

# **The Detection of Earnings Manipulation**

Messod D. Beneish\*

June 1999

Comments Welcome

\* Associate Professor, Indiana University, Kelley School of Business, 1309 E. 10th Street  
Bloomington, Indiana 47405, dbeneish@indiana.edu, 812 855-2628. I have benefited from the  
comments of Vic Bernard, Jack Ciesielski, Linda DeAngelo, Martin Fridson, Cam Harvey, David  
Hsieh, Charles Lee, Eric Press, Bob Whaley, Mark Zmijewski, workshop participants at Duke,  
Maryland, Michigan, and Université du Québec à Montréal. I am indebted to David Hsieh for his  
generous econometric advice and the use of his estimation subroutines. I thank Julie Azoulay,  
Pablo Cisilino and Melissa McFadden for expert assistance.

## **Abstract**

### **The Detection of Earnings Manipulation**

The paper profiles a sample of earnings manipulators, identifies their distinguishing characteristics, and estimates a model for detecting manipulation. The model's variables are designed to capture either the effects of manipulation or preconditions that may prompt firms to engage in such activity. The results suggest a systematic relation between the probability of manipulation and financial statement variables. The evidence is consistent with accounting data being useful in detecting manipulation and assessing the reliability of accounting earnings.

In holdout sample tests, the model identifies approximately half of the companies involved in earnings manipulation prior to public discovery. Because companies discovered manipulating earnings see their stocks plummet in value, the model can be a useful screening device for investing professionals. While the model is easily implemented-- the data can be extracted from an annual report--, the screening results require further investigation to determine whether the distortions in financial statement numbers result from earnings manipulation or have another structural root.

## **Introduction**

The extent to which earnings are manipulated has long been a question of interest to analysts, regulators, researchers, and other investment professionals. While the SEC's recent commitment to vigorously investigate earnings manipulation (see Levitt (1998)) has sparked renewed interest in the area, there has been little in the academic and professional literature on the detection earnings manipulation.<sup>1</sup>

This paper presents a model to distinguish manipulated from non-manipulated reporting. I define earnings manipulation as an instance where management violates Generally Accepted Accounting Principles (GAAP) in order to beneficially represent the firm's financial performance. I use financial statement data to construct variables that seek to capture the effects of manipulation and preconditions that may prompt firms to engage in such activity. Since manipulation typically consists of an artificial inflation of revenues or deflation of expenses, I find that variables that take into account the simultaneous bloating in asset accounts have predictive content. I also find that sales growth has discriminatory power since the primary characteristic of sample manipulators is that they have high growth prior to periods during which manipulation is in force.

I conduct tests using a sample of 74 firms that manipulate earnings and all COMPUSTAT firms matched by two-digit SIC for which data are available in the period 1982-1992. I find that sample manipulators typically overstate earnings by recording fictitious, unearned, or uncertain revenue, recording fictitious inventory, or improperly capitalizing costs. The context of earnings

---

<sup>1</sup> A model for detecting earnings manipulation is in Beneish (1997). The model in that paper differs from the model in this study in three ways: (i) the model is estimated with 64 treatment firms vs 74 firms in the present study, (ii) control firms are COMPUSTAT firms with the largest unexpected accruals vs. COMPUSTAT firms in the same industry in the present study, and (iii) the set of explanators differ across studies, with the present study presenting a more parsimonious model.

manipulation is an annual report or a 10-K for about two-thirds of the sample and a security offering prospectus (initial, secondary, debt offering) for the remaining third. Sample manipulators are relatively young, growth firms as such characteristics make it more likely that firms come under the scrutiny of regulators (see Beneish (1999)).

I estimate a model for detecting earnings manipulation using sample manipulators and industry-matched firms in the period 1982-1988 and evaluate the model's performance on a holdout sample in the period 1989-1992. The model distinguishes manipulators from non-manipulators, and has pseudo- $R^2$ s of 30.6% and 37.1% for two different estimation methods. The evidence indicates that the probability of manipulation increases with: (i) unusual increases in receivables, (ii) deteriorating gross margins, (iii) decreasing asset quality (as defined later), (iv) sales growth, and (v) increasing accruals. I show that the model discriminates manipulators from non-manipulators in the holdout sample. The results are robust to different estimates of the prior probability of earnings manipulation, several specifications of the model and various transformations of the explanatory variables. The results are also insensitive to the choice of estimation and holdout samples.

The evidence needs to be interpreted in light of possible sample selection biases. The estimation addresses the bias arising from oversampling manipulators, but it is based on a sample of discovered manipulators. It is possible that there are successful, unidentified manipulators, and the results need to be interpreted assuming that sample manipulators represents a substantial portion of the manipulators in the population. Given this caveat, the evidence of a systematic relation between the likelihood of manipulation and financial statement variables suggests that accounting data are useful in detecting manipulation and assessing the reliability of accounting

earnings.

### **Sample**

The sample of earnings manipulators is obtained either from firms subject to the SEC's accounting enforcement actions or from a news media search. I identified firms subject to accounting enforcement actions by the SEC using Accounting and Auditing Enforcement Releases (AAERs) numbers 132 to 502 issued from 1987 to 1993. Of 363 AAERs examined (#372 to #379 were not assigned by the SEC), I eliminated 80 AAERs relating to financial institutions, 15 relating to auditing actions against independent CPAs, nine relating to 10-Q violations that were resolved in annual filings, and 156 relating to firms for which no financial statement data are available on either COMPUSTAT, S&P Corporate Text or 10-K microfiche. The SEC search yields a final sample of 103 AAERs relating to 49 firms that violate GAAP.

I also conducted an extensive news media search on LEXIS/NEXIS in the period January 1987 to April 1993.<sup>2</sup> The search identified 80 firms mentioned in articles discussing earnings manipulation. In addition to eight firms that are identified by the SEC search, I eliminate ten firms for which no financial statement data are available on either COMPUSTAT, S&P Corporate Text or 10-K microfiche, five financial institutions, and 17 firms mentioned in articles with no discussion of an accounting or disclosure problem.<sup>3</sup>

---

<sup>2</sup> Specifically, the search encompassed the following data bases in LEXIS/NEXIS: *Barron's*, *Business Week*, *Business Wire*, *Corporate Cash Flow*, *Disclosure Online*, *Forbes*, *Fortune*, *Institutional Investor*, *Investor Daily*, *Money*, *The Courier Journal*, *The New York Times*, *The Wall Street Journal*, *The Washington Post*, and *The Reuter Business Report*. I used the following keywords: "earnings management;" "earnings manipulation;" "cooking the books;" "financial statements or reports" with adjectives such as deceptive, false, fraudulent, misleading, illusive, inappropriate, misstated, and spurious; and "inflated or overstated" with either profits, earnings, or income.

