



## Predicting Accounting Fraud: Evidence from Japan<sup>\*</sup>

MINGZI SONG

*FINANCIAL TECHNOLOGY  
RESEARCH INSTITUTE INC., TOKYO*

NAOTO OSHIRO

*FINANCIAL TECHNOLOGY  
RESEARCH INSTITUTE INC., TOKYO*AKINOBU SHUTO<sup>†</sup>*Graduate School of Economics  
THE UNIVERSITY OF TOKYO*

### ABSTRACT

This study develops a prediction model for identifying accounting fraud by analyzing the accounting information for Japanese firms. In particular, we (1) explore the characteristics of accounting fraud firms by analyzing financial information obtained from annual reports (*yukashoken-houkokusho* in Japanese) and (2) develop a model for predicting accounting fraud based on the characteristics of Japanese fraud firms. To identify the characteristic of fraud firms, we focus on 39 variables for the eight factors of “accruals quality,” “performance,” “nonfinancial measures,” “off-balance-sheet activities,” “market-related incentives,” “conservatism,” “real-activities manipulation,” and “Japanese-specific factors.” Through our univariate analysis and model building process, we find that “accrual quality,” “market-related incentives,” “real-activities manipulation,” “conservatism” and “Japanese-specific factors” are generally useful for detecting accounting fraud. We also conduct several analyses that test the predictive ability of our models, including (1) the detection rates of fraud firms, (2) Type I and Type II error rates, (3) marginal effect analysis on independent variables, and (4) robustness tests on time periods and industry clustering. We find that our models have generally higher predictive power in detecting accounting fraud. We expect that our models can be used widely in various accounting and finance practices.

*JEL Classification:* M41

*Keywords:* Earnings Quality; Accounting Fraud; Accrual Quality; F-Score; Japan

<sup>\*</sup> Acknowledgements: The authors appreciate the helpful comments and suggestions received from Masahiro Enomoto (editor), Hidetoshi Yamaji (editor), anonymous referee, Shota Otomasa, and the participants of the workshops conducted at The University of Tokyo. All errors are the responsibility of the author.

<sup>†</sup> Corresponding Author. *Address* Graduate School of Economics, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN. *Telephone* +81-3-5841-5523 *E-mail* shuto@e.u-tokyo.ac.jp

*Received August 13, 2016; accepted September 21, 2016; available online November 19, 2016 (Advance publications by J-STAGE)*

*DOI: 10.11640/tjar.6.2016.01*

## 1. Introduction

Recent accounting scandals such as Toshiba and Olympus Corporation in Japan have significantly damaged the reliability of accounting information and Japanese capital markets. Exploring the cause and consequence of accounting fraud is of critical importance to the efficient functioning of capital markets (Dechow et al. 2011). In particular, determining how to detect earnings manipulation has been one of the major concerns for both accounting research and business practice (Beneish 1999a; Lee et al. 1999; Ettredge et al. 2006; Brazel et al. 2009; Dechow et al. 2011). This study develops a prediction model for detecting accounting fraud among Japanese firms. Specifically, we explore the characteristics of accounting fraud firms by analyzing financial information obtained from the annual reports (*yukashoken-houkokusho* in Japanese) of Japanese firms and then develop a model for predicting accounting frauds based on the characteristics of Japanese fraud firms. In developing the model, we focus on the quality of earnings such as earnings management and accounting conservatism.

We use two data sources to identify accounting fraud firms in Japan. First, we collect firms that have been accused or had administrative monetary penalties imposed by the Securities and Exchange Surveillance Commission (SESC) for misstatements of financial reports on material issues (SESC firms). Second, we focus on firms that have reported accounting fraud via *the Timely Disclosure Rules* of the stock exchange (TDR firms). As no prior studies have examined accounting fraud in Japanese firms, our subsidiary purpose is to construct a database of accounting fraud.

Beneish (1999a) examines the relationship between financial statement data and earnings manipulation. He provides a probit model to detect earnings manipulation by analyzing eight financial ratios. Prior studies also reveal that some information is useful for detecting accounting fraud, such as accounting accruals (Lee et al., 1999), deferred tax (Ettredge et al., 2006) and nonfinancial measures (Brazel et al., 2009).

Dechow et al. (2011) comprehensively examine the relationship between accounting information and accounting misstatement. They focus on five factors—accruals quality, performance, nonfinancial measures, off-balance-sheet activities, and market-related incentives—for use in identifying misstatements and develop a model for predicting misstatements by analyzing 28 financial characteristics variables. They reveal that the output of their prediction model, *F-Score*, has a strong predictive ability for misstatements.

Although we generally follow the analysis procedure employed by Dechow et al. (2011), we extend their study in several ways. First, we conduct a detailed analysis of accruals quality. Dechow et al. (2011) do not use discretionary accruals to construct their prediction models; we do use them, based on various estimation models. The discretionary accruals are generally used to capture earnings management behavior (Dechow et al. 2010). The research suggests that discretionary accruals are strongly associated with the incidence of accounting fraud (Dechow et al. 1996). In addition to the level of discretionary accruals, we use the absolute value and standard deviation of discretionary accruals in order to capture the effect of accruals reversal.

Second, while earnings management has generally been classified into accrual-based and real earnings management, Dechow et al. (2011) focus only on accrual-based earnings management. They note that an important avenue for future research is seeking a better understanding of the role of real transaction or cash-flow management (Dechow et al. 2011, p. 77). Accordingly, we use real-activities manipulation variables based on Roychowdhury (2006).

Third, we also use the degree of accounting conservatism since recent studies have revealed that accounting conservatism can increase the efficiency of contracts and reduce information asymmetry among investors (Watts 2003). We predict that the use of accounting conservatism is negatively correlated with accounting fraud.

Finally, as most prior studies, including Dechow et al. (2011), examine accounting fraud or misstatement among US firms, we add variables that capture the specific features of Japanese firms. Specifically, we use ownership by financial institution to capture the effect of main banks. We also use ownership by business corporation to proxy for cross-shareholdings.

Thus, in addition to the five factors used by Dechow et al. (2011)—accruals quality, performance, nonfinancial measures, off-balance-sheet activities, and market-related incentives—we focus on three additional factors: real-activities manipulation, accounting conservatism, and Japanese-specific factors. We calculate 38 variables based on the above eight factors and examine the relationship between the variables and accounting fraud.

Our sample consists of 241 SESC firm-year observations and 620 TDR firm-year observations. The key results for the SESC firms, our main concern, are as follows. First, we conduct a univariate analysis to compare variables between fraud and non-fraud firms. The results indicate that most measures of accrual quality are unusually high in fraud years relative to the broad population of firms. We also find that the measures for market-related incentives, real-activities manipulation, accounting conservatism, and Japanese-specific factors are significantly higher for fraud firms than for non-fraud firms.

Second, by extracting the variables that display significant differences in the univariate analysis, we develop a prediction model for detecting accounting fraud. Model 1 includes variables obtained from the annual reports. Model 2 adds variables for market-related incentives. The output of these models is a scaled logistic probability for each firm-year, which we term the *F-Score* following Dechow et al. (2011). We use a backward elimination technique to identify the incremental benefit for predicting misstatement from including information beyond the financial statements. After performing the backward elimination, we retain the following variables in Model 1: soft assets, absolute value of discretionary accruals, new financing, accounting conservatism, real-activities manipulation, and ownership by business corporation. For Model 2, we retain soft assets, absolute value of discretionary accruals, new financing, real earnings management, ownership by business corporation, and book-to-market ratio. The results indicate that many of the variables added to the model of Dechow et al. (2011) are employed in the prediction models, suggesting that our models have generally higher predictive power in detecting accounting fraud.

Finally, we conduct several analyses to test the predictive ability of our models, including (1) the detection rates of fraud firms, (2) Type I and II error rates, (3) marginal effect analysis on independent variables, and (4) robustness tests on time periods and industry clustering. We find that our models have generally higher predictive power to detect accounting fraud. For example, we rank firm-years into five portfolios based on the magnitude of their *F-Score* and report the frequency with which fraud and non-fraud firms fall into each quintile. If our models have better predictive ability in identifying fraud firms, we expect the fraud firms to be clustered in the fifth portfolio (Dechow et al. 2011). Our results indicate that 59.39% of fraud firms are in Quintile 5 for Model 1 and that 90.83% of fraud firms are in Quintile 5 for Model 2. These results are higher than our expected level of 20% and the results obtained in Dechow et al. (2011).

This study significantly contributes to the literature and has implications for accounting

practice. First, it contributes to the accounting fraud literature in its detailed analysis of how earnings quality could improve the explanatory power of a prediction model. While many studies examine the relationship between accruals quality and accounting misstatements (Beneish, 1999a; Lee et al., 1999; Dechow et al., 2011), fewer studies investigate the effect of real earnings management and accounting conservatism on accounting fraud. Our results suggest that a deeper analysis of earnings quality from various dimensions could increase the predictive power of models for detecting accounting fraud.

Second, our results suggest that the institutional features of a country could be important factors in the characteristics of accounting fraud. We examine the Japanese institutional features of ownership by financial institutions and business corporations, and find that corporate ownership has a significant effect on accounting fraud. Xu and Zhang (2009) develop a model for predicting bankruptcy in the Japanese market, finding that incorporating the unique Japanese institutional features of main banks and business groups into their model improves its ability to predict the bankruptcy of Japanese listed companies, consistent with our results.

Our results also have several implications for stakeholders such as auditors, regulators, investors, and other financial statement users. For instance, auditors and regulators could improve the efficiency of their decision making by estimating the probability of accounting fraud, and investors might consider an investment strategy based on the reliability of the available accounting information (i.e., *F-Score*).

The remainder of this paper is organized as follows. Section 2 reviews the research on this topic. Section 3 explains the variable measurements used in this study. Section 4 outlines the sample selection process and the descriptive statistics. Section 5 reports the empirical results on the characteristics of fraud firms and develops a prediction model for identifying accounting fraud. Section 6 summarizes the results of the additional analysis. Finally, section 7 concludes the study with a summary.

## 2. Literature Review

In examining accounting fraud among US firms, most studies focus on firms that have been subject to enforcement actions by the Securities and Exchange Commission (SEC) for allegedly misstating their financial statements. Detailed information on such misstatements is reported in the Accounting and Auditing Enforcement Releases (AAERs) issued by the SEC.

Beneish (1999a) is a pioneering study that develops a model for identifying fraud firms by analyzing financial statement information. He identifies 74 AAER firms operating from 1982 and 1992 and matches the sample to 2,332 Compustat non-fraud samples using two-digit SIC industry and year codes. Using the eight financial statement ratios, he develops a probit model to estimate a likelihood of an earnings overstatement. Beneish (1999a) indicates that five financial ratios have relatively high explanatory power for accounting manipulation: the day's sales in receivables index, gross margin index, asset quality index, sales growth index, and accruals.

Lee et al. (1999) show that the difference between earnings and operating cash flow (i.e., accounting accruals) is larger for fraud firms than for non-fraud firms based on a sample of 56 fraud cases covering 1978 to 1991. Ettredge et al. (2006) reveal that deferred taxes can be useful for predicting misstatements after controlling for other factors for 169 AAER firms. Brazel et al. (2009) investigate whether publicly available nonfinancial measures can be used to assess the likelihood of accounting fraud, showing that non-financial measures such as number of patents,

employees, and products can be effectively used to assess fraud risk in 50 AAER firms.

