# The Detection of Earnings Manipulation

Messod D. Beneish

*Presented are a profile of a sample of earnings manipulators, their distinguishing characteristics, and a suggested model for detecting manipulation. The model's variables are designed to capture either the financial statement distortions that can result from manipulation or preconditions that might prompt companies to engage in such activity. The results suggest a systematic relationship between the probability of manipulation and some financial statement variables. This evidence is consistent with the usefulness of accounting data in detecting manipulation and assessing the reliability of reported earnings. The model identifies approximately half of the companies involved in earnings manipulation prior to public discovery. Because companies that are discovered manipulating earnings see their stocks plummet in value, the model can be a useful screening device for investment professionals. The screening results, however, require determination of whether the distortions in the financial statement numbers result from earnings manipulation or have another structural root.*

The extent to which earnings are manipulated has long been of interest to analysts, regulators, researchers, and other investment professionals. The U.S. SEC's recent commitment to vigorous investigation of earnings manipulation (see Levitt 1998) has sparked renewed interest in the area, but the academic and professional literature contains little discussion of the detection of earnings manipulation.

This article presents a model to distinguish manipulated from nonmanipulated reporting.[1] Earnings manipulation is defined as an instance in which a company's managers violate generally accepted accounting principles (GAAP) to favorably represent the company's financial performance. To develop the model, I used financial statement data to construct variables that would capture the effects of manipulation and preconditions that might prompt companies to engage in such activity.

## Sample

The sample consisted of 74 companies that manipulated earnings and all Compustat companies matched by two-digit SIC numbers for which data were available for the 1982–92 period. The 74 earnings manipulators were either companies subject to the SEC's accounting enforcement actions or were identified as manipulators by the news media.

To identify companies subject to accounting enforcement actions by the SEC, I used Accounting and Auditing Enforcement Releases (AAERs) #132 through #502, issued from 1987 to 1993. After elimination of inappropriate companies from the 363 AAERs examined (#372 to #379 were not assigned by the SEC), this search yielded a final sample of 103 AAERs relating to 49 companies that violated GAAP.[2]

An extensive media search on LEXIS-NEXIS for January 1987 to April 1993 identified 80 companies mentioned in articles discussing earnings manipulation.[3] After elimination of eight companies that were also identified by the AAER search and of other inappropriate companies from this search, the final sample from the new search was 40 companies.[4]

For the remaining 40, I required *ex post* evidence of manipulation. That is, I required that the companies had to restate earnings to comply with GAAP at the request of auditors or following the conclusion of an internal investigation. This requirement made this sample consistent with the sample from the AAER search, in the sense that a restatement (in addition to a permanent injunction from future violations of security laws) is usually the outcome of a successful SEC investigation.

*Messod D. Beneish is an associae professor at the Kelley School of Business, Indiana University.*

This criterion was imposed to exclude companies that managed earnings within GAAP and to reduce the likelihood that companies would be considered manipulators in the sample on the basis of media articles based on self-serving rumors by interested parties. For example, articles by Hector (1989) and Khalaf (1992) discussed changes in "useful lives" at General Motors Corporation, unusual charges at General Electric Company, and short sellers' interest in Advanta Corporation. None of the companies was subsequently required to reverse the effects of its accounting decisions; thus, these companies were excluded from the manipulator sample. Similarly, some companies—for example, Battle Mountain Gold Company and Blockbuster Entertainment—voluntarily changed their accounting choices or estimates as a result of pressure from the investment community. Because their accounting choices were initially within GAAP and they did not restate earnings, they also were not considered manipulators. These eliminations reduced the news media sample by 15 companies.

The 25 manipulators identified by the news media search had size, leverage, liquidity, profitability, and growth characteristics that were similar to those of the 49 SEC manipulators, which suggests that manipulators found by the different searches were not drawn from different populations.

## Sample versus Control Companies

The final sample consisted of 74 companies that manipulated earnings, which were matched to 2,332 Compustat nonmanipulators by two-digit SIC industry and year for which the financial statement data used in the model were available.[5] The distribution of manipulators by two-digit SIC groups indicates a concentration of companies in manufacturing (SIC 30–39) and personal and business services (SIC 70–79), which together represented 45 percent of the sample. Of the 63 four-digit SIC codes represented, only 4 companies were in software (SIC 7372), 3 were in computers (SIC 3571), and 3 were in audiovisual retail stores (SIC 5731).

**Table 1** contains a comparison of manipulators' financial characteristics with those of industry-matched controls in the fiscal year prior to the year that contained the public disclosure of the earnings manipulation. In the prior year, manipulators were smaller (when size was measured either in terms of total assets or in terms of sales), less profitable, and more leveraged than control companies. Manipulators also experienced higher growth. The median sales growth of manipulators (34.4 percent) is significantly larger than that of controls (9.4 percent), which raises the question: Is growth exogenous to earnings manipulation or a result of manipulation? I found that in the year prior to the fiscal year of manipulation, manipulators also had significantly higher growth than nonmanipulators (medians were 29.4 percent versus 10.6 percent), which suggests that growth originates exogenously. This profile of manipulators as companies with high-growth prospects could explain why I found manipulators to have, on average, lower total assets than control companies but similar market values of equity.

