041*	1.74
Lagged mkt-adj return	433	**0.273**	0.019	153,903	**0.055**	-0.055	+	**0.218**	*0.002*	3.08
Book-to-market	565	**0.541**	0.374	158,312	**0.663**	0.573	-	**-0.122**	*0.001*	-4.01
Earnings-to-price	358	**0.068**	0.046	104,646	**0.087**	0.069	-	**-0.019**	*0.001*	-4.89

All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers. Note that even though we restrict our sample to 1979-2002 in the cross-sectional analysis, for some variables, the number of observations appears to be slight larger in Table 5. This is because in the time-series analysis, we eliminate those observations with available data only in either manipulation period or non-manipulation period to make the comparison meaningful.

Table 6: Correlation matrix between variables (Spearman above diagonal, Pearson below diagonal)

Variable		Accrual Quality							Performance			Non financial			Market-related				
	manipflag	***Sloan_acc***	***rsst_acc***	***ch_rec***	***ch_inv***	***da***	***resid***	***sresid***	***ch_cs***	***ch_cm***	***ch_e***	***ch_emp***	***bcklog***	***oplease***	***issue***	***cff***	***ret***$_t$	***ret***$_{t-1}$	***bm***
manipflag		0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.00	0.00	-0.01	-0.01	0.01	0.02	0.01	0.01	0.01	-0.01
Sloan_acc	0.02		0.45	0.49	0.48	0.72	-0.14	0.52	0.03	-0.11	0.23	-0.11	-0.01	0.08	0.07	-0.01	0.09	0.09	0.10
rsst_acc	0.02	0.40		0.25	0.22	0.33	-0.14	0.34	0.11	-0.02	0.35	-0.25	-0.03	0.12	0.09	0.07	0.15	0.15	0.11
ch_rec	0.02	0.50	0.32		0.29	0.29	0.00	0.29	0.07	-0.13	0.14	-0.16	0.03	0.17	0.11	0.15	0.12	0.09	-0.01
ch_inv	0.01	0.46	0.31	0.29		0.25	-0.03	0.31	0.17	-0.14	0.07	-0.09	0.05	0.18	0.10	0.11	0.08	0.10	0.02
da	0.01	0.75	0.30	0.35	0.31		-0.05	0.38	-0.03	-0.07	0.15	-0.07	0.04	0.03	0.03	-0.02	0.05	0.04	0.05
resid	0.01	0.00	-0.04	-0.01	-0.02	0.01		-0.18	0.04	0.02	-0.07	-0.02	0.03	-0.02	-0.08	0.30	0.00	0.00	-0.20
sresid	0.03	0.52	0.34	0.33	0.32	0.44	0.05		0.09	0.00	0.24	-0.07	-0.05	0.08	0.08	0.01	0.06	0.06	0.07
ch_cs	0.02	0.14	0.28	0.21	0.33	-0.01	-0.03	0.18		0.42	0.10	-0.02	-0.17	0.13	0.08	0.18	0.04	0.09	-0.04
ch_cm	0.00	-0.25	-0.09	-0.33	-0.32	-0.24	0.00	-0.08	0.20		0.10	0.02	-0.11	-0.03	0.00	0.00	-0.02	-0.02	-0.01
ch_earn	-0.01	0.14	0.25	0.13	0.07	0.08	-0.02	0.19	0.17	0.08		-0.18	-0.05	0.01	0.00	-0.05	0.17	-0.01	0.03
ch_emp	-0.01	-0.13	-0.18	-0.18	-0.11	-0.09	-0.01	-0.07	-0.03	0.04	-0.08		0.04	-0.01	-0.06	-0.33	-0.11	-0.08	0.00
ch_backlog	-0.01	-0.03	-0.06	0.02	0.04	0.04	-0.02	-0.09	-0.24	-0.16	-0.05	0.07		0.01	-0.02	0.07	0.07	-0.06	-0.02
oplease	0.02	0.10	0.17	0.18	0.19	0.05	-0.03	0.09	0.21	-0.07	-0.02	-0.01	-0.01		0.07	0.15	0.03	0.06	0.02
leasedum	0.02	-0.04	0.05	-0.01	0.05	0.00	-0.09	0.03	0.07	0.02	-0.02	0.03	-0.01	0.19	0.17	0.01	0.02	0.02	-0.09
issue	0.02	0.06	0.11	0.12	0.11	0.04	-0.06	0.07	0.14	-0.01	-0.01	-0.08	-0.01	0.10		0.14	0.04	0.07	-0.09
cff	0.03	0.17	0.11	0.25	0.19	0.10	0.09	0.18	0.18	-0.09	-0.09	-0.24	-0.01	0.13	0.25		0.08	0.10	-0.17
leverage	0.00	-0.05	-0.10	-0.05	0.00	-0.01	-0.02	-0.03	0.02	-0.01	-0.01	-0.02	0.01	0.03	0.21	-0.04	-0.04	-0.04	-0.09
ret$_t$	0.00	0.12	0.24	0.19	0.12	0.07	-0.03	0.07	0.13	-0.05	0.34	-0.10	0.06	0.05	0.04	-0.05		-0.01	-0.11
ret$_{t-1}$	0.01	0.16	0.25	0.18	0.17	0.10	-0.03	0.09	0.18	-0.04	0.01	-0.11	-0.05	0.10	0.09	0.06	0.09		-0.08
bm	-0.02	0.03	-0.08	-0.08	-0.03	0.00	-0.11	-0.04	-0.12	-0.02	-0.04	-0.01	0.02	-0.04	-0.15	-0.25	-0.18	-0.13	
ep	-0.02	0.03	-0.04	-0.05	0.03	-0.05	0.09	0.04	-0.04	0.02	0.12	-0.04	-0.04	-0.06	-0.11	-0.21	-0.03	0.01	0.53