Okumura (2014) comprehensively investigates misstatements by Japanese firms through analyses on the features of misstatement firms in Japan, the relationship between misstatements and stock price, and the effect of corporate governance on misstatements. Okumura (2014) is closely related to our study in that, in Appendix A, he examines the predictive ability of discretionary accruals for detecting misstatement, and finds a significant relationship between discretionary accruals based on various types of Jones models and misstatements. The results suggest that discretionary accruals are useful for detecting financial misstatements among Japanese firms.<sup>1</sup>

Finally, our research design basically follows the outline of Dechow et al. (2011), the main purpose of which is to analyze the financial characteristics of misstating firms and to develop a model for predicting misstatements. They analyze 494 AAER firms and use 28 financial characteristics variables associated with “accruals quality,” “performance,” “nonfinancial measures,” “off-balance-sheet activities,” and “market-related incentives.”

First, they conduct a univariate analysis to clarify the financial characteristics of misstating firms. Specifically, they compare variables between misstatement and non-misstatement firms. They consider most important the comparison between 494 AAER firms and other firms (COMPUSTAT), which reveals that several factors—such as accruals, the extent of leasing, the number of employees, financing, and prior stock price performance—show significant differences. Second, using the variables displaying significant effects in the univariate analysis, they present models for predicting misstatements. To test their predictive ability, they conduct various analyses. The results generally show that their measure of the likelihood of manipulation (*the F-Score*) is a useful tool for identifying misstatement firms.

### 3. Research Design

#### 3.1 Identification of accounting fraud firms

We form two groups of firms to identify those most likely to engage in accounting fraud: firms that have (1) been accused or had administrative monetary penalties imposed by the Securities and Exchange Surveillance Commission and (2) announced accounting fraud at their timely disclosures.

For (1), we obtain information from the website of the Financial Services Agency (<http://www.fsa.go.jp/policy/kachoukin/index.html>)<sup>2</sup> and denote as the year of fraud the year in which the firms suffered administrative monetary penalties or the equivalent (SESC firms), excluding incidents such as misstatements of the register of shareholders, unfair financing, or insider trading, as these types of fraud are not associated with accounting fraud.

For (2), we identify firms that have announced improper accounting via the timely disclosure required by the stock exchange (TDR firms). We searched for these firms using keywords expressing suspicion of accounting fraud in timely disclosures, available from TDnet (a timely

<sup>1</sup> Our study extends the analyses in Okumura (2014) as follows. First, while Okumura (2014) focuses exclusively on discretionary accruals for detecting accounting fraud, we develop a better prediction model based on all the variables available from annual reports. Second, our sample includes, in addition to restatement firms, firms that have been accused, or have had administrative monetary penalties imposed by, the Securities and Exchange Surveillance Commission (SESC), thereby constituting a more reliable sample for detecting accounting fraud.

<sup>2</sup> The last access date on this web site was July 21, 2015.

disclosure information browsing service) on the Tokyo Stock Exchange homepage and identified the year when the accounting fraud occurred.<sup>3</sup>

We use sets of four keywords (in Japanese) for this search. First, we use *teisei* (“correction”), *tekisetsu* (“appropriate”), *fusei-kaikei* (“accounting fraud”). These keywords are expected to extract all firms that have corrected their *yukashoken-houkokusho* (“annual report”) or *kessan-tanshin* (“earnings briefing”). Then, we clarify the reason for the correction by reviewing each report and identifying the firms that have corrected an earnings briefing due to accounting fraud. We exclude firms that announced a correction as being due to a simple error. Second, we use *chien* (“delay”), *enki* (“postpone”), and *kanri* (“supervision”) as keywords to find firms that delayed their submission of financial statements or earnings briefings. These words are used to identify the firms that delayed submission due to accounting fraud or improper sales. Third, we use *iinkai* (“committee”), *cyousa* (“investigation”), and *setchi* (“installation”/“establishment”) as keywords to extract the firms for which third-party committees were established due to alleged accounting fraud. Most of the firms that have carried out accounting fraud are identified through these keywords. Finally, we use *kadai* (“excessive”), *fusei* (“fraud”), and *syori* (“process”) to extract the firms against which allegations have been made with respect to the adequacy of their accounting or cheating by employees.

For the firms extracted using the groups of four keywords, we impose further screening under the following conditions. First, we select the firms for which earnings are manipulated upward through accounting fraud with respect to accounts on the balance sheet or income statements. We focus on net assets in the balance sheet and net income in the income statements. Second, we exclude cases such as mistakes of the unit on account, modifications of explanatory materials, revisions of management earnings forecasts, and modifications of segment information, as they are likely to have little association with accounting fraud.

While the first selection criterion (i.e., SESC firms) is more important, both selection criteria have unique features. The advantage of using SESC firms is that they may provide more reliable fraud samples, as they were fully investigated by the committee. One disadvantage is that the SESC may not be able to investigate all fraud firms because such investigations are expensive. Furthermore, a sample selection bias may occur due to some particular tendency in their research policy; for example, the committee might tend to investigate larger firms more often than smaller ones.<sup>4</sup>

The benefit of the second selection criteria (i.e., TDR firms) is that it can collect a larger number of fraud samples than the SESC firm criterion can. However, TDR firms, extracted through restatement announcements by timely disclosure, might include corrections of simple, unintentional, and minor errors. Thus, we identify the firms that are more likely to commit accounting fraud by sorting TDR firms according to the magnitude of the restatement (the details are provided below).

We use the abovementioned sample selection criteria so that one could complement the other and compensate for its shortcomings. Though the samples selected by the first criterion are more reliable, they may not include latent fraud firms. Thus, we also use the second criterion to select a larger number of fraud firms. We are also interested in seeing how the results differ

---

<sup>3</sup> When we cannot identify the exact year of the accounting fraud, we assume that it occurred in the latest fiscal year, and all prior years are omitted as non-discriminable.

<sup>4</sup> Dechow et al. (2011) also discuss the same limitation, whose research handles samples alleged by SEC.

TABLE 1 TYPES OF DEPENDENT VARIABLE

Type	Definition	Number	Total Number
Misstate1	Firms that have alleged or imposed administrative monetary penalty from the Securities and Exchange Surveillance Commission	241	241
Misstate2	Firms that have announced their accounting fraud at the timely disclosures	620	861
Misstate3	Firms with correction ratio of net income $\geq 0.05$ and correction ratio of net assets $\geq 0.01$ in Misstate2	99	340
Misstate4	Firms with correction ratio of net income $\geq 0.10$ and correction ratio of net assets $\geq 0.02$ in Misstate2	72	313
Misstate5	Firms with correction ratio of net income $\geq 0.15$ and correction ratio of net assets $\geq 0.03$ in Misstate2	62	303
Misstate6	Firms with correction ratio of net income $\geq 0.20$ and correction ratio of net assets $\geq 0.03$ in Misstate2	58	299
Misstate7	Firms with correction ratio of net income $\geq 0.25$ and correction ratio of net assets $\geq 0.04$ in Misstate2	43	284
Misstate8	Firms with correction ratio of net income $\geq 0.30$ and correction ratio of net assets $\geq 0.04$ in Misstate2	42	283

*Note:* Misstate1 includes firms that have been accused or had administrative monetary penalties imposed by the SESC. Misstate2 includes firms identified by performing a plurality of keyword searches related to accounting fraud at the timely disclosure. Misstate3 to Misstate8 calculate the correction ratio by the accounting fraud of net income and net assets and are defined by correction ratio increases through Misstate3 to Misstate8. For more information on variable definitions, please refer to Section 3.1.

among the subsamples based on the two criteria. Specifically, we use the variables described below as a proxy for accounting fraud.

Misstate1 comprises SESC firms, our primary concern. Table 1 provides details on the variables for accounting fraud firms. Misstate1 has 241 firm-year observations. Misstate2 comprises firms that announce accounting fraud by timely disclosure (i.e., TDR firms). Misstate2 contains 620 firm-year observations.<sup>5</sup> We set variables through Misstate3 to Misstate8 in order to gradually capture the degree of accounting fraud. To measure the variables, we focus on the total amount of correction in net earnings (net assets)<sup>6</sup> through restatements of financial statements and calculate the ratio of misstatements to reported earnings (reported net assets).<sup>7</sup>

We find that the number of subsamples decreases from Misstate3 through to Misstate8,

<sup>5</sup> We included several firms that announced their improper accounting through their timely disclosure in Misstate1 since they are strongly suspected of conducting accounting fraud. The firms are defined as having gone bankrupt or being delisted from the stock market before the submission of restatements. These observations total seven firm-years.

<sup>6</sup> Specifically, we include the corrections of the net assets, caused by manipulation of assets and/or liabilities.

<sup>7</sup> We calculate two ratios, (pre-correction of net income – post-correction of net income)/5-year average of pre-correction of net income and (pre-correction of net assets – post-correction of net assets)/pre-correction of net assets. The former is called a “correction ratio of net income,” and the latter is called a “correction ratio of net assets.” We classify the sample into six subsamples (Misstate3 to Misstate8) with 0.05 and 0.01 intervals of net income and net assets, respectively.

suggesting that the number decreases as the impact of accounting fraud increases. In the later analyses, the misstatement variables are defined as the sum of observations in Misstate1 and from Misstate2 to Misstate8. For example, the Misstate8 variables include a total of 283 observations, the sum of Misstate1's 241 and Misstate8's 42.

### 3.2 Measurement of explanatory variables

This section describes the variables used to identify accounting fraud firms. The variables are based on those used in Dechow et al. (2011), including the factors of accruals quality, performance, nonfinancial measures, off-balance-sheet activities, and market-related incentives.<sup>8</sup> We use two additional factors, conservatism and real-activities manipulation, in order to capture the quality of earnings more comprehensively. We also use discretionary accruals based on various estimation models since they are a major proxy for earnings management behaviors. Furthermore, because our sample consists of Japanese firms, we add Japanese-specific factors as variables. Details on the definition of each variable are provided in Table 2.<sup>9</sup>

#### (1) Accrual quality-related variables

Accruals and discretionary accruals are traditional variables used to capture earnings management behaviors (Dechow et al. 2010; Shuto 2010). Dechow et al. (1996) find that firms alleged by the SEC to have committed GAAP violations via earnings manipulation also engaged in earnings management within GAAP in the form of accruals management. These results suggest a relationship between accrual quality and accounting fraud.

First, we employ working capital accruals (*WC accruals*) as a proxy for an accruals variable. *WC accruals* reflect short-term accruals relating to working capital (Teoh et al. 1998; Allen et al. 2013). Next, we use *RSST accruals* as an extended definition of *WC accruals* to include changes in long-term operating assets and long-term operating liabilities (Richardson et al. 2005). We also examine two accruals components, *Change in receivables* and *Change in inventory*. These variables are expected to have higher flexibility for management among accruals-related accounts (Richardson et al. 2005), and greater variable values are likely to reflect a higher probability of accounting fraud.