Sample manipulators typically overstated earnings by recording fictitious, unearned, or uncertain revenues, recording fictitious inventory, or improperly capitalizing costs. The context of earnings manipulation was an annual report or a 10-K for about two-thirds of the sample and was a security offering prospectus (initial, secondary, debt offering) for the remaining third. Sample manipulators were relatively young growth companies,

**Table 1. Characteristics of Sample and Control Companies: Fiscal Year Prior to Public Disclosure, 1982–92 Data**

| | Manipulators | | Nonmanipulators | | Wilcoxon Z | Median $\chi^2$ |
|---|---|---|---|---|---|---|
| Characteristic | Mean | Median | Mean | Median | *p*-Value | *p*-Value |
| *Size (millions)* | | | | | | |
| Total assets | $467.33 | $43.20 | $1,140.37 | $ 95.84 | 0.003 | 0.004 |
| Sales | 469.87 | 53.56 | 1,295.22 | 122.54 | 0.001 | 0.007 |
| Market value | 323.72 | 74.90 | 813.35 | 64.59 | 0.884 | 0.701 |
| *Liquidity/leverage* | | | | | | |
| Working capital to total assets | 0.26 | 0.28 | 0.30 | 0.31 | 0.472 | 0.345 |
| Current ratio | 2.54 | 1.83 | 2.54 | 2.11 | 0.103 | 0.473 |
| Total debt to total assets | 0.58 | 0.58 | 0.51 | 0.52 | 0.027 | 0.098 |
| *Profitability/growth* | | | | | | |
| Return on assets | –1% | 3% | 3% | 5% | 0.063 | 0.078 |
| Sales growth | 58 | 34 | 13 | 9 | 0.000 | 0.001 |

*Note*: The Wilcoxon rank-sum and the median $\chi^2$ tests were used to evaluate the null hypothesis that the size, liquidity, profitability, and growth characteristics of manipulators and nonmanipulators indicate that the groups were drawn from the same population.

which have characteristics that make a company more likely to come under the scrutiny of regulators (see Beneish forthcoming 1999).

## The Model

I estimated a model for detecting earnings manipulation using sample manipulators and industry-matched companies in the 1982–88 period and evaluated the model's performance on a holdout sample in the 1989–92 period. The model for detection of earnings manipulation is

$$M_i = \beta' X_i + \tilde{\varepsilon}_i,$$

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

**Estimation.** In this study, earnings manipulators were oversampled relative to their true proportion in the population. The econometric justification is that such a state-based sample is likely to generate a larger number of manipulators than a random sample would generate, which would make the identification of a model for classifying earnings manipulation difficult. However, because estimation of a dichotomous-state model that ignores the state-based sample procedures would yield asymptotically biased coefficient estimates, I used weighted exogenous sample maximum likelihood (WESML) probit as well as unweighted probit.[6] The estimation sample spanned the 1982–88 period and consisted of 50 manipulators and 1,708 controls.

Using WESML required an estimate of the proportion of companies in the population that manipulate earnings. Assuming that the population from which the companies were sampled is the population of Compustat companies, one estimate of the proportion of manipulators equals 0.0069 (50/7,231). Because I have no way of assessing the validity of this assumption, I also evaluated the sensitivity of the model to other specifications of the prior probability of manipulation.

**Variables: Can Accounting Data Be Used to Detect Earnings Manipulation?** If financial statement manipulations encompass not only earnings but also other signals that investors and analysts rely on, then the discriminatory power of accounting data is diminished, the results of this study are biased against rejection of a null hypothesis on the variables' coefficients, and the usefulness of accounting information for detecting earnings manipulation is limited. In the absence of an economic theory of manipulation, I relied on three sources for choosing explanatory variables based on financial statement data. First, I considered signals about future prospects that appear in the academic and practitioner literature.[7] The presumption was that earnings manipulation is more likely when companies' future prospects are poor. Second, I considered variables based on cash flows and accruals (Healy 1985; Jones 1991). Third, I considered variables drawn from positive theory research, which hypothesizes contract-based incentives for earnings management (Watts and Zimmerman 1986).

The result of the search for variables based on financial statement data was a model that includes eight variables.[8] The variables were measured from data from the fiscal year of the first reporting violation (e.g., the first year for which the company was subsequently required to restate). I designated seven of the eight variables as indexes 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 with those in the year prior.[9]

The computation of each variable and its Compustat number are in the notes to **Table 2**.

- *Days' sales in receivables index*. The *DSRI* is the ratio of days' sales in receivables in the first year in which earnings manipulation was 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 could also suggest revenue inflation. Thus, I expected a large increase in the *DSRI* to be associated with a higher likelihood that revenues and earnings are overstated.
- *Gross margin index*. The *GMI* is the ratio of the gross margin in year $t - 1$ to the gross margin in year $t$. When the *GMI* is greater than 1, gross margins have deteriorated. Lev and Thiagarajan suggested that deterioration of gross margin is a negative signal about a company's prospects. So, if companies with poorer prospects are more likely to engage in earnings manipulation, I expected a positive relationship between *GMI* and the probability of earnings manipulation.[10]
- *Asset quality index*. Asset quality in a given year is the ratio of noncurrent assets other than property, plant, and equipment (PP&E) to total assets and measures the proportion of total assets for which future benefits are potentially less certain. The asset quality index (*AQI*) is the ratio of asset quality in year $t$ to asset quality in year $t - 1$. The *AQI* is an aggregate measure of the change in asset realization risk, which was suggested by Siegel. If

**Table 2. Distribution of Variables for Manipulators and Nonmanipulators in the Estimation Sample**

| | Manipulators ($N$ = 50) | | Nonmanipulators ($N$ = 1,708) | | | |
|---|---|---|---|---|---|---|
| Characteristic | Mean | Median | Mean | Median | Wilcoxon Z $p$-Value | Median Test $p$-Value |
| *DSRI* | 1.465 | 1.281 | 1.031 | 0.996 | 0.000 | 0.000 |
| *GMI* | 1.193 | 1.036 | 1.014 | 1.001 | 0.006 | 0.007 |
| *AQI* | 1.254 | 1.000 | 1.039 | 1.000 | 0.096 | 0.246 |
| *SGI* | 1.607 | 1.411 | 1.134 | 1.106 | 0.000 | 0.000 |
| *DEPI* | 1.077 | 0.966 | 1.001 | 0.974 | 0.307 | 0.774 |
| *SGAI* | 1.041 | 0.960 | 1.054 | 1.010 | 0.271 | 0.389 |
| *LVGI* | 1.111 | 1.030 | 1.037 | 1.000 | 0.394 | 0.077 |
| *TATA* | 0.031 | 0.034 | 0.018 | 0.013 | 0.000 | 0.002 |

$N$ = number of companies.