All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers.

Table 7 Panel A: Logistic regressions (dependent variable is equal to one if the firm-year manipulated earnings, zero otherwise) examining the determinants of manipulations.

Variable	**Model 1** **Financial Statement Variables**	**Model 2** **Add Off-balance sheet and Non-financial Variables**	**Model 3** **Add Stock Market-based Variables**
	Coefficient Estimate (Wald Chi-square) (P-value)	Coefficient Estimate (Wald Chi-square) (P-value)	Coefficient Estimate (Wald Chi-square) (P-value)
Intercept	**-6.753**	**-7.181**	**-6.591**
	1490.6	1098.6	717.1
	0.001	0.001	0.001
RSST accruals	**0.773**	**0.683**	**0.931**
	22.0	13.3	13.5
	0.001	0.001	0.001
Change in receivables	**3.201**	**3.041**	**2.181**
	44.7	35.5	11.5
	0.001	0.001	0.001
Change in inventory	**2.465**	**2.722**	**2.767**
	16.1	18.5	13.3
	0.001	0.001	0.001
Change in cash sales	**0.108**	**0.085**	**0.079**
	7.7	3.6	1.7
	0.006	0.059	0.195
Change in earnings	**-0.995**	**-1.167**	**-1.412**
	22.0	26.3	21.7
	0.001	0.001	0.001
Actual issuance	**0.938**	**0.838**	**0.480**
	26.8	18.1	5.0
	0.001	0.001	0.025
Abnormal change in employees		**-0.215**	**-0.242**
		5.8	4.5
		0.016	0.034
Existence of operating leases		**0.613**	**0.516**
		19.2	10.9
		0.001	0.001
Book to market			**-0.128**
			3.3
			0.068
Lagged market-adjusted stock return			**0.063**
			4.1
			0.044
Manipulating Firm-years:	494	451	362
Non-manipulating Firm-years:	143,452	130,312	95,170

Table 7 Panel B: Examination of the detection rates of manipulating and non-manipulating firms for each Model reported in Panel A.