We also examine *%Soft assets*, defined as the percentage of assets on the balance sheet that are neither cash nor P&E (Dechow et al. 2011). Barton and Simko (2002) show that firms with greater net operating assets have more accounting flexibility with which to manage earnings. We assume that *%Soft assets* has a positive relationship with the probability of accounting fraud since a firm with higher *%Soft assets* has greater accounting flexibility with which to meet short-term earnings goals.

Finally, we examine discretionary accruals on the basis of three estimation models: the

<sup>8</sup> The definition of the variables used in this study basically follows that of Dechow et al. (2011) to ensure comparability between the two studies.

<sup>9</sup> Although the variables used in this study basically follow those used in Dechow et al. (2011), we do not use the following variables: (1) earnings quality, as in Dechow and Dichev (2002), (2) differed tax expense, and (3) pension plan assets. Variables (2) and (3) were excluded because our database does not contain these data. Variable (1) requires information in the next year ( $t+1$ ) in order to measure the earnings quality in the current year ( $t$ ) and therefore cannot be used, since we are building a prediction model. Dechow et al. (2011) also excluded this variable for the same reason.



TABLE 2 VARIABLE DEFINITIONS

Variable	Pred sign	Calculation
<i>Misstate</i>		Indicator variable equal to 1 for accounting fraud firm-years and zero otherwise
<b>Accruals quality related variables</b>		
<i>WC accruals</i>	+	$((\Delta \text{Current Assets} - \Delta \text{Cash and Short-term Investments}) - (\Delta \text{Current Liabilities} - \Delta \text{Debt in Current Liabilities} - \Delta \text{Taxes Payable})) / \text{Average total assets}$
<i>RSST accruals</i>		$\text{RSST} = (\Delta \text{WC} + \Delta \text{NCO} + \Delta \text{FIN}) / \text{Average total assets}$
	+	$\text{WC} = (\text{Current Assets} - \text{Cash and Short-term Investments}) - (\text{Current Liabilities} - \text{Debt in Current Liabilities}),$
	+	$\text{NCO} = (\text{Total assets} - \text{Current Assets} - \text{Investments and Advances}) - (\text{Total Liabilities} - \text{Current Liabilities} - \text{Long-term Debt}),$
	+	$\text{FIN} = (\text{Short-term Investments} + \text{Long-term Investments}) - (\text{Long-term Debt} + \text{Debt in Current Liabilities})$
<i>Change in receivables</i>	+	$\Delta \text{Accounts Receivable} / \text{Average total assets}$
<i>Change in inventory</i>	+	$\Delta \text{Inventory} / \text{Average total assets}$
<i>% Soft assets</i>	+	$(\text{Total Assets} - \text{PP\&E} - \text{Cash and Short-term Investments}) / \text{Total Assets} \times 100$
<i>MJ discretionary accruals</i>	+	Discretionary accruals following Dechow et al. (1995)
<i>MJ discretionary accruals SD</i>	+	For more information on the estimation methods, see Appendix A The standard deviation of the past five years of MJ discretionary accruals
<i>MJ discretionary accruals AB</i>	+	The absolute value of the MJ discretionary accruals
<i>PM discretionary accruals</i>	+	Discretionary accruals following Kothari et al. (2005)
<i>PM discretionary accruals SD</i>	+	For more information on the estimation methods, see Appendix A The standard deviation of the past five years of PM discretionary accruals
<i>PM discretionary accruals AB</i>	+	The absolute value of the PM discretionary accruals
<i>CFO discretionary accruals</i>	+	Discretionary accruals following Kasznik (1999)
<i>CFO discretionary accruals SD</i>	+	For more information on the estimation methods, see Appendix A The standard deviation of the past five years of CFO discretionary accruals
<i>CFO discretionary accruals AB</i>	+	The absolute value of the CFO discretionary accruals
<b>Performance variables</b>		
<i>Change in cash sales</i>	-	$\text{Change in cash sales} = (\text{cash sales}_t / \text{cash sales}_{t-1} - 1) \times 100$ $\text{cash sales} = \text{Sales} - \Delta \text{Accounts Receivable}$
<i>Change in cash margin</i>	-	$\text{Change in cash margin} = (\text{cash margin}_t / \text{cash margin}_{t-1} - 1) \times 100$ $\text{cash margin} = 1 - (\text{Cost of Good Sold} - \Delta \text{Inventory} + \Delta \text{Accounts Payable}) / (\text{Sales} - \Delta \text{Accounts Receivable})$
<i>Change in free cash flows</i>	-	$\text{Change in free cash flows} = (\text{free cash flows}_t - \text{free cash flows}_{t-1}) / \text{Average total assets}$ $\text{free cash flows} = \text{Earnings} - \text{RSST}$
<b>Nonfinancial variables</b>		
<i>Abnormal change in employees</i>	-	$(\text{Number of employees}_t / \text{Number of employees}_{t-1} - 1) \times 100 - (\text{Total Assets}_t / \text{Total Assets}_{t-1} - 1) \times 100$
<i>Abnormal change in order backlog</i>	-	$(\text{Order backlog}_t / \text{Order backlog}_{t-1} - 1) \times 100 - (\text{sales}_t / \text{sales}_{t-1} - 1) \times 100$

TABLE 2 (CONTINUED)

Variable	Pred sign	Calculation
<b>Off-balance-sheet variables</b>		
<i>Existence of operating leases</i>	+	Indicator variable equal to one when future operating lease obligations are greater than zero, and zero otherwise.
<i>Change in operating lease activity</i>	+	$((\text{future operating lease obligations}_t / \text{Average total assets}_t) / (\text{future operating lease obligations}_{t-1} / \text{Average total assets}_{t-1}) - 1) \times 100$
<b>Market-related incentives</b>		
<i>Ex ante financing need</i>	+	Indicator variable equal to one when $((\text{CFO} - \text{past three years average capital expenditures}) / (\text{Current Assets})) < -0.5$ , zero otherwise.
<i>Actual issuance</i>	+	Indicator variable equal one when $(\text{issuance of stock}) > 0$ or $(\text{commercial paper} + \text{issuance of bond}) > 0$ , zero otherwise.
<i>CFF</i>	+	Financing activities net cash flows / Average total assets
<i>Leverage</i>	+	Long-term Debt / Total assets
<i>Market adjusted stock-return</i>	+	$(\text{Return}_t - \text{Return}_{t-1}) / \text{Return}_{t-1} - (\text{TOPIX}_t - \text{TOPIX}_{t-1}) / \text{TOPIX}_{t-1}$
<i>Lagged market adjusted stock-return</i>	+	Market adjusted stock-return in the previous year
<i>Book-to-market</i>	-	Equity / Market Value
<i>Earnings-to-price</i>	-	Earnings / Market Value
<b>Accounting conservatism variables</b>		
<i>C score RR</i>	+	Cscore following Khan and watts(2009), estimated by use of the raw return. See Appendix B for details of estimation method,
<i>C score RRSD</i>	+	The standard deviation of the past five years C score RR
<i>C score AR</i>	+	Cscore following Khan and watts(2009), by used of the abnormal return. See Appendix B for details of estimation method.
<i>C score ARSD</i>	+	The standard deviation of <i>C score AR</i> in the past five years
<b>Real activities manipulation variables</b>		
<i>AB cash flow</i>	-	Abnormal cash flows following Roychowdhury(2006) See Appendix C for details of estimation method.
<i>Discretionary expense</i>	-	Abnormal Discretionary expense following Roychowdhury(2006) See Appendix C for details of estimation method.
<i>AB product cost</i>	+	Abnormal product cost following Roychowdhury(2006) See Appendix C for details of estimation method.
<b>Japanese specific factors-related variables</b>		
<i>FIN</i>	-	Number of shares held by financial institutions / Number of shares outstanding $\times 100$
<i>CORP</i>	+	Number of shares held by the other corporation / Number of shares outstanding $\times 100$

*Note:* Predicted sign shows the expected sign of the relationship between the occurrence of accounting fraud and each independent variable.

modified Jones model (*MJ discretionary accruals*) of Dechow et al. (1995); the performance-matched discretionary accruals model (*PM discretionary accruals*) of Kothari et al. (1999); and the CFO modified Jones model (*CFO discretionary accruals*) of Kasznik (1999). We estimate the parameters of each model using cross-sectional regression by industry-year. Details on the estimation methods for the models are elaborated in Appendix A. Earnings management research generally assumes that, while positive discretionary accruals imply income-increasing earnings management, negative discretionary accruals imply income-decreasing earnings

management.

We also examine the standard deviation of the discretionary accruals (*discretionary accruals SD*) and the absolute value of the discretionary accruals (*discretionary accruals AB*). Managers are likely to use a large amount of discretionary earnings when they have a strong incentive to conduct aggressive earnings management. Such earnings management causes a reversal of accruals in later years and amplifies the standard deviation of discretionary accruals and earnings. Therefore, we examine the absolute values and their standard deviation over the last five years in order to capture the effect of such accruals reversals.<sup>10</sup> We expect that the variables related to discretionary accruals have a positive relationship with the probability of accounting fraud.

## (2) Performance variables

We focus on a set of variables concerning a firm's financial performance to examine whether managers engage in accounting fraud to mask their deteriorating performance (Dechow et al. 1996; Dechow et al. 2011; Beneish 1997, 1999b). The first variable is change in cash sales (*Change in cash sales*) and change in cash margin (*Change in cash margin*). *Change in cash sales* is defined as sales amount after accruals-based sales are excluded. *Change in cash margin* is equal to cash sales less cash cost of goods sold; the influence of accruals such as receivables and inventory are excluded from this variable. We assume that, when these performance variables decline, managers are more likely to inflate earnings by boosting accruals.

A change in ROA (*Change in return on assets*) is an important variable for managers who are conscious about growth of earnings (Graham et al 2005). Therefore, increased ROA during an accounting fraud period might be caused by a manager's earnings management. *Change in free cash flows* is a more fundamental measure than earnings because this variable is not affected by accruals. We posit that managers have an incentive to increase earnings when free cash flow is decreasing.

## (3) Nonfinancial variables

We examine two nonfinancial measures. First, we use number of employees because managers attempting to mask deteriorating financial performance will reduce employees in order to boost earnings (Brazel et al. 2009). We assume that the change in the number of employees (*Abnormal change in employees*) has a negative relationship with the probability of accounting fraud.

The second variable is the order backlog (*Abnormal change in order backlog*). Decreasing the order backlog is indicative of lower future sales and earnings (Rajgopal et al. 2003). Therefore, we assume that managers facing a decreasing backlog have an incentive to engage in accounting fraud.

## (4) Off-balance-sheet variables

---

<sup>10</sup> In particular, our assumption is as follows. Panel C in Table 3 indicates that fraud firms tend to conduct accounting fraud over multiple years. Because firms are not able to continue to increase earnings by managing accruals, fraud firms are likely to report large negative discretionary accruals because of accruals reversal during the fraud period. Further, firm managers might conduct illegal accounting fraud after implementing income-increasing earnings management within GAAP, resulting in large negative discretionary accruals in fraud years. This suggests that the dispersion of discretionary accruals among fraud firms will be greater around the fraud year.