*Notes*: The definition and Compustat data item number (in brackets) for each variable are as follows (LTD = long-term debt; $t$ and $t-1$ are defined in the text):

$$DSRI = \frac{\text{Receivables}_t[2]/\text{Sales}_t[12]}{\text{Receivables}_{t-1}/\text{Sales}_{t-1}}.$$

$$GMI = \frac{(\text{Sales}_{t-1}[12] - \text{Cost of goods sold}_{t-1}[41])/\text{Sales}_{t-1}[12]}{(\text{Sales}_t[12] - \text{Cost of goods sold}_t[41])/\text{Sales}_t[12]}.$$

$$AQI = \frac{(1 - \text{Current assets}_t[4] + \text{PP\&E}[8])/\text{Total assets}_t[6]}{(1 - \text{Current assets}_{t-1} + \text{PP\&E}_{t-1})/\text{Total assets}_{t-1}}.$$

$$SGI = \frac{\text{Sales}_t[12]}{\text{Sales}_{t-1}}.$$

$$DEPI = \frac{\text{Depreciation}_{t-1}[14-65]/(\text{Depreciation}_{t-1} + \text{PP\&E}_{t-1}[8]}{\text{Depreciation}_t/(\text{Depreciation}_t + \text{PP\&E}_t)}.$$

$$SGAI = \frac{\text{Sales, general, and administrative expense}_t[189]/\text{Sales}_t[12]}{\text{Sales, general, and administrative expense}_{t-1}/\text{Sales}_{t-1}}.$$

$$LVGI = \frac{(\text{LTD}_t[9] + \text{Current liabilities}_t[5])/\text{Total assets}_t[6]}{(\text{LTD}_{t-1} + \text{Current liabilities}_{t-1})/\text{Total assets}_{t-1}}.$$

$$TATA = \frac{\begin{array}{l}\Delta\text{Current assets}_t[4] - \Delta\text{Cash}_t[1] - \Delta\text{Current liabilities}_t[5] \\ -\Delta\text{Current maturities of LTD}_t[44] - \Delta\text{Income tax payable}_t[71] \\ -\text{Depreciation and amortization}_t[14]\end{array}}{\text{Total assets }[6]}.$$

The Wilcoxon and median tests compared the distribution of manipulator companies' characteristics with the corresponding distribution for nonmanipulators. The reported $p$-values indicate the smallest probability of incorrectly rejecting the null hypothesis of no difference.

the *AQI* is greater than 1, the company has potentially increased its involvement in cost deferral.[11] An increase in asset realization risk indicates an increased propensity to capitalize, and thus defer, costs. Therefore, I expected to find a positive relationship between the *AQI* and the probability of earnings manipulation.

*Sales growth index*. The *SGI* is the ratio of sales in year $t$ to sales in year $t-1$. Growth does not imply manipulation, but growth companies are viewed by professionals as more likely than other companies to commit financial statement fraud, because their financial positions 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 operations in periods of high growth (National Commission on Fraudulent Financial Reporting; Loebbecke, Eining, and Willingham 1989). If growth companies face large stock price losses at the first indication of a slowdown, they may have greater incentives than nongrowth companies to manipulate earnings.

About this point, Fridson (1993) commented:

> Almost invariably, companies try to dispel the impression that their growth is decelerating, since that perception can be so costly to them. (pp. 7–8)

I thus foresaw a positive relationship between the *SGI* and the probability of earnings manipulation.

*Depreciation index*. The *DEPI* is the ratio of the rate of depreciation in year $t - 1$ to the corresponding rate in year $t$. The depreciation rate in a given year is equal to Depreciation/(Depreciation + Net PP&E). A *DEPI* greater than 1 indicates that the rate at which assets are being depreciated has slowed—raising the possibility that the company has revised upward the estimates of assets' useful lives or adopted a new method that is income increasing.[12] I thus expected a positive relationship between the *DEPI* and the probability of manipulation.

*Sales, general, and administrative expenses index*. The *SGAI* is the ratio of sales, general, and administrative expenses to sales in year $t$ relative to the corresponding measure in year $t - 1$. The use of this variable follows the recommendation of Lev and Thiagarajan that analysts interpret a disproportionate increase in sales as a negative signal about a company's future prospects. I expected to find a positive relationship between the *SGAI* and the probability of manipulation.

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

*Total accruals to total assets*. Total accruals were calculated as the change in working capital accounts other than cash less depreciation. Either total accruals or a partition of total accruals was used in prior work to assess the extent to which managers make discretionary accounting choices to alter earnings (see, for example, Healy or Jones). I used total accruals to total assets (*TATA*) to proxy for the extent to which cash underlay reported earnings, and I expected higher positive accruals (less cash) to be associated with a higher likelihood of earnings manipulation.

*Distribution of variables*. The explanatory variables in the model were primarily based on year-to-year changes, which introduces a potential problem when the denominator is small. To deal with this problem, I winsorized the data at the 1 percent and 99 percent percentiles for each variable. In addition, the denominator of the *AQI* variable was sometimes zero because assets in the reference year (period $t - 1$) consisted exclusively of current assets and PP&E. Because the *AQI* was not defined in such cases, I set its value to 1 (its neutral value) instead of treating the observation as missing. Similarly, I set the *DEPI* and *SGAI* to values of 1 when elements of the computation (amortization of intangibles, Compustat #65, and SG&A, Compustat #189) were not available on the Compustat tapes. I found that estimating the model with and without those observations yielded similar results.

Table 2 contains a comparison of the distribution of these variables for manipulators and nonmanipulators 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 than nonmanipulators.

## Results

Panel A of **Table 3** reports the results of the WESML probit and unweighted probit estimations of the model. The likelihood ratio test indicates that for both estimations, with $\chi^2$ statistics ($p$-values) of 34.5 (0.00) and 129.2 (0.00), the model has significant power. The pseudo-$R^2$s of 0.306 and 0.371 for, respectively, WESML and unweighted probit indicate that the model has descriptive validity.