<sup>3</sup> For example, in an article on the manipulation of earnings at Chambers Development, Flynn and Zellner (1992) discuss other firms in the waste management industry such as Sanifill Inc., and Waste Management, without referring to any accounting measurement or disclosure problem.

I require *ex post* evidence of manipulation for the remaining 40 firms. That is, I require that firms restate earnings to comply with GAAP at the request of auditors or following the conclusion of an internal investigation. This requirement makes sample entry consistent with the SEC search in the sense that a restatement is usually the outcome of successful SEC investigations (in addition to a permanent injunction from future violations of security laws). This criterion eliminates fifteen firms and is imposed to exclude firms that manage earnings within GAAP and to reduce the likelihood that news media articles are not based on self-serving rumors by interested parties. For example, articles by Hector (1989) and Khalaf (1992) discuss changes in useful lives at General Motors, unusual charges at General Electric and short sellers' interest in Advanta Corp. Neither firm is subsequently required to reverse the effects of its accounting decisions and thus, the firms are excluded from the sample. Similarly, some firms such as Battle Mountain Gold and Blockbuster voluntarily change their accounting choices or estimates as a result of pressure from the investment community. Their choices are initially within GAAP and they do not restate.

The 25 additional manipulators identified by the news media search have similar size, leverage, liquidity, profitability and growth characteristics than the 49 SEC manipulators suggesting that manipulators in both searches are not drawn from different populations. The final sample consists of 74 firms that manipulated earnings and 2332 COMPUSTAT non-manipulators matched by 2-digit SIC industry and year for which financial statement data used in the model are available.<sup>4</sup> There are 63 different four-digit SIC codes represented, with four firms in SIC 7372

---

<sup>4</sup> I treat firms in the same industry for which my searches do not identify an instance of manipulation as non-manipulators. Since successful manipulators would not be identified by the searches, it is possible that the control

(Software), and three firms in both SIC 3571 (Computers), and SIC 5731 (Audio-visual retail stores). The distribution of manipulators by two-digit SIC groups indicates a concentration of firms in manufacturing (SIC 30-39) and Personal and Business Services industry groups (SIC 70 to 79) which together represent 45% of the sample.

### **Characteristics of Sample Firms**

In Table 1, I compare manipulators' financial characteristics to those of industry-matched controls. I find that, in the fiscal year prior to the year containing the public disclosure of earnings manipulation, manipulators are smaller (when size is measured either in terms total assets or sales), less profitable, and more levered. Manipulators also differ from control firms in that they experience higher growth. The median sales growth of manipulators (34.4%) is significantly larger than that of controls (9.4%). This raises the question of whether growth is exogenous or results from manipulation. In the year prior to the fiscal year of manipulation, I find that manipulators also have significantly higher growth than non-manipulators (medians are 29.4% v. 10.6%) suggesting that growth originates exogenously. This profile of manipulators as firms with high growth prospects could explain why I find that manipulators have, on average, lower total assets but similar market value of equity than control firms.

### **Method**

This section discusses the estimation of the earnings manipulation detection model and the selection of the model's variables. The model is written as follows:

$$M_i = b' X_i + \tilde{\epsilon}_i \tag{1}$$

---

sample of 2332 contains manipulators. This biases against discriminating manipulators from non-manipulators, making the tests more conservative.

where  $M$  is a dichotomous variable coded 1 for manipulators and 0 otherwise,  $X$  is the matrix of explanatory variables, and  $\tilde{\epsilon}$  is a vector of residuals.

### **Estimation**

Earnings manipulators are oversampled relative to their true proportion in the population. The econometric justification for such a state-based sample is that a random sample would likely generate a smaller number of manipulators, thus making the identification of an earnings manipulation classification model difficult. However, since estimation of a dichotomous state model that ignores the state-based sample procedures yields asymptotically biased coefficient estimates, I use weighted exogenous sample maximum likelihood probit (WESML).<sup>5</sup> The estimation sample spans the period 1982-1988 and consists of 50 manipulators and 1708 controls.

Using WESML requires an estimate of the proportion of firms in the population that manipulate earnings. Assuming that the population from which the firms are sampled is the population of COMPUSTAT firms, one estimate of the proportion of manipulators in the population equals .0069 (50/7231). Because I have no way of assessing the validity of this assumption, I also evaluate the sensitivity of the model to the different specifications of the prior probability of manipulation. Having discussed estimation issues, I now turn to the composition of the  $X$  matrix.

### **Can accounting data be used to detect earnings manipulation?**

If financial statement manipulations encompass not only earnings but also other signals that

---

<sup>5</sup> WESML accounts for state-based sampling by weighing the likelihood function according to the proportion of earnings manipulators in the sample and in the population. For a discussion of the implications of using state-based samples see Hsieh, et al. (1985). Prior research has employed weighted probit to predict audit qualifications, (Dopuch et al. (1987)) and bankruptcy (Zmijewski (1984)).

investors and analysts rely on, then the discriminatory power of accounting data is diminished. This would bias the results against rejection of a null hypothesis on the variables' coefficients, and limit the usefulness of using accounting information for detecting earnings manipulation. In the absence of an economic theory of manipulation, I rely on three sources to choose explanatory variables based on financial statement data. First, I consider signals about future prospects that appear in the academic and practitioner literature.<sup>6</sup> The presumption is that earnings manipulation is more likely when firms' future prospects are poor. Second, I consider variables based on cash flows and accruals (Healy (1985), Jones (1991)). Third, I consider variables drawn from positive theory research that hypothesizes contract-based incentives for earnings management (Watts and Zimmerman (1986)).