The most important off-balance-sheet financing variable is operating lease (Dechow et al. 2011). Accounting for operating leases allows firms to record lower expenses in the early period of the lease term. Thus, managers who are excessively conscious of window-dressing are more likely to increase operating lease activity. Further, we expect that the use of an operating lease increases during accounting fraud. Consequently, we examine two variables: the use of the operating lease (*Existence of operating lease*) and its change (*Change in operating lease activity*).

#### (5) Market-related incentive variables

One of the major incentives for earnings management is maintaining a high stock price (Dechow et al. 2011). We investigate two motivations relating to stock price (Dechow et al. 2011). The first motivation is to keep a high stock price in order to reduce the cost of raising new equity. We use four variables to capture this kind of motivation. First, we use an indicator variable identifying whether the firm has issued new equity or debt (*Actual issuance*). The second variable is the net amount of new financing raised (*CFF*). The third variable is a need of ex ante financing (*Ex ante financing need*), as some firms may have wished to raise new capital but could not because they were unable to secure favorable terms. The fourth variable is financing leverage (*Leverage*). Firms with higher leverage will have incentives to boost financial performance to both satisfy financial covenants in existing debt contracts and raise new debt on more favorable terms.

The second managerial motivation to maintain a high stock price is management compensation tied to stock price performance. Managers might also be forced to retire when stock prices decrease. Because managers are conscious of stock-based performance, they are likely to have an incentive to increase earnings. Here, we examine three variables: market adjusted stock return (*Market adjusted stock-return*), book-to-market (*Book-to-market*), and PER (*Earnings-to-price*). Thus, we expect that while *Market adjusted stock-return* has positive sign, *Book-to-market* and *Earnings-to-price* have negative signs.

#### (6) Accounting conservatism variables

The variables described so far basically follow the variables employed by Dechow et al. (2011). Although Dechow et al. (2011) focused on variables relating to accrual qualities as proxies for earnings quality, we also focus on other quality measures: “conservatism” and “real activities manipulation.” Conservatism reflects accountants’ tendency to require a higher degree of verification to recognize good news than to recognize bad news in financial statements (Basu 1997). Higher conservatism is thus identified when the degree of verification needed to recognize good news (positive earnings) is higher than what is needed to recognize bad news (negative earnings).

Watts (2003) argues that adopting accounting conservatism, which tends to produce conservatively measured earnings and net assets, could facilitate contract efficiency and provide reliable information to the securities market. Thus, we expect that a higher degree of conservatism improves earnings quality and restricts occurrences of accounting fraud.

We employ the method of Khan and Watts (2009) to measure the degree of conservatism. Their model is an extended version of Basu (1997) with cross-sectional regression, allowing it to measure conservatism for an individual firm. In the model, stock price returns are used as a proxy for economic loss. We measure two types of conservatism on the basis of raw stock return (*C score RR*) and market-adjusted stock return (*C score AR*). Similar to the research method for discretionary accruals, we also examine the standard deviation (*C score RRSD*) and absolute value

(*C score ARSD*) of these variables. Details on the calculation of conservatism are provided in Appendix B.

(7) Real activities manipulation variables

Earnings management is divided into two types: accrual-based earnings management and real earnings management. Accrual-based earnings management constitutes the discretionary behaviors that occur through the discretion of recognition and estimates in accrual accounting, such as fictional accounting adjustments without a change in cash flow. We have already discussed the variables of accrual-based earnings management in subsection (1).

Real earnings management constitutes discretionary behaviors performed to manage earnings by altering a firm's real economic activities, such as a reduction in R&D activities. We employ methods proposed by Roychowdhury (2006) to comprehensively capture a manager's real earnings management.

Roychowdhury (2006) classifies a firm's real earnings management into three categories: 1) sales manipulation, 2) overproduction, and 3) reduction of discretionary expenditures. The first manipulation increases annual sales temporarily by discounting or altering sales term, causing an abnormally low cash flow over sales amount. The second manipulation boosts earnings by reducing total cost of production through excessive overproduction. In the third manipulation, managers discretionally adjust a certain amount of accrued costs (such as for R&D or advertising), resulting in abnormally high discretionary expenditures over sales amounts.

To capture these influences of real earnings management, we examine an abnormal operating cash flow (*AB cash flow*), abnormal discretionary expense (*Discretionary expense*), and abnormal product cost (*AB product cost*). Details on these variables are provided in Appendix C.<sup>11</sup>

(8) Japanese-specific factors-related variables

We examine Japanese-specific factors related to occurrences of accounting fraud. Xu and Zhang (2009) developed a bankruptcy prediction model for Japanese firms and found that the model's performance improved when they added shareholdings by financial institutions such as main banks and a cross-shareholding among business corporations. We thus expect that these factors are useful for detecting accounting fraud since anecdotal evidence suggests that firms facing bankruptcy tend to commit it.

Shuto (2010) also finds that, while firms with higher ownership by financial institutions are not likely to conduct earnings management to avoid earnings reductions, firms with higher cross-shareholding tend to engage in earnings management to attain short-earnings targets. These results suggest that shareholdings by financial institutions restrict managers' opportunistic behaviors through shareholders' strict monitoring and that cross-shareholdings do not function as a mutual monitoring system and do not prevent earnings management.

Thus, we analyze ownership by financial institutions (*FIN*) and cross-shareholdings (*CORP*) as Japanese-specific factors. We assume that *FIN* (*CORP*) has a negative (positive) relationship with the probability of accounting fraud.

---

<sup>11</sup> Our real earnings management variables are calculated using un-restated (i.e., manipulated) data, as described in footnote 10. Thus, we cannot deny the possibility that the variables might not reflect the firm's real economic activity.

## 4. Sample selection and statistics

### 4.1 Sample selection

We obtained our initial samples from all the listed companies on the Japanese stock market for the period from 2000 to 2014, and eliminated the following kinds of firms:

- (1) financial services companies,
- (2) companies adopting an accounting standard other than J-GAAP,
- (3) companies without sufficient data to compute the *F-Score*.

The sample selection procedure leaves us with 65,440 firm-year observations. As shown in Table 1, the SESC firms (i.e., Misstate1) provide 241 observations for 83 firms. The TDR firms, who announced possible improper disclosures on their financial statements, provide 620 observations for 89 firms. As described in section 3, Misstate2 includes 861 observations for 172 firms.

The data used in this analysis are obtained from the Nikkei NEEDS financial statements database and the Nikkei NEEDS market database. Special treatment is needed when analyzing the financial statements of fraud firms. The financial statements included in the database were immediately overwritten retroactively once restatement data became available from the firm. Since the purpose of this research is to build a prediction model, these overwritten data are not suitable for our use. We thus restored all the restated data to the original statements.<sup>12</sup> We use these restored datasets for all analyses of the fraud firms.

### 4.2 Features of the fraud samples

As a preliminary analysis, we first observe the features of the fraud firms. Table 3 summarizes the Misstate1 sample. Panel A shows the distribution of samples sorted by year. Most fraud samples occur between 2003 and 2012, around the middle of our sample period. As one of the reasons for this biased distribution, we might point out that no SESC investigations took place before 2000 since the framework of the SESC's administrative monetary penalty was established in 2005. Moreover, it is possible that there are fewer samples after 2013 because few SESC investigations have occurred recently; the number of fraud firms may increase in the coming years.

Panel B indicates the distribution sorted by industry. The panel shows that there are more fraud firms in *Service*, *Trading*, *Electronics*, and *Construction*. We also find that the percentage of fraud firms is high in *Warehouse*, *Power generation*, and *Precision machinery*. Panel C presents the number of restatements caused by fraud statements. A total of 32.53% of fraud firms have restated their financial statements once, and 67.47% restated them more than twice. This result suggests that most of the fraud firms have restated their financial statements.

Table 4 summarizes the contents of the misstatements—thus, the methods of accounting fraud. The table shows that the overstatement of assets is the most popular kind, including accounting receivables, inventory, and work-in-process. The magnitude of the misstatement relating to these assets is 47.15% of total revenue on average, and the number of misstatements

---

<sup>12</sup> We first collected the original financial statements as pre-restated statements, and restored all restated figures to the original ones in the database. By contrast, Dechow et al. (2011) do not restore their database (Compustat) even for a company with an amended 10-K because they find via a random sampling test that these misstatement companies are less likely to file amended financial statements.

TABLE 3 FREQUENCY OF POPULATION BY CALENDAR YEAR, INDUSTRY, AND ACCOUNTING FRAUD FIRM-YEAR

*Panel A: Frequency of the population by calendar year*

year	Accounting fraud firm-years	Non-accounting fraud firm-years	Total firm-years	Percentage (Total firm-years)	Percentage (Accounting fraud firm-years)
2000	1	4,603	4,604	7.04%	0.02%
2001	2	4,717	4,719	7.21%	0.04%
2002	6	4,783	4,789	7.32%	0.13%
2003	13	4,787	4,800	7.33%	0.27%
2004	19	4,697	4,716	7.21%	0.40%
2005	30	4,658	4,688	7.16%	0.64%
2006	36	4,612	4,648	7.10%	0.77%
2007	29	4,573	4,602	7.03%	0.63%
2008	30	4,315	4,345	6.64%	0.69%
2009	25	4,179	4,204	6.42%	0.59%
2010	17	4,068	4,085	6.24%	0.42%
2011	11	3,949	3,960	6.05%	0.28%
2012	14	3,827	3,841	5.87%	0.36%
2013	6	3,754	3,760	5.75%	0.16%
2014	2	3,677	3,679	5.62%	0.05%
Total	241	65,199	65,440	100.00%	0.37%

*Panel B: Frequency of the population by industry*

Industry	Accounting fraud firm-years	Non-accounting fraud firm-years	Total firm-years	Percentage (Total firm-years)	Percentage (Accounting fraud firm-years)
Foods	7(3)	2,485	2,492	3.81%	0.28%
Textiles and Apparels	4(1)	994	998	1.53%	0.40%
Pulp and Paper		473	473	0.72%	0.00%
Chemicals		3,445	3,445	5.26%	0.00%
Pharmaceutical		1,016	1,016	1.55%	0.00%
Oil		191	191	0.29%	0.00%
Rubber Products		345	345	0.53%	0.00%
Ceramics		1,184	1,184	1.81%	0.00%
Iron and Steel		886	886	1.35%	0.00%
Nonferrous Metals	2(1)	2,327	2,329	3.56%	0.09%
Machinery	13(5)	4,146	4,159	6.36%	0.31%
Electronics	24(6)	4,813	4,837	7.39%	0.50%
Shipbuilding		100	100	0.15%	0.00%
Automotive	8(1)	1,371	1,379	2.11%	0.58%

*(Continued)*

*(-Continued)*

Transportation Equipment		237	237	0.36%	0.00%
Precision Machinery	8(2)	947	955	1.46%	0.84%
Other Products		2,069	2,069	3.16%	0.00%
Fishery		160	160	0.24%	0.00%
Mining		182	182	0.28%	0.00%
Construction	24(7)	3,404	3,428	5.24%	0.70%
Trading	29(12)	6,807	6,836	10.45%	0.42%
Retail	14(5)	4,716	4,730	7.23%	0.30%
Real Estate	4(3)	2,960	2,964	4.53%	0.13%
Rail and bus		1,194	1,194	1.82%	0.00%
Land Transportation		639	639	0.98%	0.00%
Marine Transportation		379	379	0.58%	0.00%
Air Transportation		157	157	0.24%	0.00%
Warehouse	10(2)	806	816	1.25%	1.23%
Telecommunications	8(1)	1,313	1,321	2.02%	0.61%
Power Generation	4(1)	203	207	0.32%	1.93%
Gas		314	314	0.48%	0.00%
Services	82(33)	14,936	15,018	22.95%	0.55%
Total	241(83)	65,199	65,440	100.00%	0.37%

*Panel C: Frequency of accounting fraud firm-years*

accounting fraud firm-years	Number of firms	Percentage of firms
1	27	32.53%
2	15	18.07%
3	9	10.84%
4	11	13.25%
5	17	20.48%
6	2	2.41%
8	2	2.41%
Total	83	100.00%

*Note:* Brackets in the "Accounting fraud firm-years" in panel B represent number of firms.

and their amounts are also large. Fictional and inflated sales also seem to be popular methods.