Because the coefficient estimates have similar magnitudes and significance across estimations, I focus on the results of the unweighted probit estimation. The variable *DSRI* has a positive coefficient (0.920) and is significant at the 5 percent level with an asymptotic $t$-statistic of 6.02. This result is consistent with disproportionate increases in receivables, raising the likelihood that a company has inflated revenues. The variable *GMI* has a positive coefficient (0.528) that is more than 2 standard deviations from zero. This result is consistent with companies that face poor prospects having greater incentives for earnings manipulation. The *AQI* also has a significant positive coefficient (0.404, $t$-statistic 3.20), which is consistent with the likelihood of earnings manipulation increasing when companies change their accounting treatment of cost deferral. The *SGI* has a positive coefficient that is more than 5 standard deviations from zero, which is consistent with growth companies that are facing growth deceleration having incentives to manipulate earnings. The *TATA* variable has a significant positive coefficient,

**Table 3. WESML Probit and Unweighted Probit Estimations of the Model**
**($t$-statistics in parentheses except as noted)**

*A. Estimation results*

| | Constant | *DSRI* | *GMI* | *AQI* | *SGI* | *DEPI* | *SGAI* | *TATA* | *LVGI* | Pseudo $R^2$ | $\chi^2$-Statistic ($p$-Value) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Predicted sign | | (+) | (+) | (+) | (+) | (+) | (+) | (+) | (+) | | |
| WESML | –4.954 | 0.789 | 0.459 | 0.306 | 0.701 | 0.033 | –0.006 | 3.937 | –0.264 | 0.306 | 34.50 |
| | (–11.80) | (6.40) | (3.02) | (2.82) | (3.43) | (0.15) | (–0.04) | (3.07) | –(0.83) | | (0.00) |
| Unweighted probit | –4.840 | 0.920 | 0.528 | 0.404 | 0.892 | 0.115 | –0.172 | 4.679 | –0.327 | 0.371 | 129.20 |
| | (–11.01) | (6.02) | (2.20) | (3.20) | (5.39) | (0.70) | (–0.71) | (3.73) | –(1.22) | | (0.00) |

*B. Estimated probabilities of manipulation*

| | WESML Probit | | | | Unweighted Probit | | | |
|---|---|---|---|---|---|---|---|---|
| | Estimation Sample | | Holdout Sample | | Estimation Sample | | Holdout Sample | |
| | Manipulators | Non-manipulators | Manipulators | Non-manipulators | Manipulators | Non-manipulators | Manipulators | Non-manipulators |
| Mean | 0.107 | 0.006 | 0.097 | 0.007 | 0.237 | 0.022 | 0.181 | 0.019 |
| Standard deviation | 0.175 | 0.021 | 0.223 | 0.044 | 0.275 | 0.051 | 0.288 | 0.063 |
| Maximum | 0.851 | 0.615 | 0.999 | 0.999 | 0.980 | 0.960 | 0.999 | 0.999 |
| Median | 0.024 | 0.003 | 0.009 | 0.002 | 0.099 | 0.011 | 0.037 | 0.009 |
| Minimum | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.004 | 0.001 |
| Wilcoxon-Z | 8.049 | | 4.721 | | 8.314 | | 4.630 | |
| ($p$-value) | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| Median $\chi^2$ | 23.785 | | 13.995 | | 26.667 | | 11.056 | |
| ($p$-value) | (0.000) | | (0.003) | | (0.000) | | (0.001) | |

*Notes:* The estimation sample consisted of the pre-1989 manipulators and their controls; the holdout sample consisted of the post-1988 manipulators and controls. The WESML probit assumed that the prior probability of manipulation was 0.0069; the unweighted probit assumed a prior probability of 0.02844. Sensitivity analysis of the prior probability of manipulation yielded coefficient estimates of similar magnitude and significance. When the prior probability of manipulation was specified as 0.0059, 0.0079, 0.0089, and 0.0099, the estimation yielded $\chi^2$ statistics of, respectively, 29.60, 39.22, 43.95, and 48.65, significant at the 1 percent level or lower. The Wilcoxon $Z$ and median $\chi^2$ tests were for the null hypothesis that the estimated probabilities for manipulators and nonmanipulators were drawn from the same distribution. The pseudo-$R^2$ is equal to $(L_\Omega^{2/n} - L_\omega^{2/n})/(1 - L_\omega^{2/n})$, where $L_\Omega$ is the log likelihood for the WESML probit model (unconstrained), $L_\omega$ is the log likelihood with only the constant term in the model (constrained), and $n$ is the number of observations (see Maddala 1983).The log likelihood ratio test statistic (which equals –2 times the difference in the log likelihood of the unconstrained and constrained models) is asymptotically distributed $\chi^2$ with degrees of freedom equal to the difference in the number of parameters of the two models.

which is consistent with manipulators having less cash behind their accounting income.

The coefficients on the *LVGI*, *DEPI*, and *SGAI* variables are not significant. Possibly, these variables are associated with earnings *management*, not manipulation. For example, a change from accelerated to straight-line depreciation or a revision that lengthens assets' useful lives would result in higher values of the *DEPI* but are an instance of earnings management, so the company would not be included in the manipulator sample. Similarly, for the *LVGI* variable, incentives to comply with debt covenants might be insufficient to induce earnings manipulation because the costs of noncompliance are small (Beneish and Press estimated these costs at 1–2 percent of market value of equity).[13]

Panel B of Table 3 shows the estimated probabilities of earnings manipulation for both the estimation and the holdout samples. For the estimation sample, the model estimated by using WESML predicted a higher average (median) probability of earnings manipulation for manipulators, 0.107 (0.024), than for nonmanipulators, 0.006 (0.003). Similarly, the model estimated by using unweighted probit predicted a higher average (median) probability for manipulators, 0.237 (0.099), than for nonmanipulators, 0.022 (0.011). Wilcoxon and median tests rejected the null hypothesis that estimated probabilities for manipulators and nonmanipulators were drawn from the same distribution.