The model includes eight variables.<sup>7</sup> The variables are measured using data from the fiscal year of the first reporting violation, e.g., the first year for which the firm is subsequently required to restate. I designate seven of the eight variables as indices because they are intended to capture distortions that could arise from manipulation by comparing financial statement measures in the year of the first reporting violation to the year prior. The variables are thus not measured contemporaneously with manipulation discovery since, in line with Feroz, Park, and Pastena

---

<sup>6</sup> Specifically, I use constructs that analysts consider as indicators of future performance. See for example, O'Glove (1987), Kellogg and Kellogg (1991), Siegel, (1991), Fridson (1993) and Lev and Thiagarajan (1993).

<sup>7</sup> I also considered but did not include in the model five other types of variables to examine whether the model could be improved: (i) variables isolating the income effect of non-recurring items (the ratio unusual items to pretax income (COMPUSTAT #17/#170), and the ratio of net non-operating income to income before extraordinary items (COMPUSTAT #61/#18), (ii) variables capturing the rate and changes in the rate of intangible amortization as well as variables identifying the funding status of pension funds, (iii) cash flow based variables (the cash flow adequacy ratio and the cash flow coverage of debt service), and (iv) signals of earnings quality documented by Lev and Thiagarajan (1993): changes in the receivable provision, changes in capital expenditures, changes in the effective tax rate, changes in employee productivity, and a variable indicating whether the firm used LIFO to value its inventory, and (v) a day sales in inventory index similar to the day sales in receivables index. None of these variables improved the model performance, and are not reported.

(1991), manipulation becomes public on average 19 months after the end of the fiscal year of the first reporting violation. I discuss below the measurement of each variable, and how I expect it to affect the likelihood of manipulation.

### 1. Days Sales in Receivables Index (DSRI):

DSRI is the ratio of days sales in receivable in the first year in which earnings manipulation is uncovered (year t) to the corresponding measure in year t-1. This variable gauges whether receivables and revenues are in or out-of-balance in two consecutive years. A large increase in days sales in receivables could be the result of a change in credit policy to spur sales in the face of increased competition, but disproportionate increases in receivables relative to sales may also be suggestive of revenue inflation. I thus expect a large increase in days sales in receivables to be associated with a higher likelihood that revenues and earnings are overstated.

### 2. Gross Margin Index (GMI):

GMI is ratio of the gross margin in year t-1 to the gross margin in year t. When GMI is greater than 1, it indicates that gross margins have deteriorated. Lev and Thiagarajan (1993) suggest that gross margin deterioration is a negative signal about firms' prospects. If firms with poorer prospects are more likely to engage in earnings manipulation, I expect a positive relation between GMI and the probability of earnings manipulation.<sup>8</sup>

### 3. Asset Quality Index (AQI):

Asset quality in a given year is the ratio of non-current assets other than property plant and equipment (PPE) to total assets and measures the proportion of total assets for which future benefits are potentially less certain. AQI is the ratio of asset quality in year t, relative to asset quality in year t-1. AQI is an aggregate measure of the change in the asset realization risk analysis suggested by Siegel (1991). If AQI is greater than 1 it indicates that the firm has potentially increased its involvement in cost deferral.<sup>9</sup> I thus expect a positive relation

---

<sup>8</sup> It is possible that manipulation of inventories and other production costs can lead to increasing gross margins. This would suggest that either increased or decreased gross margins can increase the likelihood of manipulation. While Kellogg and Kellogg (1991, p. 10-16) state: "Barring unusual circumstances, the percentage of gross profit to sales should remain the same from year to year." it is difficult to determine what "the same" means. I considered a variable relating gross margin changes to inventory changes but it did not enhance the specification of the model.

<sup>9</sup> It is possible that part of the increase is attributable to acquisitions involving Goodwill. However, sample manipulators undertake few acquisitions and those are primarily stock-for-stock exchanges accounted for using pooling of interests. Nevertheless, I also calculate this variable using the ratio of non-current assets other than PPE and Goodwill to total assets and find similar results.

between AQI and the probability of earnings manipulation. An increase in asset realization risk indicates an increased propensity to capitalize and thus defer costs.

#### 4. Sales Growth Index (SGI):

SGI is the ratio of sales in year  $t$  to sales in year  $t-1$ . Growth does not imply manipulation, but growth firms are viewed by professionals as more likely to commit financial statement fraud because their financial position and capital needs put pressure on managers to achieve earnings targets (National Commission on Fraudulent Financial Reporting (1987), National Association of Certified Fraud Examiners (1993)). In addition, concerns about controls and reporting tend to lag behind operations in periods of high growth (National Commission on Fraudulent Financial Reporting (1987), Loebecke et al. (1989)). If growth firms face large stock prices losses at the first indication of a slowdown, they may have greater incentives to manipulate earnings. To this effect, Fridson (1993, pp. 7-8) states: "Almost invariably, companies try to dispel the impression that their growth is decelerating, since that perception can be so costly to them." I thus expect a positive relation between SGI and the probability of earnings manipulation.

#### 5. Depreciation Index (DEPI):

DEPI is the ratio of the rate of depreciation in year  $t-1$  vs the corresponding rate in year  $t$ . The depreciation rate in a given year equals is equal to  $\text{depreciation}/(\text{depreciation} + \text{net PPE})$ . A DEPI greater than 1 it indicates that the rate at which assets are depreciated has slowed down--raising the possibility that the firm has revised upwards the estimates of assets useful lives or adopted a new method that is income increasing.<sup>10</sup> I thus expect a positive relation between DEPI and the probability of manipulation.

#### 6. Sales General and Administrative Expenses Index (SGAI):

SGAI is calculated as the ratio of SGA to sales in year  $t$  relative to the corresponding measure in year  $t-1$ . The variable is used following Lev and Thiagarajan's (1993) suggestion that analysts would interpret a disproportionate increase in sales as a negative signal about firms future prospects. I expect a positive relation between SGAI and the probability of manipulation.

#### 7. Leverage Index (LVGI):

---

<sup>10</sup> To allow for the possibility that firms manipulate earnings by using lower depreciation rates than comparable firms in the industry, I estimated the model using the depreciation rate instead changes in the depreciation rate. This variable did not enhance the specification of the model and did not alter the magnitude or the significance of the coefficients on the other variables.