TABLE 4 TYPES OF ACCOUNTING FRAUD

Type of accounting fraud	Method-1	Method-2	Method-3	Total	Percentage (Total / 83)	Average percentage (max(modified net income or modified sales) / sales)	Average percentage (max(modified net income or modified sales) / total assets)
Misstated sales	39	6	45	45	54.22%	123.15%	16.53%
Misstated Cost of Goods Sold	4	2	6	6	7.23%	6.68%	9.86%
Misstated expense or loss (impairment loss, goodwill, etc.)	9	4	2	15	18.07%	51.90%	14.26%
Misstated assets (Accounts receivable, work-in-process, inventory, fixed assets, etc.)	7	39	7	53	63.86%	47.15%	42.92%
Misstated liabilities (Bad debt allowance, short-term Debt, etc.)	10	6	2	18	21.69%	9.44%	12.94%
Misstated stock	4	4	1	9	10.84%	36.14%	35.30%
Embezzlement, misappropriation	2	1	3	3	3.61%	15.32%	8.79%
Trading impersonation, circulation transactions, fictitious capital increase	5	2	2	9	10.84%	1108.16%	2.64%
Others (Consolidated Remove, etc.)	3	3	6	6	7.23%	82.05%	45.54%
Total	83	67	14	164	197.59%	120.53%	18.38%

*Note:* "Method-1" represents main method of accounting fraud. "Method-2" and "Method-3" represent the reasons for cases of multiple techniques. Therefore, the total percentage exceeds 100%.

## 5. Results

### 5.1 Univariate analysis

#### (1) Analysis of Misstate1

We describe the statistical features of the accounting fraud firms by comparing samples between the 241 fraud observations identified as Misstate1 and the 65,199 non-fraud observations. Table 5 summarizes the results of the univariate analyses, providing the mean and median for fraud and non-fraud firms, as well as the results of a *t*-test between them. The shadows in the table indicate the significant mean differences with the expected sign.

The table shows that many variables of accruals quality show significant differences between fraud and non-fraud firms. First, soft asset ratio (*% Soft assets*) is strongly significant, which implies that firms with higher soft asset ratios have more flexibility in earnings manipulation. Second, the values of almost all the discretionary accruals variables of the fraud firms are higher than are those of the non-fraud firms, as expected. This indicates that the discretionary accruals are useful for detecting accounting fraud. Furthermore, their derivatives, such as the standard deviation of discretionary accruals (*discretionary accruals SD*) and their absolute value (*discretionary accruals AB*) indicate more significant differences on the basis of *t*-statistics. This result suggests that a consideration of accrual reversal may improve accounting fraud prediction.

None of the variables of performance, nonfinancial, or off-balance-sheet is significant. The non-significance of the performance variables is consistent with the result in Dechow et al. (2011). On the other hand, most of variables of market show significant differences. We find that the book-to-market (*Book-to-market*) has strong significant differences, with a *t*-value of 25.3, the highest value among the results for all the variables. The firms with higher financial needs (*Actual issuance*), higher leverage company (*Leverage*), and lower PER (*Earnings-to-price*) are more likely to conduct accounting fraud.

We also find that all variables of accounting conservatism have significant differences. The variables calculated by using the market adjusted return (*C score AR*, and *C score ARSD*) show more significant results than do those calculated by using the raw return. The results suggest that the firms with higher conservatism are less likely to be accounting fraud firms.

The table reveals that the abnormal operating cash flow (*AB cash flow*) and the abnormal production cost (*AB product cost*) among the variables of real activities manipulation are statistically significant, with the expected sign, indicating that firms involved in real earnings management such as the manipulation of sales or production costs tend to engage in accounting fraud.

The ratio of business corporation ownership (*CORP*) of the Japanese-specific factors has significant differences, although ownership by financial institution (*FIN*) has no significant results. Cross-shareholdings among business companies may reduce the mutual monitoring function of shareholders and increase the opportunity for accounting fraud.

Finally, Table 6 presents the results of Wilcoxon rank sum tests as a non-parametric measure. The results are similar to those of the *t*-test reported in Table 5, except for a few variables of performance. Some variables show significant differences between fraud and non-fraud firms.<sup>13</sup>

<sup>13</sup> In addition to the main analyses, Dechow et al. (2011) conducted other comparative analyses: (1) between the fraud year and the other years for fraud companies and (2) between the fraud year and the previous years for fraud companies. Although we conduct the same analyses, we omit the results because of space constraints. The results are generally consistent with those of the main analyses. However, for (1), we observed a declining tendency in the explanatory power of the accruals variables.

TABLE 5 DESCRIPTIVE STATISTICS ON ACCOUNTING FRAUD FIRM-YEARS VERSUS NON-ACCOUNTING FRAUD FIRM-YEARS: T-TEST WITH MISSTATE1

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Accounting fraud – Non-Acc fraud		
	N	Mean	Median	N	Mean	Median		Diff. in mean	t-statistics	p-value
<b>Accruals quality related variables</b>										
<i>WC accruals</i>	239	0.180	0.167	64,191	0.146	0.140	+	0.034	-2.689	0.008
<i>RSS1 accruals</i>	239	0.016	0.017	64,191	0.014	0.012	+	0.003	-0.244	0.808
<i>Change in receivables</i>	239	0.010	0.000	64,191	0.006	0.001	+	0.004	-0.587	0.558
<i>Change in inventory</i>	239	0.006	0.000	64,191	0.003	0.000	+	0.002	-0.431	0.667
<i>% Soft assets</i>	241	74.261	75.535	65,199	65.173	67.025	+	9.088	-10.873	<.0001
<i>MJ discretionary accruals</i>	239	0.041	0.021	64,191	0.014	-0.007	+	0.027	-1.608	0.109
<i>MJ discretionary accruals SD</i>	193	0.137	0.084	43,215	0.113	0.045	+	0.024	-1.891	0.060
<i>MJ discretionary accruals AB</i>	239	0.173	0.117	64,191	0.141	0.081	+	0.032	-2.800	0.006
<i>PM discretionary accruals</i>	239	0.063	0.027	64,191	0.001	-0.008	+	0.062	-2.903	0.004
<i>PM discretionary accruals SD</i>	193	0.172	0.080	43,215	0.094	0.044	+	0.077	-3.939	<.0001
<i>PM discretionary accruals AB</i>	239	0.199	0.120	64,191	0.124	0.079	+	0.075	-4.007	<.0001
<i>CFO discretionary accruals</i>	238	0.013	0.006	57,091	0.001	0.001	+	0.012	-1.788	0.075
<i>CFO discretionary accruals SD</i>	153	0.052	0.046	36,654	0.034	0.025	+	0.018	-6.324	<.0001
<i>CFO discretionary accruals AB</i>	238	0.063	0.048	57,091	0.033	0.022	+	0.029	-8.384	<.0001
<b>Performance variables</b>										
<i>Change in cash sales</i>	238	8.069	0.724	63,071	5.151	2.135	-	2.918	-1.224	0.222
<i>Change in cash margin</i>	238	0.172	-13.650	63,055	-3.219	-1.406	-	3.391	-0.154	0.878
<i>Change in return on assets</i>	238	-0.023	-0.006	63,098	0.000	0.001	+	-0.024	2.458	0.015
<i>Change in free cash flows</i>	238	0.004	0.001	63,098	0.005	0.003	-	-0.001	0.135	0.893
<b>Nonfinancial variables</b>										
<i>Abnormal change in employees</i>	239	-5.204	-1.050	63,376	-1.188	-0.695	-	-4.016	1.615	0.108
<i>Abnormal change in order backlog</i>	107	22.684	-1.731	21,651	6.677	-1.141	-	16.007	-1.104	0.272
<b>Off-balance-sheet variables</b>										
<i>Existence of operating leases</i>	239	0.259	0.000	63,809	0.325	0.000	+	-0.066	2.309	0.022
<i>Change in operating lease activity</i>	56	27.460	-17.200	18,633	66.390	-12.936	+	-38.930	1.627	0.109

(Continued)

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Accounting fraud – Non-Acc fraud		
	N	Mean	Median	N	Mean	Median		Diff. in mean	t-statistics	p-value
<b>Market-related incentives</b>										
<i>Ex ante financing need</i>	228	0.268	0.000	51,451	0.374	0.000	+	-0.106	3.599	<.0001
<i>Actual issuance</i>	240	0.558	1.000	62,355	0.256	0.000	+	0.302	-9.395	<.0001
<i>CFF</i>	238	0.087	0.020	61,966	-0.001	-0.014	+	0.088	-7.232	<.0001
<i>Leverage</i>	241	0.122	0.103	65,199	0.105	0.060	+	0.017	-2.272	0.024
<i>Market adjusted stock-return</i>	221	-0.082	-0.068	49,677	-0.009	-0.009	+	-0.074	1.799	0.073
<i>Lagged market adjusted stock-return</i>	206	-0.095	-0.090	48,327	-0.017	-0.018	+	-0.078	1.904	0.058
<i>Book-to-market</i>	233	0.193	0.000	51,252	1.245	1.061	-	-1.052	25.332	<.0001
<i>Earnings-to-price</i>	233	-0.191	0.007	51,252	-0.007	0.046	-	-0.185	7.106	<.0001
<b>Accounting conservatism variables</b>										
<i>C score RR</i>	233	0.124	0.022	51,252	0.020	-0.031	+	0.104	-5.633	<.0001
<i>C score RRSB</i>	149	0.345	0.276	33,678	0.187	0.121	+	0.158	-3.591	<.0001
<i>C score AR</i>	233	0.270	0.236	51,252	0.099	0.072	+	0.172	-10.700	<.0001
<i>C score ARSD</i>	149	0.354	0.183	33,678	0.135	0.089	+	0.220	-5.141	<.0001
<b>Real activities manipulation variables</b>										
<i>AB cash flow</i>	239	-0.060	-0.039	61,965	0.001	0.001	-	-0.061	6.189	<.0001
<i>Discretionary expense</i>	237	0.013	-0.007	63,928	0.000	-0.014	-	0.013	-1.350	0.178
<i>AB product cost</i>	238	0.051	0.059	63,098	0.000	0.019	+	0.052	-3.022	0.003
<b>Japanese specific factors-related variables</b>										
<i>FIN</i>	225	1.091	0.064	56,545	0.427	0.035	-	0.664	-4.811	<.0001
<i>CORP</i>	236	7.931	0.197	59,398	2.417	0.058	+	5.515	-7.002	<.0001

Note: All variables are defined in Table 2. Shaded variables are those with p-values less than 10% and with the correct sign.