	Model 1			Model 2			Model 3		
	N	**Minimum *F-Score***	**% of Total**	**N**	**Minimum *F-Score***	**% of Total**	**N**	**Minimum *F-Score***	**% of Total**
Quintile 1									
Manipulate Firms	38	0.058	7.69%	28	0.087	6.21%	23	0.079	6.35%
No-Manipulate Firms	28,751	0.000	20.04%	26,124	0.000	20.05%	19,083	0.000	20.05%
Quintile 2									
Manipulate Firms	68	0.495	13.77%	57	0.464	12.64%	52	0.554	14.36%
No-Manipulate Firms	28,721	0.495	20.02%	26,096	0.464	20.03%	19,055	0.554	20.02%
Quintile 3									
Manipulate Firms	52	0.845	10.53%	71	0.725	15.74%	54	0.765	14.92%
No-Manipulate Firms	28,738	0.845	20.03%	26,082	0.725	20.02%	19,052	0.765	20.02%
Quintile 4									
Manipulate Firms	**106**	**0.962**	**21.46%**	**85**	**0.986**	**18.85%**	**72**	**0.976**	**19.89%**
No-Manipulate Firms	28,683	0.962	19.99%	26,068	0.986	20.00%	19,035	0.976	20.00%
Quintile 5									
Manipulate Firms	**230**	**1.224**	**46.56%**	**210**	**1.276**	**46.56%**	**161**	**1.240**	**44.48%**
No-Manipulate Firms	28,559	1.224	19.91%	25,942	1.276	19.91%	18,945	1.240	19.91%

Note: All observations are ranked based on their predicted probabilities (*F-Scores*) and sorted into Quintiles. Minimum *F-Score* is the minimum scaled predicted probability based on estimates in Panel A to enter each quintile.

Panel C: *F-Score* cut-off set at 1.00

	Model 1 Predicted			Model 2 Predicted			Model 3 Predicted		
Observed	Manip.	No-Manip.		Manip.	No-Manip.		Manip.	No-Manip.	
Manipulate	**324**	170	494	**293**	158	451	**228**	134	362
No-Manipulate	51,061	**92,391**	143,452	50,954	**79,358**	130,312	35,502	**59,668**	95,170
	51,385	92,561	**143,946**	51,247	79,516	**130,763**	35,730	59,802	**95,532**
Manipulate	**65.59%**	34.41%	0.34%	**64.97%**	35.03%	0.34%	**62.98%**	37.02%	0.38%
No-Manipulate	35.59%	**64.41%**	99.66%	39.10%	**60.90%**	99.66%	37.30%	**62.70%**	99.62%

	Model 1		Model 2	Model 3
Correct classification	**64.41%**	(1)	**60.91%**	**62.70%**
Sensitivity	**65.59%**	(2)	**64.97%**	**62.98%**
Type I errors	35.47%	(3)	38.97%	37.16%
Type II errors	34.41%	(4)	35.03%	37.02%

Notes:
(1) Correct classification is calculated as (324+92,391/143,946)
(2) Sensitivity is calculated as (324/494)
(3) *Type I* errors are calculated as (51,061/143,452)
(4) *Type II* errors are calculated as (170/494)
All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers.

Table 8 Panel A: Logistic regressions (dependent variable is equal to one if the firm-year manipulated earnings, zero otherwise) examining the determinants of manipulations estimated using the time period 1979-1998.

Variable	**Model 1** **Financial Statement Variables**	**Model 2** **Add Off-balance sheet and Non-financial Variables**	**Model 3** **Add Stock Market-based Variables**
	Coefficient Estimate (Wald Chi-square) (P-value)	Coefficient Estimate (Wald Chi-square) (P-value)	Coefficient Estimate (Wald Chi-square) (P-value)
Intercept	**-6.684**	**-7.066**	**-6.502**
	1324.4	949.2	597.4
	0.001	0.001	0.001
RSST accruals	**0.891**	**0.879**	**1.035**
	17.5	13.2	10.4
	0.001	0.001	0.001
Change in receivables	**3.127**	**2.850**	**1.626**
	30.9	22.1	4.5
	0.001	0.001	0.034
Change in inventory	**2.821**	**3.271**	**3.015**
	16.4	20.7	11.9
	0.001	0.001	0.001
Change in cash sales	**0.097**	**0.064**	**0.092**
	3.9	1.2	1.7
	0.048	0.272	0.198
Change in earnings	**-0.984**	**-1.183**	**-1.275**
	12.2	14.9	10.2
	0.001	0.001	0.001
Actual issuance	**0.744**	**0.672**	**0.343**
	15.0	10.2	2.2
	0.001	0.001	0.138
Abnormal change in employees		**-0.136**	**-0.214**
		1.6	2.5
		0.208	0.116
Existence of operating leases		**0.536**	**0.464**
		12.1	7.1
		0.001	0.008
Book to Market			**-0.163**
			3.320
			0.069
Lagged market-adjusted stock return			**0.083**
			6.5
			0.011
Manipulating Firm-years:	361	327	260
Non-manipulating Firm-years:	114,426	104,160	76,144