LVGI is the ratio of total debt to total assets in year  $t$  relative to the corresponding ratio in year  $t-1$ . A LVGI greater than 1 indicates an increase in leverage. The variable is included to capture debt covenants incentives for earnings manipulation. Assuming that leverage follows a random walk, LVGI implicitly measures the leverage forecast error. I use the change in leverage in the firms' capital structure given evidence in Beneish and Press (1993) that such changes are associated with the stock market effect of default.

#### 8. Total Accruals to Total Assets (TATA):

Total accruals are calculated as the change in working capital accounts other than cash less depreciation. Either total accruals or a partition thereof has been used in prior work to assess the extent to which managers make discretionary accounting choices to alter earnings (see for example Healy (1985), Jones (1991)). I use total accruals to total assets to proxy for the extent to which cash underlies reported earnings, and expect higher positive accruals (less cash) to be associated with a higher likelihood of earnings manipulation.

The explanatory variables in the model are primarily based on year-to-year changes and this introduces a potential problem when the denominator is small. To deal with this problem, I winsorized the data at the 1% and 99 % percentiles for each variable. In addition, there were instances where the denominator of the Asset Quality Index variable was zero as assets in the reference year (period  $t-1$ ) consisted exclusively of current assets and PPE. Since in such cases the Asset Quality Index was not defined, I set its value to one (its neutral value) instead of treating the observation as missing. Similarly, I set the Depreciation and SGA indices to values of one, when elements of the computation (Amortization of Intangibles (#65), and SG&A (#189)) were not available on the COMPUSTAT tapes. I found that estimating the model after excluding those observations yielded similar results.

In table 2, I compare the distribution of these variables for manipulators and non-manipulators in the estimation sample. The results indicate that on average, manipulators have significantly larger increases in days sales in receivables, greater deterioration of gross margins and asset quality, higher growth, and larger accruals.

## Estimation Results and Holdout Sample Tests

Table 3, Panel A reports the results of the WESML probit and unweighted probit estimations of the model. The likelihood ratio test indicates that for both estimations the model has significant power, with  $\chi^2$  statistics (p-values) of 34.5 (.00) and 129.2 (.00). The model has descriptive validity with pseudo- $R^2$ s of 30.6% and 37.1% for WESML and unweighted probit respectively. Since coefficient estimates have similar magnitudes and significance across estimations, I discuss the results of the unweighted probit estimation.

The variable days sales in receivables index has a positive coefficient, .920, and is significant at the 5% level with an asymptotic t-statistic of 6.02. This is consistent with disproportionate increases in receivables raising the likelihood that a firm has inflated revenues. The variable gross margin index has a positive coefficient of .528 that is over two standard deviations from zero. This is consistent with firms facing poor prospects having greater incentives for earnings manipulation. The asset quality index also has a significant positive coefficient (.404, t-statistic 3.20), consistent with the likelihood of earnings manipulation increasing when firms' change their involvement in cost deferral. The sales growth index has a positive coefficient that is over five standard deviations from zero, consistent with growth firms facing growth deceleration having more incentives to manipulate earnings. The accruals to total assets variable has a significant positive coefficient consistent with less cash being behind accounting income for manipulators.

The coefficients on the leverage, depreciation and SGA variables are not significant. It is possible that these variables are associated with earnings management, not manipulation. For

example, a change from accelerated depreciation to straight line or a revision that lengthens useful lives, would result in higher values of the depreciation index. However, this is an instance of earnings management and the firm would not be included in the sample. Similarly, for the leverage variable, incentives to comply with debt covenants may be insufficient to induce earnings manipulation because the costs of non-compliance are small (Beneish and Press (1993) estimate these costs to range between 1 and 2% of market value of equity).<sup>11</sup>

In table 3, Panel B I report the estimated probabilities of earnings manipulation for both the estimation and holdout samples. For the estimation sample, the model estimated using WESML predicts higher average (median) probability of earnings manipulation .107 (.024) for manipulators than for non-manipulators .006 (.003). Similarly, the model estimated using unweighted probit predicts higher average (median) probabilities for manipulators .237 (.099) than for non-manipulators .022 (.011). Wilcoxon and median tests reject the null hypothesis that estimated probabilities for manipulators and non-manipulators are drawn from the same distribution. Results for the holdout sample of 24 manipulators and 624 controls are similar to the estimation sample findings. The model predicts that manipulators are, on average, about 10 times more likely to manipulate earnings. The distributions of estimated probabilities for manipulators and non-manipulators based on unweighted probit illustrate these differences. For example, in the estimation sample, nearly all the non-manipulators (93.4%) have an estimated probability less than .05 compared to 38.0% of the manipulators. Similarly, in the holdout sample, 56.1% of the non-manipulators have an estimated probability less than .01, compared to 20.8% of the manipulators.

---

<sup>11</sup> I also consider three alternative definitions of leverage: total debt to market value of equity, total debt to book value of equity, and long-term debt to total assets as well as using leverage level variables instead of changes. None of the alternative leverage measures attains significance. .

## **Robustness Checks**

I assess the robustness of the results in three ways. First, even though collinearity is not likely to be a problem (none of the 36 Pearson correlation coefficients is greater than .25), I drop up to four variables from the model to assess the stability of the coefficient estimates. Dropping the depreciation, leverage, SGA, and accrual variables one at a time and in combination yields similar results for the remaining variables. Second, I assess the sensitivity of the WESML estimation results to the specification of the prior probability of manipulation. In addition the estimations based on prior probabilities of .0069 and .02844 (implicit in unweighted probit), I estimate the model with four alternative prior probabilities of earnings manipulation, namely .0059, .0079, .0089, .0099. The four new estimations yield similar results with  $\chi^2$  statistics ranging between 29.61 and 48.65 and pseudo-R<sup>2</sup>s ranging from 29.81% and 32.65%. Moreover, the coefficients estimates are similar in size and significance across the four new specifications of the prior probability of manipulation.

Third, while the holdout sample is chosen to be independent from the estimation sample, I assess the sensitivity of the results to the choice of estimation and holdout samples. To do so, I generate 100 random samples of 50 manipulators and 1500 controls with the RANUNI function in SAS and use these to estimate the model 100 times. Similarly, I obtain 100 random holdout samples by treating the complement of 24 manipulators and 832 controls to each random estimation sample as a holdout sample and reproduce the tests on estimated probabilities. The results are reported in Table 4 and the evidence suggests that the findings are not sensitive to the choice of estimation/holdout samples.