TABLE 6 DESCRIPTIVE STATISTICS ON ACCOUNTING FRAUD FIRM-YEARS VERSUS NON-ACCOUNTING FRAUD FIRM-YEARS: WILCOXON RANK SUM TEST WITH MISSTATE1

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Wilcoxon rank sum test	
	N	Mean	Median	N	Mean	Median		z-statistics	p-value
<b>Accruals quality related variables</b>									
<i>WC accruals</i>	239	0.180	0.167	64,191	0.146	0.140	+	2.523	0.006
<i>RSST accruals</i>	239	0.016	0.017	64,191	0.014	0.012	+	0.593	0.276
<i>Change in receivables</i>	239	0.010	0.000	64,191	0.006	0.001	+	-1.242	0.107
<i>Change in inventory</i>	239	0.006	0.000	64,191	0.003	0.000	+	-1.950	0.026
<i>% Soft assets</i>	241	74.261	75.535	65,199	65.173	67.025	+	9.427	<.0001
<i>MJ discretionary accruals</i>	239	0.041	0.021	64,191	0.014	-0.007	+	2.292	0.011
<i>MJ discretionary accruals SD</i>	193	0.137	0.084	43,215	0.113	0.045	+	7.768	<.0001
<i>MJ discretionary accruals AB</i>	239	0.173	0.117	64,191	0.141	0.081	+	5.410	<.0001
<i>PM discretionary accruals</i>	239	0.063	0.027	64,191	0.001	-0.008	+	3.541	<.0001
<i>PM discretionary accruals SD</i>	193	0.172	0.080	43,215	0.094	0.044	+	8.508	<.0001
<i>PM discretionary accruals AB</i>	239	0.199	0.120	64,191	0.124	0.079	+	5.875	<.0001
<i>CFO discretionary accruals</i>	238	0.013	0.006	57,091	0.001	0.001	+	1.894	0.029
<i>CFO discretionary accruals SD</i>	153	0.052	0.046	36,654	0.034	0.025	+	6.991	<.0001
<i>CFO discretionary accruals AB</i>	238	0.063	0.048	57,091	0.033	0.022	+	9.368	<.0001
<b>Performance variables</b>									
<i>Change in cash sales</i>	238	8.069	0.724	63,071	5.151	2.135	-	-1.737	0.041
<i>Change in cash margin</i>	238	0.172	-13.650	63,055	-3.219	-1.406	-	-2.250	0.012
<i>Change in return on assets</i>	238	-0.023	-0.006	63,098	0.000	0.001	+	-4.168	<.0001
<i>Change in free cash flows</i>	238	0.004	0.001	63,098	0.005	0.003	-	-1.527	0.063
<b>Nonfinancial variables</b>									
<i>Abnormal change in employees</i>	239	-5.204	-1.050	63,376	-1.188	-0.695	-	0.174	0.431
<i>Abnormal change in order backlog</i>	107	22.684	-1.731	21,651	6.677	-1.141	-	-0.423	0.336
<b>Off-balance-sheet variables</b>									
<i>Existence of operating leases</i>	239	0.259	0.000	63,809	0.325	0.000	+	-2.166	0.015
<i>Change in operating lease activity</i>	56	27.460	-17.200	18,633	66.390	-12.936	+	-0.439	0.330

(Continued)

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Wilcoxon rank sum test	
	N	Mean	Median	N	Mean	Median		z-statistics	p-value
<b>Market-related incentives</b>									
<i>Ex ante financing need</i>	228	0.268	0.000	51,451	0.374	0.000	+	-3.303	<.0001
<i>Actual issuance</i>	240	0.558	1.000	62,355	0.256	0.000	+	10.691	<.0001
<i>CFF</i>	238	0.087	0.020	61,966	-0.001	-0.014	+	7.937	<.0001
<i>Leverage</i>	241	0.122	0.103	65,199	0.105	0.060	+	3.841	<.0001
<i>Market adjusted stock-return</i>	221	-0.082	-0.068	49,677	-0.009	-0.009	+	-2.935	0.002
<i>Lagged market adjusted stock-return</i>	206	-0.095	-0.090	48,327	-0.017	-0.018	+	-2.345	0.010
<i>Book-to-market</i>	233	0.193	0.000	51,252	1.245	1.061	-	-22.454	<.0001
<i>Earnings-to-price</i>	233	-0.191	0.007	51,252	-0.007	0.046	-	-10.156	<.0001
<b>Accounting conservatism variables</b>									
<i>C score RR</i>	233	0.124	0.022	51,252	0.020	-0.031	+	5.851	<.0001
<i>C score RRSB</i>	149	0.345	0.276	33,678	0.187	0.121	+	9.910	<.0001
<i>C score AR</i>	233	0.270	0.236	51,252	0.099	0.072	+	12.231	<.0001
<i>C score ARSD</i>	149	0.354	0.183	33,678	0.135	0.089	+	11.947	<.0001
<b>Real activities manipulation variables</b>									
<i>AB cash flow</i>	239	-0.060	-0.039	61,965	0.001	0.001	-	-9.300	<.0001
<i>Discretionary expense</i>	237	0.013	-0.007	63,928	0.000	-0.014	-	1.486	0.069
<i>AB product cost</i>	238	0.051	0.059	63,098	0.000	0.019	+	3.465	<.0001
<b>Japanese specific factors-related variables</b>									
<i>FIN</i>	225	1.091	0.064	56,545	0.427	0.035	-	5.178	<.0001
<i>CORP</i>	236	7.931	0.197	59,398	2.417	0.058	+	5.770	<.0001

Note: All variables are defined in Table 2. Shaded variables are those with p-values less than 10% and with the correct sign.

## (2) Analysis of Misstate 2

Table 7 shows the results for Misstate2. The results are generally consistent with the results for Misstate 1 seen in Table 5, except that the variable of operating lease (*Existence of operating leases*) is significant with the expected sign, whereas leverage (*Leverage*) is not. The results of a Wilcoxon rank sum test for Misstate2 are summarized in Table 8. The results are also similar to those of Table 6, except that the variables of off-balance-sheet are more significant.

Table 9 shows the AR (accuracy ratio) of all variables for Misstate1 and Misstate2. The AR is a popular indicator, especially for corporate failure prediction models. In this performance indicator, AR reaches 1 when the model's prediction performance becomes perfect.<sup>14</sup> The results show that 1) the variables with higher AR values are common to both Misstate1 and Misstate2, and 2) the AR values for Misstate1 are higher than are those for Misstate2. This suggests that the discrimination performance of accounting fraud deteriorated for Misstate2; we infer that this occurred because Misstate2 contains more indefinite samples with less significant incidents (i.e., a small amount of fraud).

## 5.2 Parameter estimation of fraud detection model

### (1) Method of estimation

In this section, we develop a prediction model to identify accounting fraud on the basis of the results of the univariate analysis in the previous section. We present the model based on Misstate1 as the main result since, in univariate analysis, the explanatory powers (the AR value) of the variables of Misstate 1 are generally higher than are those of Misstate 2. The results based on Misstate 2 to Misstate 8 are presented as additional analyses.

We performed variable selections and parameter estimations for each model following the procedures of Dechow et al. (2011). Specifically, (1) candidate variables for use in the prediction model are selected among those in Table 5 or Table 6 with a significance level of 10% or higher with the correct expected sign; (2) a dependent variable is equal to 1 for firm-years involving Misstate1 and 0 otherwise; and (3) a logistic regression is employed to determine a battery of variables and their parameters. Our logistic regression is based on a backward method with a restriction of the correct parameter sign by eliminating a variable step-by-step from the highest *p*-value.<sup>15</sup>

Dechow et al. (2011) build three models with different sets of parameters. Model 1 includes only financial-statement variables as predictors. Model 2 adds nonfinancial and off-balance-sheet variables to Model 1, and Model 3 adds market-based measures to Model 2 (aside from the variables used in Model 1).

Our research follows the procedures of Dechow et al. (2001) but extend it in the following ways. First, we employ variables relating to discretionary accruals as significant variables, which Dechow et al. (2011) do not use. Although the level and absolute value of discretionary accruals have significant differences in the univariate analysis, both variables should be used in the estimation procedure. However, we use only the absolute value of discretionary accruals in the

<sup>14</sup> Please see Appendix D for details on the construct of AR and its method of calculation.

<sup>15</sup> Other than the abovementioned method of backward elimination with the restriction of the correct parameter sign, we employed other methods of variable selection, such as standard backward elimination, forward elimination, and stepwise elimination. The backward elimination with sign restriction provided the best results in terms of correct signs and higher AR.

TABLE 7 DESCRIPTIVE STATISTICS ON ACCOUNTING FRAUD FIRM-YEARS VERSUS NON-ACCOUNTING FRAUD FIRM-YEARS: T-TEST WITH MISSTATE2

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Accounting fraud – Non-Acc fraud		
	N	Mean	Median	N	Mean	Median		Diff. in mean	t-statistics	p-value
<b>Accruals quality related variables</b>										
<i>WC accruals</i>	859	0.175	0.163	63,571	0.146	0.139	+	0.029	-5.089	<.0001
<i>RSS T accruals</i>	859	0.004	0.009	63,571	0.014	0.012	+	-0.009	2.202	0.028
<i>Change in receivables</i>	859	0.006	0.001	63,571	0.006	0.001	+	0.000	0.154	0.878
<i>Change in inventory</i>	859	0.001	0.000	63,571	0.003	0.000	+	-0.003	1.347	0.178
<i>% Soft assets</i>	861	70.289	70.938	64,579	65.139	67.009	+	5.150	-11.122	<.0001
<i>MJ discretionary accruals</i>	859	0.027	0.010	63,571	0.014	-0.007	+	0.013	-0.947	0.344
<i>MJ discretionary accruals SD</i>	783	0.132	0.059	42,625	0.113	0.045	+	0.019	-1.760	0.079
<i>MJ discretionary accruals AB</i>	859	0.172	0.088	63,571	0.140	0.081	+	0.032	-2.275	0.023
<i>PM discretionary accruals</i>	859	0.041	0.009	63,571	0.000	-0.008	+	0.041	-4.451	<.0001
<i>PM discretionary accruals SD</i>	783	0.116	0.055	42,625	0.094	0.043	+	0.022	-3.089	0.002
<i>PM discretionary accruals AB</i>	859	0.154	0.089	63,571	0.124	0.079	+	0.030	-3.831	<.0001
<i>CFO discretionary accruals</i>	854	0.006	0.003	56,475	0.001	0.001	+	0.006	-2.245	0.025
<i>CFO discretionary accruals SD</i>	709	0.041	0.035	36,098	0.034	0.025	+	0.007	-6.561	<.0001
<i>CFO discretionary accruals AB</i>	854	0.042	0.029	56,475	0.033	0.022	+	0.009	-6.223	<.0001
<b>Performance variables</b>										
<i>Change in cash sales</i>	856	5.377	1.800	62,453	5.159	2.134	-	0.218	-0.232	0.816
<i>Change in cash margin</i>	856	-1.226	-3.270	62,437	-3.233	-1.400	-	2.007	-0.234	0.815
<i>Change in return on assets</i>	856	-0.010	-0.001	62,480	0.000	0.001	+	-0.010	2.665	0.008
<i>Change in free cash flows</i>	856	0.004	0.002	62,480	0.005	0.003	-	-0.002	0.459	0.646
<b>Nonfinancial variables</b>										
<i>Abnormal change in employees</i>	858	-1.944	-0.134	62,757	-1.193	-0.699	-	-0.751	0.814	0.416
<i>Abnormal change in order backlog</i>	373	7.444	0.372	21,385	6.744	-1.160	-	0.700	-0.147	0.883
<b>Off-balance-sheet variables</b>										
<i>Existence of operating leases</i>	858	0.393	0.000	63,190	0.324	0.000	+	0.069	-4.098	<.0001
<i>Change in operating lease activity</i>	320	96.763	-9.337	18,369	65.742	-12.988	+	31.021	-0.844	0.399