Table 8 Panel B: Examination of detection rates for each model reported in Panel A using a hold-out sample for the time period 1999-2002.

	Model 1			**Model 2**			**Model 3**		
	N	**Minimum *F-Score***	**% of Total**	**N**	**Minimum *F-Score***	**% of Total**	**N**	**Minimum *F-Score***	**% of Total**
Quintile 1									
Manipulate Firms	8	0.159	6.02%	7	0.159	5.65%	9	0.500	8.82%
No-Manipulate Firms	5,823	0.000	20.06%	5,248	0.000	20.07%	3,816	0.000	20.06%
Quintile 2									
Manipulate Firms	17	0.541	12.78%	14	0.510	11.29%	13	0.558	12.75%
No-Manipulate Firms	5,815	0.509	20.03%	5,241	0.510	20.04%	3,813	0.558	20.04%
Quintile 3									
Manipulate Firms	17	0.859	12.78%	17	0.764	13.71%	14	0.764	13.73%
No-Manipulate Firms	5,815	0.779	20.03%	5,239	0.764	20.03%	3,812	0.764	20.04%
Quintile 4									
Manipulate Firms	**28**	**0.890**	**21.05%**	**23**	**0.956**	**18.55%**	**20**	**0.970**	**19.61%**
No-Manipulate Firms	5,804	0.890	20.00%	5,232	0.956	20.01%	3,806	0.940	20.00%
Quintile 5									
Manipulate Firms	**63**	**1.145**	**47.37%**	**63**	**1.210**	**50.81%**	**46**	**1.205**	**45.10%**
No-Manipulate Firms	5,769	1.145	19.88%	5,192	1.179	19.85%	3,779	1.205	19.86%

Note: All observations are ranked based on their *F-Scores* and sorted into Quintiles. Minimum *F-Score* is the minimum scaled predicted probability based on estimates in Panel A to enter each quintile.

Panel C: *F-Score* cut-off set to 1.00

	Model 1 Predicted			**Model 2 Predicted**			**Model 3 Predicted**		
Observed	Manip.	No-Manip.		Manip.	No-Manip.		Manip.	No-Manip.	
Manipulate	**77**	56	133	**80**	44	124	62	40	102
No-Manipulate	8,246	**20,780**	29,026	8,800	**17,352**	26,152	6,554	12,472	19,026
	8,323	20,836	**29,159**	8,800	17,396	**26,276**	6,113	13,011	**19,128**
Manipulate	**57.89%**	42.11%	0.46%	**64.52%**	35.48%	0.47%	**60.78%**	39.22%	0.53%
No-Manipulate	28.41%	**71.59%**	99.54%	33.65%	**66.35%**	99.53%	34.45%	**65.55%**	99.47%

	Model 1		Model 2	Model 3
Correct classification	**71.53%**	(1)	**66.34%**	**65.53%**
Sensitivity	**57.89%**	(2)	**64.52%**	**60.78%**
Type I errors	28.28%	(3)	33.49%	34.26%
Type II errors	42.11%	(4)	35.48%	39.22%

Notes:
(1) Correct classification is calculated as (77+ 20,780/29,159)
(2) Sensitivity is calculated as (77/133)
(3) *Type I* errors are calculated as (8,323/29,026)
(4) *Type II* errors are calculated as (56/133)
All variables are defined in Table 3. Each of the continuous variables (except stock return variables) is winsorized at 1% and 99% to mitigate outliers.