Overall, the estimation results provide evidence of a systematic relation between the

likelihood of manipulation and financial statement data. Since the model distinguishes manipulators from non-manipulators, I assess its usefulness as a classification tool. Below, I present evidence on the probability cut-offs associated with differing costs of making classification errors.

### **The Model as a Classification Tool**

The model can make two types of errors: it can classify a firm as a non-manipulator when it manipulates (thereafter Type I error) and it can classify a firm as a manipulator when it does not manipulate (thereafter Type II error). The probability cut-offs that minimize the expected costs of misclassification depend on costs associated with the relative cost of making an error of either type. While decision makers' objective functions are not observable, classification error costs likely differ across decision makers. For example, investors are likely to have high Type I error costs since the stock price loss associated with the discovery of the manipulation is dramatic whereas their Type II error cost would be low given the availability of substitutes. On the other hand, a regulator's objective function requires balancing the protection of the investing public against the costs of falsely accusing a firm. Their relative costs cannot be measured but it is likely that their Type II error costs are higher than those of investors.

The cost of Type I and Type II errors is not amenable to objective measurement and I consider relative costs ranging from 1:1 to 100:1. For investors, however, the relevant range is likely between 20:1 and 30:1. To explain, following Beneish (1999), the typical manipulator loses approximately 40% of its market value on a risk-adjusted basis in the quarter containing the discovery of the manipulation. Assuming that, on a similar basis, a typical firm's equity appreciates between 1 and 2% per quarter, it takes 20 to 40 non-manipulators in the investor's

portfolio to offset a single manipulator in that quarter. As such, one possibility is that investors view a type I error as 20 to 40 times as costly as a type II error.

I compute the probability cut-offs that minimize the expected costs of misclassification (ECM) and present the evidence in Table 5.<sup>12</sup> The results are similar across estimation methods and I focus the discussion on the unweighted probit estimation in Panel B. In the estimation sample, at relative error costs of 10:1, the model classifies firms as manipulators when the estimated probabilities exceed .0685 (a score greater than -1.49); it misclassifies 42% of the manipulators and 7.6% of the non-manipulators. Similarly, at relative error costs of 20:1 or 30:1, the model classifies firms as manipulators when the estimated probabilities exceed .0376 (a score greater than -1.78); it misclassifies 26% of the manipulators and 13.8% of the non-manipulators. In the holdout sample, at relative error costs of 20:1 or 30:1, the model classifies firms as manipulators when the estimated probabilities exceed .0376 (a score greater than -1.78); it misclassifies 50% of the manipulators and 7.2% of the non-manipulators.

I also report the performance of the unweighted probit model in Figures 1 and 2. The figures contain the following information (1) the probability cut-offs associated with each relative error cost assumption, (2) the percentage of correctly classified manipulators, and (3) the percentage of incorrectly classified non-manipulators. For the estimation sample in figure 1, the percentage of correctly classified manipulators ranges from 58 to 76%, while the percentage of incorrectly classified non-manipulators ranges from 7.6% to 17.5%. For the holdout sample in

---

<sup>12</sup> Expected costs of misclassification are computed as follows:

$$ECM = P(M) P_I C_I + (1-P(M)) P_{II} C_{II} \quad (2)$$

where  $P(M)$  is the prior probability of encountering earnings manipulators (.0069 for WESML and .02844 for unweighted probit),  $P_I$  and  $P_{II}$  are the conditional probabilities of a Type I and Type II errors and  $C_I$  and  $C_{II}$  are the costs of Type I and Type II errors.

figure 2, the percentage of correctly classified manipulators ranges from 37.5 to 56%, while the percentage of incorrectly classified non-manipulators ranges from 3.5% to 9.1%.

While these results suggest that the model identifies potential manipulators, it does so with large error rates in the range of error costs that are likely to be of relevance to investors. Since instances of discovered manipulations are rare, this raises the question of whether the model is useful relative to a naïve strategy that classifies all firms as non-manipulators.

I thus compare (in Table 5) the model's expected costs of misclassification to those of a naive strategy of classifying all firms as non-manipulators. The naive strategy makes no Type II errors ( $P_{II}=0$ ) and the conditional probability of a Type I error ( $P_I$ ) is one. Thus,  $ECM(\text{naive})=P(M) C_I$  or  $.0069 C_I$  for the WESML comparison and  $.02844 C_I$  for the unweighted probit comparison. For both the estimation and the holdout samples, I find that the model has lower expected misclassification costs than a naive strategy when the cost of a type I error is greater than that of a type II error. For example in the estimation sample in Panel B, the ratio of the cost of the model's errors to the cost of errors from a naive strategy is .496 and .417 at relative errors costs of 20:1 and 30:1. Similarly, for the holdout sample, the ratio of the cost of the model's errors to the cost of errors from a naive strategy is .623 and .582 at relative errors costs of 20:1 and 30:1. The preceding evidence thus suggests that the model is cost effective relative to a naive strategy that treats all firms as non-manipulators.

## **Conclusion**

The evidence in this paper is based on a sample of firms whose manipulation of earnings was publicly discovered. Such firms likely represent the upper tail of the distribution of firms that seek to influence their reported earnings and the evidence should be interpreted in this light.

Given this caution, the evidence of a systematic association between earnings manipulation and financial statement data is of interest to both accounting researchers and professionals because it suggests that accounting data not only meet the test of providing useful information but also enable an assessment of reliability. The explicit classification model only requires two years of data (one annual report) to evaluate the likelihood of manipulation and can be inexpensively applied by the SEC, auditors, and investors to screen a large number of firms and identify potential manipulators for further investigation.

While the model is cost-effective relative to a strategy of treating all firms as non-manipulators, its large rate of classification errors makes further investigation of the results an important element to the model's implementation. That is, since the model's variables exploit distortions in financial statement data that could result from manipulation, one must recognize that such distortions can have an alternative origin. For example, they could be the result of a material acquisition during the period examined, a material shift in the firm's value maximizing strategy, or a significant change in the firm's economic environment.

One limitation of the model is that it is estimated using financial information for publicly traded companies. Therefore, it cannot be reliably used to study privately-held firms. Another limitation is that the earnings manipulation in the sample involves earnings overstatement rather than understatement and therefore, the model cannot be reliably used to study firms operating in circumstances that are conducive to decreasing earnings.