(Continued)



Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Accounting fraud – Non-Acc fraud		
	N	Mean	Median	N	Mean	Median		Diff. in mean	t-statistics	p-value
<b>Market-related incentives</b>										
<i>Ex ante financing need</i>	836	0.281	0.000	50,843	0.375	0.000	+	-0.094	5.955	<.0001
<i>Actual issuance</i>	858	0.368	0.000	61,737	0.256	0.000	+	0.113	-6.795	<.0001
<i>CFF</i>	856	0.029	-0.006	61,348	-0.001	-0.014	+	0.030	-6.529	<.0001
<i>Leverage</i>	861	0.107	0.082	64,579	0.105	0.060	+	0.003	-0.693	0.488
<i>Market adjusted stock–return</i>	822	-0.039	-0.044	49,076	-0.008	-0.009	+	-0.031	1.867	0.062
<i>Lagged market adjusted stock–return</i>	794	-0.053	-0.054	47,739	-0.017	-0.017	+	-0.037	2.190	0.029
<i>Book-to-market</i>	841	0.828	0.655	50,644	1.247	1.061	-	-0.419	13.092	<.0001
<i>Earnings-to-price</i>	841	-0.067	0.033	50,644	-0.006	0.046	-	-0.060	5.637	<.0001
<b>Accounting conservatism variables</b>										
<i>C score RR</i>	841	0.218	-0.031	50,644	0.017	-0.031	+	0.201	-1.982	0.048
<i>C score RRSB</i>	685	1.598	0.165	33,142	0.158	0.121	+	1.439	-3.116	0.002
<i>C score AR</i>	841	0.403	0.125	50,644	0.095	0.072	+	0.309	-2.242	0.025
<i>C score ARSD</i>	685	1.538	0.116	33,142	0.107	0.089	+	1.431	-3.373	0.001
<b>Real activities manipulation variables</b>										
<i>AB cash flow</i>	857	-0.031	-0.019	61,347	0.001	0.001	-	-0.031	8.414	<.0001
<i>Discretionary expense</i>	857	0.009	-0.005	63,308	0.000	-0.014	-	0.009	-2.072	0.039
<i>AB product cost</i>	856	0.019	0.022	62,480	0.000	0.019	+	0.019	-2.668	0.008
<b>Japanese specific factors-related variables</b>										
<i>FIN</i>	830	0.652	0.038	55,940	0.426	0.035	-	0.226	-3.592	<.0001
<i>CORP</i>	845	4.479	0.078	58,789	2.409	0.058	+	2.069	-6.170	<.0001

Note: All variables are defined in Table 2. Shaded variables are those with p-values less than 10% and with the correct sign.

TABLE 8 DESCRIPTIVE STATISTICS ON ACCOUNTING FRAUD FIRM-YEARS VERSUS NON-ACCOUNTING FRAUD FIRM-YEARS: WILCOXON RANK SUM TEST WITH MISSTATE2

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Wilcoxon rank sum test	
	N	Mean	Median	N	Mean	Median		z-statistics	p-value
<b>Accruals quality related variables</b>									
<i>WC accruals</i>	859	0.175	0.163	63,571	0.146	0.139	+	5.118	<.0001
<i>RSST accruals</i>	859	0.004	0.009	63,571	0.014	0.012	+	-2.869	0.002
<i>Change in receivables</i>	859	0.006	0.001	63,571	0.006	0.001	+	-1.747	0.040
<i>Change in inventory</i>	859	0.001	0.000	63,571	0.003	0.000	+	-2.368	0.009
<i>% Soft assets</i>	861	70.289	70.938	64,579	65.139	67.009	+	8.816	<.0001
<i>MJ discretionary accruals</i>	859	0.027	0.010	63,571	0.014	-0.007	+	3.529	<.0001
<i>MJ discretionary accruals SD</i>	783	0.132	0.059	42,625	0.113	0.045	+	8.477	<.0001
<i>MJ discretionary accruals AB</i>	859	0.172	0.088	63,571	0.140	0.081	+	3.636	<.0001
<i>PM discretionary accruals</i>	859	0.041	0.009	63,571	0.000	-0.008	+	4.712	<.0001
<i>PM discretionary accruals SD</i>	783	0.116	0.055	42,625	0.094	0.043	+	8.717	<.0001
<i>PM discretionary accruals AB</i>	859	0.154	0.089	63,571	0.124	0.079	+	3.522	<.0001
<i>CFO discretionary accruals</i>	854	0.006	0.003	56,475	0.001	0.001	+	2.410	0.008
<i>CFO discretionary accruals SD</i>	709	0.041	0.035	36,098	0.034	0.025	+	8.473	<.0001
<i>CFO discretionary accruals AB</i>	854	0.042	0.029	56,475	0.033	0.022	+	6.446	<.0001
<b>Performance variables</b>									
<i>Change in cash sales</i>	856	5.377	1.800	62,453	5.159	2.134	-	-2.024	0.021
<i>Change in cash margin</i>	856	-1.226	-3.270	62,437	-3.233	-1.400	-	-1.206	0.114
<i>Change in return on assets</i>	856	-0.010	-0.001	62,480	0.000	0.001	+	-3.405	<.0001
<i>Change in free cash flows</i>	856	0.004	0.002	62,480	0.005	0.003	-	-2.126	0.017
<b>Nonfinancial variables</b>									
<i>Abnormal change in employees</i>	858	-1.944	-0.134	62,757	-1.193	-0.699	-	1.758	0.039
<i>Abnormal change in order backlog</i>	373	7.444	0.372	21,385	6.744	-1.160	-	0.346	0.365
<b>Off-balance-sheet variables</b>									
<i>Existence of operating leases</i>	858	0.393	0.000	63,190	0.324	0.000	+	4.273	<.0001
<i>Change in operating lease activity</i>	320	96.763	-9.337	18,369	65.742	-12.988	+	1.635	0.051

(Continued)

Variable	Accounting fraud firm-years			Non-accounting fraud firm-years			Predicted sign	Wilcoxon rank sum test	
	N	Mean	Median	N	Mean	Median		z-statistics	p-value
<b>Market-related incentives</b>									
<i>Ex ante financing need</i>	836	0.281	0.000	50,843	0.375	0.000	+	-5.546	<.0001
<i>Actual issuance</i>	858	0.368	0.000	61,737	0.256	0.000	+	7.492	<.0001
<i>CFF</i>	856	0.029	-0.006	61,348	-0.001	-0.014	+	7.702	<.0001
<i>Leverage</i>	861	0.107	0.082	64,579	0.105	0.060	+	3.264	0.001
<i>Market adjusted stock-return</i>	822	-0.039	-0.044	49,076	-0.008	-0.009	+	-2.530	0.006
<i>Lagged market adjusted stock-return</i>	794	-0.053	-0.054	47,739	-0.017	-0.017	+	-2.614	0.005
<i>Book-to-market</i>	841	0.828	0.655	50,644	1.247	1.061	-	-16.267	<.0001
<i>Earnings-to-price</i>	841	-0.067	0.033	50,644	-0.006	0.046	-	-5.283	<.0001
<b>Accounting conservatism variables</b>									
<i>C score RR</i>	841	0.218	-0.031	50,644	0.017	-0.031	+	1.064	0.144
<i>C score RRSB</i>	685	1.598	0.165	33,142	0.158	0.121	+	11.056	<.0001
<i>C score AR</i>	841	0.403	0.125	50,644	0.095	0.072	+	11.297	<.0001
<i>C score ARSD</i>	685	1.538	0.116	33,142	0.107	0.089	+	13.731	<.0001
<b>Real activities manipulation variables</b>									
<i>AB cash flow</i>	857	-0.031	-0.019	61,347	0.001	0.001	-	-10.843	<.0001
<i>Discretionary expense</i>	857	0.009	-0.005	63,308	0.000	-0.014	-	3.905	<.0001
<i>AB product cost</i>	856	0.019	0.022	62,480	0.000	0.019	+	1.129	0.129
<b>Japanese specific factors-related variables</b>									
<i>FIN</i>	830	0.652	0.038	55,940	0.426	0.035	-	3.127	0.001
<i>CORP</i>	845	4.479	0.078	58,789	2.409	0.058	+	5.508	<.0001

Note: All variables are defined in Table 2. Shaded variables are those with p-values less than 10% and with the correct sign.

TABLE 9 ACCURACY RATIO (AR) FOR EACH INDEPENDENT VARIABLE

Variable	AR whit Misstate1		AR whit Misstate2	
	Value	rank	Value	rank
<b>Accruals quality related variables</b>				
<i>WC accruals</i>	0.094	21	0.101	15
<i>RSST accruals</i>	0.022	28	-0.057	35
<i>Change in receivables</i>	-0.046	31	-0.035	30
<i>Change in inventory</i>	-0.073	34	-0.047	32
<i>% Soft assets</i>	0.351	8	0.175	9
<i>MJ discretionary accruals</i>	0.086	22	0.070	18
<i>MJ discretionary accruals SD</i>	0.324	11	0.177	8
<i>MJ discretionary accruals AB</i>	0.202	17	0.072	17
<i>PM discretionary accruals</i>	0.132	19	0.093	16
<i>PM discretionary accruals SD</i>	0.354	6	0.182	7
<i>PM discretionary accruals AB</i>	0.220	15	0.070	19
<i>CFO discretionary accruals</i>	0.071	24	0.048	23
<i>CFO discretionary accruals SD</i>	0.327	10	0.186	6
<i>CFO discretionary accruals AB</i>	0.351	7	0.128	11
<b>Performance variables</b>				
<i>Change in cash sales</i>	0.065	25	0.040	25
<i>Change in cash margin</i>	0.084	23	0.024	26
<i>Change in return on assets</i>	-0.156	38	-0.068	37
<i>Change in free cash flows</i>	0.057	26	0.042	24
<b>Nonfinancial variables</b>				
<i>Abnormal change in employees</i>	-0.007	29	-0.035	31
<i>Abnormal change in order backlog</i>	0.024	27	-0.010	29
<b>Off-balance-sheet variables</b>				
<i>Existence of operating leases</i>	-0.066	33	0.069	20
<i>Change in operating lease activity</i>	-0.034	30	0.053	22
<b>Market-related incentives</b>				
<i>Ex ante financing need</i>	-0.106	36	-0.094	39
<i>Actual issuance</i>	0.302	12	0.113	12
<i>CFE</i>	0.298	13	0.153	10
<i>Leverage</i>	0.142	18	0.064	21
<i>Market adjusted stock-return</i>	-0.114	37	-0.051	33
<i>Lagged market adjusted stock-return</i>	-0.095	35	-0.054	34
<i>Book-to-market</i>	0.851	1	0.327	1
<i>Earnings-to-price</i>	0.385	5	0.106	14
<b>Accounting conservatism variables</b>				
<i>C score RR</i>	0.222	14	0.021	28
<i>C score RRSD</i>	0.470	3	0.246	3
<i>C score AR</i>	0.464	4	0.227	4
<i>C score ARSD</i>	0.566	2	0.306	2
<b>Real activities manipulation variables</b>				
<i>AB cash flow</i>	0.348	9	0.215	5
<i>Discretionary expense</i>	-0.056	32	-0.078	38
<i>AB product cost</i>	0.130	20	0.022	27
<b>Japanese specific factors-related variables</b>				
<i>FIN</i>	-0.200	39	-0.063	36
<i>CORP</i>	0.217	16	0.110	13

Note: All variables are defined in Table 2.