Results for the holdout sample of 24 manipulators and 624 controls were similar to the estimation sample findings. The model predicted that manipulators are, on average, about 10 times more likely to manipulate earnings. The distributions of estimated probabilities for manipulators and nonmanipulators based on unweighted probit illustrate these differences. For example, although not reported in the tables, nearly all the nonmanipulators in the estimation sample (93.4 percent) had an estimated probability of manipulation of less than 0.05 whereas the manipulators had a 38.0 percent probability of manipulating. Similarly, in the holdout sample, 56.1 percent of the nonmanipulators had an estimated probability of manipulation of less than 0.01, compared with 20.8 percent of the manipulators.

The evidence indicates that the probability of manipulation increases with (1) unusual increases in receivables, (2) deteriorating gross margins, (3) decreasing asset quality, (4) sales growth, and (5) increasing accruals.

## Robustness

I assessed the robustness of the results in three ways. First, even though collinearity was not likely to be a problem because none of the 36 Pearson correlation coefficients was greater than 0.25, I dropped up to four variables from the model to assess the stability of the coefficient estimates. Dropping the *DEPI*, *LVGI*, *SGAI*, and *TATA* variables one at a time and in combination yielded results similar to results for the remaining variables.

Second, I assessed the sensitivity of the WESML estimation results to the specification of the prior probability of manipulation. In addition to the estimations based on prior probabilities of 0.0069 and the 0.02844 implicit in unweighted probit (50/1,758), I estimated the model with four other prior probabilities of earnings manipulation; as explained in the notes to Table 3, these estimates yielded similar results in terms of both the descriptive validity of the model and the sign and significance of the coefficient estimates.

Third, although the holdout sample was chosen to be independent from the estimation sample, I assessed the sensitivity of the results to the choice of estimation and holdout samples. To do so, I generated 100 random samples of 50 manipulators and 1,500 controls with a basic random-number generator in SAS and used them to estimate the model 100 times. Similarly, I obtained 100 random holdout samples by treating the complement of 24 manipulators and 832 controls to each random estimation sample as a holdout sample and reproduced the estimated probabilities tests. 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 relationship between the likelihood of manipulation and some financial statement data. 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 Model as a Classification Tool

Because the model distinguishes manipulators from nonmanipulators, I assessed its usefulness as a classification tool. This section discusses the probability cutoffs associated with different costs of making classification errors. The model can make two types of errors: It can classify a company as a nonmanipulator when it manipulates (a Type I error), or it can classify a company as a manipulator when it does not manipulate (a Type II error). The probability cutoffs that minimize the expected costs of misclassification depend on costs associated with

**Table 4. Analysis of Sensitivity of Results to Choice of Sample**

| | Mean | Standard Deviation | Maximum | Median | Minimum | Positive | Significant at 10 Percent Level | Significant at 5 Percent Level | Significant at 2.5 Percent Level |
|---|---|---|---|---|---|---|---|---|---|
| *A. Descriptive statistics for 100 estimation samples* | | | | | | | | | |
| Constant | –4.223 | 0.549 | –3.404 | –4.040 | –5.853 | 0.0% | 100.0% | 100.0% | 100.0% |
| *DSRI* | 0.857 | 0.097 | 1.065 | 0.864 | 0.588 | 100.0 | 100.0 | 100.0 | 100.0 |
| *GMI* | 0.488 | 0.115 | 0.871 | 0.487 | 0.222 | 100.0 | 95.0 | 84.0 | 66.0 |
| *AQI* | 0.453 | 0.113 | 0.789 | 0.438 | 0.223 | 100.0 | 95.0 | 84.0 | 96.0 |
| *SGI* | 0.374 | 0.365 | 1.232 | 0.152 | 0.103 | 100.0 | 100.0 | 100.0 | 100.0 |
| *DEPI* | 0.059 | 0.183 | 0.437 | 0.097 | –0.782 | 81.0 | 37.0 | 18.0 | 10.0 |
| *SGAI* | –0.144 | 0.180 | 0.333 | –0.156 | –0.559 | 25.0 | 30.0 | 12.0 | 4.0 |
| *TATA* | 4.370 | 0.965 | 7.219 | 4.464 | 2.090 | 100.0 | 99.0 | 95.0 | 93.0 |
| *LVGI* | –0.110 | 0.165 | 0.278 | –0.114 | –0.544 | 25.0 | 8.0 | 2.0 | 1.0 |
| Pseudo-$R^2$ | 0.242 | 0.068 | 0.444 | 0.220 | 0.124 | — | — | — | — |
| $\chi^2$-statistic | 89.79 | 19.59 | 142.69 | 84.49 | 51.90 | — | 100.0 | 100.0 | — |
| *B. Descriptive statistics for estimated probabilities on 100 holdout samples* | | | | | | | | | |
| | | | | | | Wilcoxon Z (*p*-value) | Median $\chi^2$ (*p*-value) | | |
| Manipulators | 0.178 | 0.049 | 0.316 | 0.164 | 0.091 | 12.212 | 199.09 | | |
| Nonmanipulators | 0.028 | 0.02 | 0.033 | 0.028 | 0.024 | (0.000) | (0.000) | | |

*Notes*: Significance based on one-tailed test. The Wilcoxon Z and median $\chi^2$ tests were for whether the estimated probabilities for manipulators and nonmanipulators were drawn from the same distribution.

the relative cost of making an error of either type.

Although decision makers' objective functions are not observable, classification error costs are likely to differ by decision maker. For example, investors are likely to have high Type I error costs because the investment loss associated with the discovery of the manipulation is dramatic, whereas their Type II error costs would be low because of the availability of substitutes. A regulator's objective function, however, requires balancing the protection of the investing public against the costs of falsely accusing a company. Their relative costs cannot be measured, but it is likely that their Type II error costs are higher than those of investors.