## References

- Beneish, M.D. and E. Press. 1993. "Costs of Technical Violation of Accounting-Based Debt Covenants." *The Accounting Review* vol. 68, no. 2 (April):233-257.
- Beneish, M.D. 1997. "Detecting GAAP Violation: Implications for Assessing Earnings Management Among Firms with Extreme Financial Performance" *Journal of Accounting and Public Policy*, vol. 16, no. 3 (Fall):271-309.
- Beneish, M.D. 1999. "Incentives and Penalties Related to Earnings Overstatements That Violate GAAP." Forthcoming, *The Accounting Review*, vol. 74, no. 4 (October).
- Healy, P.M. 1985. "The Effect of Bonus Schemes on Accounting Decisions." *Journal of Accounting and Economics*, vol. 7, no. 1-3 (April):85-107.
- Dopuch, N., R.W. Holthausen, and R.W. Leftwich. 1987. "Predicting Audit Qualifications with Financial and Market Variables." *The Accounting Review*, vol. 62, no. 3 (July):431-454.
- Feroz, E.H., K. Park; and V.S. Pastena. 1991. "The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases." *Journal of Accounting Research*, vol. 29 (Supplement):107-148.
- Flynn, J. and W. Zellner. 1992. "Buying Trash in Big Holes--On the Balance Sheet." *Business Week* (May 11):88-89.
- Fridson, M. S. 1993. *Financial Statement Analysis: A Practitioner's Guide*. New York: John Wiley.
- Hector, G. 1989. "Cute Tricks on the Bottom Line." *Fortune* (April 24):193-200.
- Hsieh, D.A., C. F. Mansky, and D. McFadden. 1985. "Estimation of Response Probabilities from Augmented Retrospective Observations." *Journal of the American Statistical Association*, vol. 80 (September): 651-662.
- Jones, J.J. 1991. "Earnings Management During Import Relief Investigations." *Journal of Accounting Research*, vol. 29, no. 2 (Autumn):193-228.
- Kellogg, I. and L.B. Kellogg. 1991. *Fraud, Window Dressing, and Negligence in Financial Statements*. New York: McGraw-Hill.
- Khalaf, R. 1992. "Fuzzy Accounting." *Forbes* (June 22):96.
- Lev, B. and S. R. Thiagarajan. 1993. "Fundamental Information Analysis." *Journal of Accounting Research*, vol. 31, no. 2 (Autumn):190-215.

- Levitt, A. 1998. "The "Numbers Game." " Securities and Exchange Commission, Remarks by Chairman Arthur Leviitt ( September 28) [www.sec.gov](http://www.sec.gov).
- Loebbecke, J.K., M.M. Eining, and J.J. Willingham. 1989. Auditor's experience with material irregularities: Frequency, nature, and detectability. *Auditing: A Journal of Theory and Practice*, vol. 9, no. 1 (Fall):1-28.
- Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- National Association of Certified Fraud Examiners. 1993. *Cooking the Books: What Every Accountant Should Know about Fraud*. New York: NASBA.
- National Commission on Fraudulent Financial Reporting. 1987. *Report of National Commission on Fraudulent Financial Reporting*. New York: AICPA.
- O'Glove, T. L. 1987. *Quality of Earnings*. New York: The Free Press.
- Siegel, J. G. 1991. *How to Analyze Businesses, Financial Statements, and the Quality of Earnings*. 2nd Edition, New Jersey: Prentice Hall.
- Watts, R. L. and J. L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs: Prentice Hall.
- Zmijewski, M.E. 1984. "Methodological Issues Related to the Estimation of Financial Distress Prediction Models." *Journal of Accounting Research*, vol. 22 (Supplement): 59-82.

Table 1

Comparing Characteristics of 74 Sample Firms vs. 2332 Firms  
Matched by 2-digit SIC Industry in the Year Prior  
the Fiscal Year Containing the Public Disclosure of Earnings Manipulation<sup>a</sup>

---

<u>Characteristic</u>	<u>Treatment Firms</u>		<u>Control Firms</u>		<u>Wilconxon -Z P-Value<sup>b</sup></u>	<u>Median <math>\Pi^2</math> P-value</u>
	<u>Mean</u>	<u>Median</u>	<u>Mean</u>	<u>Median</u>		
<u>Size</u>						
Total assets	467.33	43.20	1140.37	95.84	.003	.004
Sales	469.87	53.56	1295.22	122.54	.001	.007
Market Value	323.72	74.90	813.35	64.59	.884	.701
<u>Liquidity/Leverage</u>						
Working capital to total assets	.26	.28	.30	.31	.472	.345
Current ratio	2.54	1.83	2.54	2.11	.103	.473
Total debt to total assets	.58	.58	.51	.52	.027	.098
<u>Profitability/Growth</u>						
Return on assets	-.01	.03	.03	.05	.063	.078
Sales Growth	.58	.34	.13	.09	.000	.001

---

<sup>a</sup> The COMPUSTAT firms in the same 2-digit SIC code for which financial statement data are available comprise the comparison sample.

<sup>b</sup> The Wilcoxon Rank-Sum and the Median tests are used to evaluate the null hypothesis that the size, liquidity, profitability, and growth characteristics of manipulators and non-manipulators are drawn from the same population.

Table 2

Potential Predictive Variables: Descriptive Statistics for the Estimation Sample of 50 pre-1989 Manipulators and their 1708 Industry-Matched Non-Manipulators

Characteristic <sup>a</sup>	Treatment Firms (N= 50)		Controls (N= 1708)		Wilcoxon-Z	Median Test
	Mean	Median	Mean	Median	P-Value <sup>b</sup>	P-Value <sup>b</sup>
Days in Receivables	1.465	1.281	1.031	.996	.000	.000
Gross Margin Index	1.193	1.036	1.014	1.001	.006	.007
Asset Quality Index	1.254	1.000	1.039	1.000	.096	.246
Sales Growth Index	1.607	1.411	1.134	1.106	.000	.000
Depreciation Index	1.077	.966	1.001	.974	.307	.774
SGA Index	1.041	.960	1.054	1.010	.271	.389
Leverage Index	1.111	1.030	1.037	1.000	.394	.077
Accruals to total assets	.031	.034	.018	.013	.000	.002

For each variable, I provide a definition and the COMPUSTAT data item number. Year t refers the first year in which earnings manipulation occurs:

$$\text{Days Sales in Receivables Index} = (\text{Receivables}_t[2]/\text{Sales}_t[12]) / (\text{Receivables}_{t-1}/\text{Sales}_{t-1})$$

$$\text{Gross Margin Index} = \left( \frac{\text{Sales}_{t-1}[12] - \text{Costs of Goods Sold}_{t-1}[41]}{\text{Sales}_{t-1}[12]} \right) / \left( \frac{\text{Sales}_t[12] - \text{Costs of Goods Sold}_t[41]}{\text{Sales}_t[12]} \right)$$

$$\text{Asset Quality Index} = \left( 1 - \frac{\text{Current Assets}_t[4] + \text{PPE}_t[8]}{\text{Total Assets}_t[6]} \right) / \left( 1 - \frac{\text{Current Assets}_{t-1} + \text{PPE}_{t-1}}{\text{Total Assets}_{t-1}} \right)$$

$$\text{Sales Growth Index} = \text{Sales}_t[12] / \text{Sales}_{t-1}$$

$$\text{Depreciation Index} = \left( \frac{\text{Depreciation}_{t-1}[14 \text{ less } 65]}{\text{Depreciation}_{t-1} + \text{PPE}_{t-1}[8]} \right) / \left( \frac{\text{Depreciation}_t}{\text{Depreciation}_t + \text{PPE}_t} \right)$$

$$\text{SGA Index} = \left( \frac{\text{SGA Expense}_t[189]}{\text{Sales}_t[12]} \right) / \left( \frac{\text{SGA Expense}_{t-1}}{\text{Sales}_{t-1}} \right)$$

$$\text{Leverage Index} = \left( \frac{\text{LTD}_t[9] + \text{Current Liabilities}_t[5]}{\text{Total Assets}_t[6]} \right) / \left( \frac{\text{LTD}_{t-1} + \text{Current Liabilities}_{t-1}}{\text{Total Assets}_{t-1}} \right)$$

$$\begin{aligned} \text{Accruals to Total Assets} = & \left[ \text{Current Assets}_t[4] - \text{Cash}_t[1] - \text{Current Liabilities}_t[5] \right. \\ & \left. - \text{Current Maturities of LTD}_t[44] - \text{Income Tax Payable}_t[71] - \text{Depreciation and} \right. \\ & \left. \text{Amortization}_t[14] \right] / \text{TA}_t[6] \end{aligned}$$

<sup>b</sup>The Wilcoxon Rank-Sum and the Median tests compare the distribution of sample firms' characteristics to the corresponding distribution for non-manipulators. The reported p-values indicate the smallest probability of incorrectly rejecting the null hypothesis of no difference.

Table 3

WESML and Unweighted Probit Estimation Results Based on an Estimation Sample of  
50 Manipulators and 1708 Non-manipulators (Panel A). Estimated Probabilities of  
Manipulation for the Estimation Sample and for a Holdout Sample of 24 Manipulators and on 624  
Non-manipulators (Panel B)<sup>a</sup>

## Panel A: Estimation Results

Predicted Sign	Constant	Days in Receivables Index <sup>b</sup>	Gross Margin Index	Asset Quality Index	Sales Growth Index	Depreciation Index	SGA Index	Accruals to total Assets	Leverage Index	Pseudo R <sup>2c</sup>	$\chi^2$ -statistic p-value <sup>c</sup>
		(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)		
WESML <sup>d</sup>	-4.954 (-11.80)	.789 (6.40)	.459 (3.02)	.306 (2.82)	.701 (3.43)	.033 (.15)	-.006 (-.04)	3.937 (3.07)	-.264 (-.83)	.306	34.50 (.00)
Unweighted Probit <sup>d</sup>	-4.840 (-11.01)	.920 (6.02)	.528 (2.20)	.404 (3.20)	.892 (5.39)	.115 (.70)	-.172 (-.71)	4.679 (3.73)	-.327 (-1.22)	.371	129.20 (.00)

## Panel B: Estimated Probabilities of Manipulation

	<u>WESML</u>				<u>Unweighted Profit</u>			
	Estimation Sample		Holdout Sample		Estimation Sample		Holdout Sample	
	<u>Manipulators</u>	<u>Non-Manipulators</u>	<u>Manipulators</u>	<u>Non-Manipulators</u>	<u>Manipulators</u>	<u>Non-Manipulators</u>	<u>Manipulators</u>	<u>Non-Manipulators</u>
Mean	.107	.006	.097	.007	.237	.022	.181	.019
St. Dev.	.175	.021	.223	.044	.275	.051	.288	.063
MaximLxn	.851	.615	.999	.999	.980	.960	.999	.999
Median	.024	.003	.009	.002	.099	.011	.037	.009
Minimum	.001	.001	.001	.001	.001	.001	.004	.001
Witcoxon-Z <sup>1</sup> (p-value)	8.049 (.000)		4.721 (.000)		8.314 (.000)		4.630 (.000)	
Median test-X <sup>1</sup> (p.value)	23.785 (.000)		13.995 (.003)		26.667 (.000)		11.056 (.001)	

<sup>a</sup> The estimation sample consists of the pre-1989 manipulators and their controls and the holdout sample of the post-1988 manipulators and their controls.

<sup>b</sup> See Table 2 for variable definitions.

<sup>c</sup> The pseudo-R<sup>2</sup> is equal to  $(L_{\Omega}^{2/n} - L_w^{2/n}) / (1 - L_w^{2/n})$  where  $L_{\Sigma}$  is the log likelihood for the WESML probit model (unconstrained),  $L_T$  is the log likelihood with only the constant term in the model (constrained) and  $n$  the number of observations (See Maddala (1983, p. 40)). The log likelihood ratio test statistic is equal to -2 times the difference in the log likelihood of the unconstrained and constrained models is asymptotically distributed  $\Pi^2$  with degrees of freedom equal to the difference in the number of parameters of the two models.