The costs of Type I and Type II errors are not amenable to objective measurement, but I considered relative costs of Type I to Type II errors ranging from 1:1 to 100:1. (For investors, however, the relevant range is likely between 20:1 and 30:1.) To explain: According to Beneish (forthcoming 1999), the typical manipulator loses approximately 40 percent 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 company's equity appreciates 1–2 percent a quarter, 20–40 nonmanipulators would be needed in the investor's portfolio to offset a single manipulator in that quarter. Therefore, one possibility is that investors would view a Type I error as 20–40 times as costly as a Type II error.

I computed the probability cutoffs that would minimize the expected costs of misclassification; the equation for computing the costs and the results are presented in **Table 5**. The results are similar across estimation methods, so I focus on the unweighted probit estimation in Panel B.

In the estimation sample, at relative error costs of 10:1, the model classified companies as manipulators when the estimated probabilities exceeded 0.0685 (a score greater than –1.49); it misclassified 42 percent of the manipulators and 7.6 percent of the nonmanipulators. At relative error costs of 20:1 or 30:1, the model classified companies as manipulators when the estimated probabilities exceeded 0.0376 (a score greater than –1.78); it misclassified

**Table 5. Probability Cutoffs that Minimize the Expected Costs of Misclassification**

| | | Estimation Sample | | | Holdout Sample | | |
|---|---|---|---|---|---|---|---|
| Relative Costs of Type I and Type II Errors | Cutoff Probability | Type I | Type II | Cost of Model Errors Relative to Naive Strategy[a] | Type I | Type II | Cost of Model Errors Relative to Naive Strategy |
| *A. WESML* | | | | | | | |
| 1:1 | 1.0000 | 1.0000 | 0.0000 | 1.000 | 1.0000 | 0.0000 | 1.000 |
| 10:1 | 0.2905 | 0.9000 | 0.0004 | 0.991 | 0.9166 | 0.0048 | 0.986 |
| 20:1 | 0.0512 | 0.5600 | 0.0409 | 0.855 | 0.7500 | 0.0112 | 0.830 |
| 30:1 | 0.0512 | 0.5600 | 0.0409 | 0.757 | 0.7500 | 0.0112 | 0.804 |
| 40:1 | 0.0223 | 0.4600 | 0.0632 | 0.688 | 0.6667 | 0.0240 | 0.753 |
| 60:1 | 0.0092 | 0.2800 | 0.1329 | 0.597 | 0.5000 | 0.0689 | 0.665 |
| 100:1 | 0.0087 | 0.2600 | 0.1417 | 0.464 | 0.5000 | 0.0753 | 0.608 |
| *B. Unweighted probit* | | | | | | | |
| 1:1 | 1.0000 | 1.0000 | 0.0000 | 1.000 | 1.0000 | 0.0000 | 1.000 |
| 10:1 | 0.0685 | 0.4200 | 0.0761 | 0.680 | 0.6250 | 0.0353 | 0.746 |
| 20:1 | 0.0376 | 0.2600 | 0.1382 | 0.496 | 0.5000 | 0.0721 | 0.623 |
| 30:1 | 0.0376 | 0.2600 | 0.1382 | 0.417 | 0.5000 | 0.0721 | 0.582 |
| 40:1 | 0.0294 | 0.2400 | 0.1747 | 0.433[b] | 0.4583 | 0.0913 | 0.628[b] |
| 60:1 | 0.0294 | 0.2400 | 0.1747 | 0.562[b] | 0.4583 | 0.0913 | 0.896[b] |
| 100:1 | 0.0294 | 0.2400 | 0.1747 | 0.819[b] | 0.4583 | 0.0913 | 1.535[b] |

*Notes*: The expected costs of misclassification, *ECM*, were computed as

$$ECM = P(M)P_I C_I + [1 - P(M)]P_{II}C_{II},$$

where $P(M)$ is the prior probability of encountering earnings manipulators (0.0069 for WESML and 0.02844 for unweighted probit), $P_I$ and $P_{II}$ are the conditional probabilities of, respectively, Type I and Type II errors, and $C_I$ and $C_{II}$ are the costs of Type I and Type II errors. Cutoff probabilities were chosen for each level of relative costs to minimize the expected costs of misclassification as defined in this equation.

[a]The expected cost of misclassification for a naive strategy that classified all companies as nonmanipulators was $0.0069C_1$ for the WESML model and $0.02844C_1$ for the unweighted probit.

[b]In these computations, the naive strategy classified all companies as manipulators. The switch to this naive strategy minimized the expected costs of misclassification because the ratio of relative costs was greater than the population proportion of manipulators. The switch occurred at 40:1 for unweighted probit (> I/0.02844).

26 percent of the manipulators and 13.8 percent of the nonmanipulators. In the holdout sample, at relative error costs of 20:1 or 30:1, the model classified companies as manipulators when the estimated probabilities exceeded 0.0376 (a score greater than –1.78); it misclassified 50 percent of the manipulators and 7.2 percent of the nonmanipulators.

**Figure 1** and **Figure 2** help clarify the classification performance of the unweighted probit model. The figures contain the following information: (1) the probability cutoffs associated with each relative error cost assumption, (2) the percentage of correctly classified manipulators, and (3) the percentage of incorrectly classified nonmanipulators. For the estimation sample in Figure 1, the percentage of correctly classified manipulators ranges from 58 percent to 76 percent whereas the percentage of incorrectly classified nonmanipulators ranges from 7.6 percent to 17.5 percent. For the holdout sample in Figure 2, the percentage of correctly classified manipulators ranges from 37.5 percent to 56.1 percent whereas the percentage of incorrectly classified nonmanipulators ranges from 3.5 percent to 9.1 percent.

These results suggest that, although the model identifies potential manipulators, it does so with large error rates in the range of error costs that are likely to be of most relevance to investors. Thus, because instances of discovered manipulations are rare, a question is raised about whether the model is more useful to investors than a naive strategy that simply classifies all companies as nonmanipulators. Table 5 contains a comparison of the model's expected costs of misclassification with those of the naive strategy.

A naive strategy makes no Type II errors ($P_{II}$ = 0), and the conditional probability of a Type I error ($P_I$) is 1. Thus, the naive strategy's expected costs of misclassification would be $ECM$(naive) = $P(M)C_I$, or $0.0069C_I$ for the WESML comparison and $0.02844C_I$ for the unweighted probit comparison.

For both the estimation and the holdout samples, the model had lower expected misclassification costs than the naive strategy when the cost of a Type I error was greater than that of a Type II error. For example, in the estimation sample in Panel B, the ratios of the cost of the model's errors to the cost of errors from a naive strategy are 0.496 and 0.417 at relative error costs of, respectively, 20:1 and 30:1. Similarly, for the holdout sample, the ratios of the cost of the model's errors to the cost of errors from a naive strategy are 0.623 and 0.582 at relative error costs of 20:1 and 30:1. The evidence suggests that the model is cost-effective in relation to a naive strategy that treats all companies as nonmanipulators.

## Conclusion

Some accounting variables can be used to identify companies that are manipulating their reported earnings. I found that, because manipulation typi-

**Figure 1. Classification Performance of the Unweighted Probit Model: Estimation Sample**

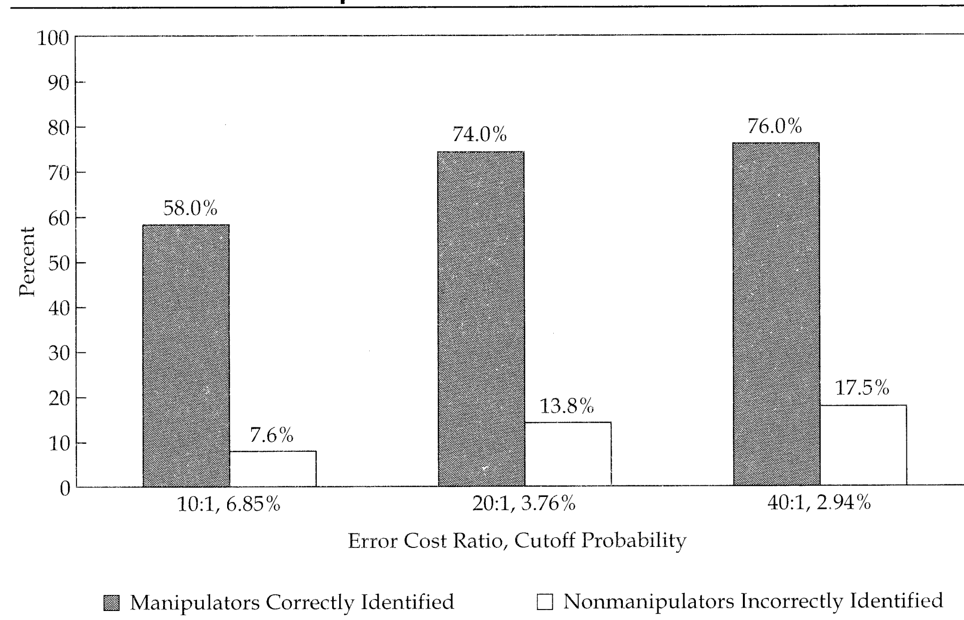
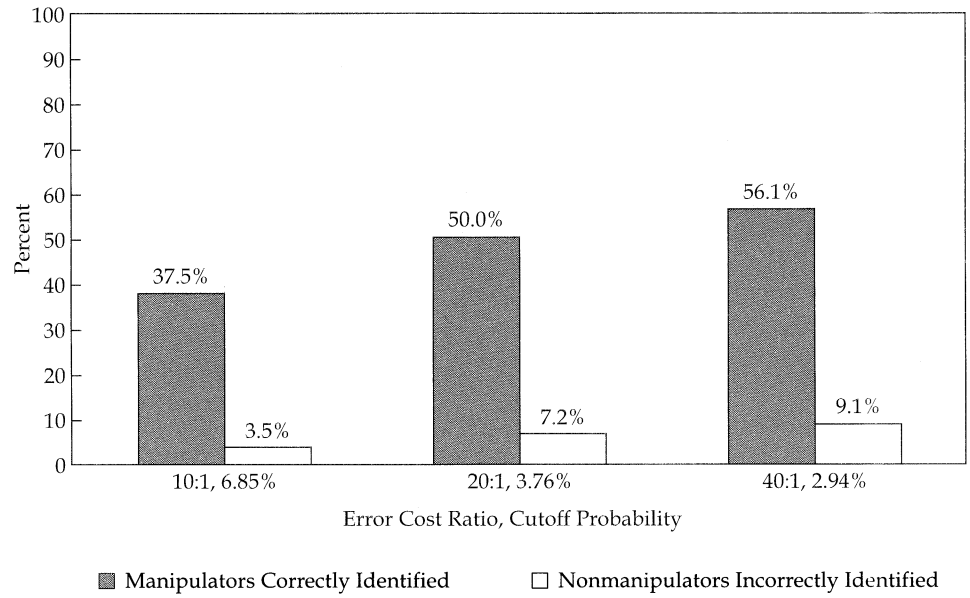


**Figure 2. Classification Performance of the Unweighted Probit Model: Holdout Sample**




cally consists of an artificial inflation of revenues or deflation of expenses, variables that take into account simultaneous bloating in asset accounts have predictive content. I also found that sales growth has discriminatory power: The primary characteristic of sample manipulators was that they had high growth prior to periods during which manipulation was in force.

The evidence presented here was based on a sample of companies whose manipulation of earnings was publicly discovered. Such companies probably represent the upper tail of the distribution of companies that seek to influence their reported earnings—successful and undiscovered manipulators undoubtedly exist—so the evidence should be interpreted in that light.

Given this caution, evidence has been presented here of a systematic association between earnings manipulation and financial statement data that is of interest to accounting researchers and investment professionals. The evidence suggests that accounting data not only meet the test of providing useful information, but they also enable an assessment of the reliability of the reporting. The explicit classification model described here requires only 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 companies and identify potential manipulators for further investigation.