<sup>d</sup> Weighted exogenous maximum likelihood probit is estimated assuming that prior probability of manipulation is .0069. Sensitivity analysis on the prior probability of manipulation yields coefficients estimates of similar magnitude are significance. When the prior probability of manipulation is specified as .0059, .0079 and .0089 the estimation yields  $\chi^2$  statistics of 29.60, 39.22, 43.95 and 48.65, significant at the 1% level or lower. Unweighted probit implicitly assumes that the prior probability of manipulation is .02844 (50/1758).

<sup>e</sup> Tests that the estimated probabilities for manipulators and non-manipulators are drawn from the same distribution.

Table 4

Sensitivity Analysis to the choice of Estimation and Holdout Samples. Descriptive Statistics for Estimation Based on 100 Random Samples of 50 Manipulators and 1500 Non-Manipulators [Panel A], and Descriptive Statistics on the Estimated Probabilities of 100 Holdout Samples of 24 Manipulators and 832 Non-Manipulators [Panel B]<sup>a</sup>

Panel A: Descriptive statistics on 100 Estimation Samples

	Mean	Standard Deviation	Max	Median	Min	Percent Positive	Percentage Significant at 10% <sup>b</sup>	Percentage Significant at 5% <sup>b</sup>	Percentage Significant at 2.5% <sup>b</sup>
Constant	-4.223	.549	-3.404	-4.040	-5.853	0.0	100.0	100.0	100.0
Days in Receivables <sup>b</sup> Index	.857	.097	1.065	.864	.588	100.0	100.0	100.0	100.0
Gross Margin Index	.488	.115	.871	.487	.222	100.0	95.0	84.0	66.0
Asset Quality Index	.453	.113	.789	.438	.223	100.0	95.0	84.0	96.0
Sales Growth Index	.374	.365	1.232	.152	.103	100.0	100.0	100.0	100.0
Depreciation Index	.059	.183	.437	.097	-.782	81.0	37.0	18.0	10.0
SGA Index	-.144	.180	.333	-.156	-.559	25.0	30.0	12.0	4.0
Accruals to Total Assets	4.370	.965	7.219	4.464	2.090	100.0	99.0	95.0	93.0
Leverage Index	-.110	.165	.278	-.114	-.544	25.0	8.0	2.0	1.0
Pseudo-R <sup>2</sup>	.242	.068	.444	.220	.124	-	-	-	-
$\chi^2$ -statistic	89.79	19.59	142.69	84.49	51.90	-	100.0	100.0	-

Panel B: Descriptive Statistics on Estimated Probabilities on 100 Holdout Samples

	Mean	Standard Deviation	Max	Median	Min	Wilcoxon-Z P-value	Median $\chi^2$ P-value
Manipulators	.178	.049	.316	.164	.091	12.212	199.09
Non-Manipulators	.028	.002	.033	.028	.024	(.000)	(.000)

<sup>a</sup> Random samples are generated using the RANUNI function in SAS. RAPUNI is used 100 times to generate 100 samples of 50 manipulators out of 74 and 1500 controls. Each time, the complement of 24 manipulators and 832 non-manipulators is considered as a holdout sample.

<sup>b</sup> Significance based on one-tailed test.

<sup>c</sup> Variables and statistics are defined in Tables 2 and 3.

<sup>d</sup> Tests that the estimated probabilities for manipulators and non-manipulators are drawn from the same distribution.

Table 5

Cut-off probabilities, Costs of Misclassification, and Probability of Type I and Type II Errors for Various Levels of Relative Costs in the Estimation Sample (50 manipulators, 1708 non-manipulators) and in the Holdout Sample (24 manipulators, 624 non-manipulators)<sup>a</sup>

## Panel A: WESML

Relative costs of Type I and Type II errors	Cut-off Probability <sup>b</sup>	Estimation Sample Probability of Classification errors		Cost of Model Errors Relative To Naïve Strategy <sup>c</sup>	Holdout Sample Probability of Classification Errors		Cost of Model Errors Relative to Naïve Strategy
		Type I	Type II		Type I	Type II	
1:1	1.0000	1.0000	.0000	1.000	1.0000	.0000	1.000
10:1	.2905	.9000	.0004	.991	.9166	.0048	.986
20:1	.0512	.5600	.0409	.855	.7500	.0112	.830
30:1	.0512	.5600	.0409	.757	.7500	.0112	.804
40:1	.0223	.4600	.0632	.688	.6667	.0240	.753
60:1	.0092	.2800	.1329	.597	.5000	.0689	.665
100:1	.0087	.2600	.1417	.464	.5000	.0753	.608

## Panel B: Unweighted Probit

Relative costs of Type I and Type II errors	Cut-off Probability	Estimation Sample Probability of Classification errors		Cost of Model Errors Relative To Naïve Strategy	Holdout Sample Probability of Classification Errors		Cost of Model Errors Relative to Naïve Strategy
		Type I	Type II		Type I	Type II	
1:1	1.0000	1.0000	.0000	1.000	1.0000	.0000	1.000
10:1	.0685	.4200	.0761	.680	.6250	.0353	.746
20:1	.0376	.2600	.1382	.496	.5000	.0721	.623
30:1	.0376	.2600	.1382	.417	.5000	.0721	.582
40:1	.0294	.2400	.1747	.433 <sup>d</sup>	.4583	.0913	.628 <sup>d</sup>
60:1	.0294	.2400	.1747	.562 <sup>d</sup>	.4583	.0913	.896 <sup>d</sup>
100:1	.0294	.2400	.1747	.819 <sup>d</sup>	.4583	.0913	1.535 <sup>d</sup>

<sup>a</sup> A Type I error is defined as classifying an observation as a non-manipulator when it manipulates. A type II error is defined as classifying and observation as a manipulator when it is a non-manipulator.

<sup>b</sup> Cut-off probabilities are chosen for each level of relative costs to minimize the expected costs of misclassification as defined in equation (2).

<sup>c</sup> A naïve strategy classifies all firms as non-manipulators. As such, the naïve strategies expected cost of misclassification is .0069  $C_1$  for the WESML model and .02844  $C_1$  for the unweighted probit.

<sup>d</sup> In these computations, the naïve strategy classifies all firms as manipulators. The switch in naïve strategies minimizes the expected costs of misclassification because the ratio of relative costs is greater than is the population proportion of manipulators. The switch occurs at 40:1 for unweighted probit (>1/.02844).

**Figure 1: The Classification Performance of the Model for Different Relative Error Cost Assumptions (Estimation Sample).**