Although the model is cost-effective relative to a strategy of treating all companies as nonmanipulators, its large rate of classification errors makes further investigation of the screening results important. The model's variables exploit distortions in financial statement data that might or might not result from manipulation. For example, the distortions could be the result of a material acquisition during the period examined, a material shift in the company's value-maximizing strategy, or a significant change in the company's economic environment.

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

*I 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, and workshop participants at Duke, the University of Maryland, the University of Michigan, and the University of Quebec at Montreal. 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.*

# Notes

1. Beneish (1997) contains a model for detecting earnings manipulation that differs from this study's model in three ways: That model was estimated with 64 sample companies (versus 74 companies in the present study), the control companies were Compustat companies with the largest unexpected accruals (versus Compustat companies in the same industry in the present study), and the set of explanatory variables in the present study provides a more parsimonious model than the previous model.
2. I eliminated 80 AAERs relating to financial institutions, 15 relating to auditing actions against independent Certified Public Accountants, 9 relating to 10-Q violations that were resolved in annual filings, and 156 relating to companies for which no financial statement data were available on either Compustat, S&P Corporate Text, or 10-K microfiche.
3. The search encompassed the following specific LEXIS-NEXIS databases: *Barron's*, *Business Week*, *Business Wire*, *Corporate Cash Flow*, *Disclosure Online*, *Forbes*, *Fortune*, *Institutional Investor*, *Investor's Business 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 "profits," "earnings," or "income").
4. I eliminated 10 companies for which no financial statement data were available on Compustat, S&P Corporate Text, or 10-K microfiche; 5 financial institutions; and 17 companies mentioned in articles with no discussion of an accounting or disclosure problem. For example, in an article on the manipulation of earnings at Chambers Development, Flynn and Zellner (1992) discussed other companies in the waste management industry, such as Sanifill and Waste Management, without referring to any accounting measurement or disclosure problems.
5. I treated companies in the same industry for which my searches did not identify an instance of manipulation as nonmanipulators. Because successful manipulators would not be identified by the searches, the control sample of 2,332 could contain manipulators. The effect would be to bias the tests against discriminating between manipulators and nonmanipulators.
6. Probit analysis is a form of regression analysis appropriate for cases in which the dependent variable is dichotomous (e.g., the firm either is or is not a manipulator). Probit coefficients are estimated by maximizing a likelihood function and are indicative of how a particular variable affects the probability of a company being a manipulator. WESML probit weighs the likelihood function according to the proportion of earnings manipulators in the sample and in the population to account for state-based sampling. For a discussion of the implications of using state-based samples, see Hsieh, Mansky, and McFadden (1985). Prior research used weighted probit to predict audit qualifications (Dopuch, Holthausen, and Leftwich 1987) and bankruptcy (Zmijewski 1984).
7. Specifically, I used constructs that analysts consider to be indicators of future performance. See, for example, O'Glove (1987); Kellogg and Kellogg (1991); Siegel (1991); Fridson (1993); Lev and Thiagarajan (1993).
8. To examine whether the model could be improved, I also considered, but did not include in the model, five other types of variables: (1) variables to isolate the income effect of nonrecurring items (the ratio of unusual items to pretax income [Compustat #17/#170] and the ratio of net nonoperating income to income before extraordinary items [Compustat #61/#18]), (2) variables to capture the rate and changes in the rate of intangible amortization and variables that identify the funding status of pension funds, (3) cash-flow-based variables (the cash flow adequacy ratio and the cash flow coverage of debt service), (4) signals of earnings quality documented by Lev and Thiagarajan, namely, 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 (5) an index of days' sales in inventory similar to the index of days' sales in receivables. None of these variables improved the model's performance and are not reported here.
9. The variables were not measured contemporaneously with manipulation discovery because, in line with Feroz, Park, and Pastena (1991), manipulation becomes public, on average, 19 months after the end of the fiscal year of the first reporting violation.
10. Manipulation of inventories and other production costs could lead to increasing gross margins, which suggests that either increased or decreased gross margins can increase the likelihood of manipulation. Kellogg and Kellogg stated, "Barring unusual circumstances, the percentage of gross profit to sales should remain the same from year to year" (pp. 10–16), but what "the same" means is difficult to determine. I considered a variable relating gross margin changes to inventory changes, but it did not enhance the specification of the model.
11. Part of the increase might be attributable to acquisitions involving goodwill, but manipulators undertake few acquisitions and those they do undertake are primarily stock-for-stock exchanges that are accounted for by using pooling of interests. Nevertheless, I also calculated the ratio of noncurrent assets other than PP&E and goodwill to total assets and found results similar to those reported here.

12. To allow for the possibility that companies manipulate earnings by using lower depreciation rates than comparable companies in their industry, I used the depreciation rate instead of changes in the depreciation rate to estimate the model. 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.

13. I also considered 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 level leverage variables instead of changes in variables. None of the alternatives attained significance.

## References


Beneish, M.D., and E. Press. 1993. "Costs of Technical Violation of Accounting-Based Debt Covenants." *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.

———. Forthcoming 1999. "Incentives and Penalties Related to Earnings Overstatements That Violate GAAP." *Accounting Review*, vol. 74, no. 4 (October).

Dopuch, N., R.W. Holthausen, and R.W. Leftwich. 1987. "Predicting Audit Qualifications with Financial and Market Variables." *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 & Sons.

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.

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 Chair Arthur Levitt at New York University Center for Law and Business (September 28). Available online at www.sec.gov/news/speeches/spch220.txt

Loebbecke, J.K., M.M. Eining, and J.J. Willingham. 1989. "Auditor's Experience with Material Irregularities: Frequency, Nature, and Detectability." *Auditing*, vol. 9, no. 1 (Fall):1–28.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, U.K.: 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 ed. Englewood Cliffs, NJ: Prentice-Hall.

Watts, R.L., and J.L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: 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.