TABLE 10 LOGISTIC REGRESSIONS WITH MISSTATE1

Variable	Model1			Model2		
	Coefficient	Wald <i>Chi-square</i>	<i>p-value</i>	Coefficient	Wald <i>Chi-square</i>	<i>p-value</i>
Intercept	-8.913	515.626	<.0001	-7.209	325.390	<.0001
% <i>Soft assets</i>	0.038	55.676	<.0001	0.036	49.701	<.0001
<i>CFO discretionary accruals AB</i>	7.654	33.854	<.0001	2.639	3.396	0.065
<i>Actual issuance</i>	0.945	44.472	<.0001	0.666	20.194	<.0001
<i>C score AR</i>	0.049	7.610	0.006			
<i>AB cash flow</i>	-1.337	8.429	0.004	-1.690	13.621	<.0001
<i>CORP</i>	0.032	51.096	<.0001	0.024	26.804	<.0001
<i>Book-to-market</i>				-1.455	139.081	<.0001
AR			0.577			0.826
Accounting fraud firm-years			229			229
Non-accounting fraud firm-years			47,709			47,709

Note: All variables are defined in Table 2.

analysis because the two variables have a strong correlation and might cause multicollinearity. Furthermore, as mentioned, the use of absolute values of discretionary accruals has an advantage in that it can capture the effect of accruals reversal. Second, we also employ the variables of conservatism, real activities manipulation, and Japanese-specific factors. All these variables are publicly available and easily obtained by financial statement users. These variables are added to all three models in the settings used by Dechow et al. (2011).

In our univariate analysis, none of the variables of nonfinancial and off-balance-sheet are significant. Therefore, we are not able to build Model 2. Thus, we build the following two models:<sup>16</sup>

Model 1: variables available from *financial statement* in the categories of accruals, performance, market-related incentive, conservatism, real activities manipulation, Japanese-specific factors

Model 2: variables in the categories of accruals, performance, conservatism, real activities manipulation, Japanese-specific factors, and market-related incentive

## (2) Estimation results

In estimating the parameters of Models 1 and 2, we incorporate the variables of the three types of discretionary accruals separately into the model because these variables have high correlations. For the same reason, the two variables in the category of conservatism are also incorporated separately into the models. Consequently, we estimate a total of six models as a combination of these variables.<sup>17</sup> From among the results of the six models, we present only the results of the model with the highest AR value.

The regression results are shown in Table 10. Six variables are selected in Model 1 by the backward elimination method with sign restriction: % *Soft assets*, *CFO discretionary accruals AB*,

<sup>16</sup> Therefore, the models in our research are equivalent to Models 1 and 3 in Dechow et al. (2011), with additional variables in the categories of conservatism, real activities manipulation, and Japanese-specific factors.

<sup>17</sup> The variables of *Ex ante financing need*, *CFF* and *Abnormal change in order backlog* are excluded, as in Dechow et al. (2011), because they are available for only a limited set of firms.

*Actual issuance*, *C score AR*, *AB cash flow*, and *CORP*. The four variables among those are newly added to our model: absolute value of discretionary accruals, conservatism score, abnormal operating cash flow to capture abnormal discretionary activities, and ownership by business corporation. The results suggest the importance of considering earnings quality and Japanese-specific factors when capturing financial fraud.

As for Model 2, the table indicates that six variables are selected: *%Soft assets*, *CFO discretionary accruals AB*, *Actual issuance*, *AB cash flow*, *CORP*, and *Book-to-market*. These variables are almost the same as the results of Model 1, except *Book-to-market* is a substitute for conservatism (*C score AR*). The results of the  $\chi^2$  value indicate that the value of *Book-to-market* is highly significant. The result suggests that managers of potential fraud firms are likely to have a strong incentive to maintain higher stock prices, consistent with our prediction.

### (3) Further validation

This section reports the estimation results for Misstate2 to Misstate8. As described in section 3.1, we define the variables from Misstate3 to Misstate8 so that we can capture the degree of accounting fraud gradually (see Table 1). Table 11 summarizes the estimation results of Model 1. The major findings are as follows. 1) The variables selected for Misstate2 to Misstate8 are almost the same as those for Misstate1; 2) the discriminant power measured by AR is less capable than for Misstate1; and 3) the discriminant power increases as it approaches Misstate8. These results suggest that the discriminant power increases as accounting fraud becomes more severe. Thus, in the following analysis on the validity of our models, we will present the results of the models on the basis of Misstate1.

## 5.3 Performance of the prediction model

### (1) Calculation of the *F-Score*

We discuss the validity of our prediction models, Model 1 and Model 2. The *F-Score* is the predicted probability of financial statement fraud in accordance with Dechow et al. (2011) as follows:

$$\text{Probability} = \frac{e^{(\text{Predicted Value})}}{1 + e^{(\text{Predicted Value})}}$$

The predicted value is the intensity parameter for the likelihood of financial statement fraud. We then derive the probability by the unconditional expectation (unconditional probability) of Misstate1 to calculate our *F-Score*. The unconditional probability is a value of the number of accounting fraud samples divided by the total number of samples. An *F-Score* of 1.00 indicates that the firm has the same probability of accounting fraud as the unconditional expectation. An *F-Score* greater than 1 indicates firms with higher probabilities of fraud. Below is an example of how this is done for Model 1 for the company named OLYMPUS (stock code #7733) in the financial statements of March 2008:

TABLE 11 LOGISTIC REGRESSIONS WITH MISSTATE8 FROM MISSTATE2

Variable	Model1			
	Coefficient estimate	Wald Chi-square	p-value	AR
<b>Misstate2</b>				
Intercept	-5.700	874.879	<.0001	0.316
% <i>Soft assets</i>	0.019	54.560	<.0001	
<i>CFO discretionary accruals AB</i>	3.103	12.564	<.0001	
<i>Leverage</i>	0.511	2.840	0.092	
<i>Actual issuance</i>	0.279	12.257	0.001	
<i>C score AR</i>	0.089	34.997	<.0001	
<i>AB cash flow</i>	-1.708	28.079	<.0001	
<i>CORP</i>	0.021	45.137	<.0001	
Accounting fraud firm-years	837			
Non-accounting fraud firm-years	47,101			
<b>Misstate3</b>				
Intercept	-8.127	608.735	<.0001	0.496
% <i>Soft assets</i>	0.037	69.961	<.0001	
<i>CFO discretionary accruals AB</i>	5.505	20.228	<.0001	
<i>Actual issuance</i>	0.610	26.478	<.0001	
<i>C score AR</i>	0.072	39.128	<.0001	
<i>AB cash flow</i>	-1.803	17.869	<.0001	
<i>CORP</i>	0.028	47.146	<.0001	
Accounting fraud firm-years	328			
Non-accounting fraud firm-years	47,610			
<b>Misstate4</b>				
Intercept	-8.205	580.678	<.0001	0.513
% <i>Soft assets</i>	0.035	61.321	<.0001	
<i>CFO discretionary accruals AB</i>	6.143	24.483	<.0001	
<i>Actual issuance</i>	0.709	33.047	<.0001	
<i>C score AR</i>	0.072	39.275	<.0001	
<i>AB cash flow</i>	-1.723	15.905	<.0001	
<i>CORP</i>	0.029	45.020	<.0001	
Accounting fraud firm-years	301			
Non-accounting fraud firm-years	47,637			
<b>Misstate5</b>				
Intercept	-8.175	564.069	<.0001	0.511
% <i>Soft assets</i>	0.034	56.252	<.0001	
<i>CFO discretionary accruals AB</i>	5.799	20.844	<.0001	
<i>Actual issuance</i>	0.768	37.601	<.0001	
<i>C score AR</i>	0.072	38.985	<.0001	

(Continued)

(-Continued)

Variable	Model1			
	Coefficient estimate	Wald Chi-square	p-value	AR
<i>AB cash flow</i>	-1.790	16.668	<.0001	
<i>CORP</i>	0.029	45.091	<.0001	
Accounting fraud firm-years	291			
Non-accounting fraud firm-years	47,647			
<b>Misstate6</b>				
Intercept	-8.174	560.169	<.0001	0.515
<i>% Soft assets</i>	0.034	54.373	<.0001	
<i>CFO discretionary accruals AB</i>	5.910	21.598	<.0001	
<i>Actual issuance</i>	0.793	39.613	<.0001	
<i>C score AR</i>	0.072	38.872	<.0001	
<i>AB cash flow</i>	-1.767	16.185	<.0001	
<i>CORP</i>	0.029	45.816	<.0001	
Accounting fraud firm-years	287			
Non-accounting fraud firm-years	47,651			
<b>Misstate7</b>				
Intercept	-8.251	548.058	<.0001	0.527
<i>% Soft assets</i>	0.033	51.188	<.0001	
<i>CFO discretionary accruals AB</i>	6.291	24.174	<.0001	
<i>Actual issuance</i>	0.855	43.641	<.0001	
<i>C score AR</i>	0.055	15.457	<.0001	
<i>AB cash flow</i>	-1.714	15.072	<.0001	
<i>CORP</i>	0.030	48.762	<.0001	
Accounting fraud firm-years	272			
Non-accounting fraud firm-years	47,666			
<b>Misstate8</b>				
Intercept	-8.228	545.642	<.0001	0.525
<i>% Soft assets</i>	0.033	50.018	<.0001	
<i>CFO discretionary accruals AB</i>	6.223	23.509	<.0001	
<i>Actual issuance</i>	0.865	44.476	<.0001	
<i>C score AR</i>	0.055	15.417	<.0001	
<i>AB cash flow</i>	-1.715	15.004	<.0001	
<i>CORP</i>	0.030	49.159	<.0001	
Accounting fraud firm-years	271			
Non-accounting fraud firm-years	47,667			

Note: All variables are defined in Table 2.



## [Model1]

Predicted Value =

$$-8.913 + 0.038 \times (\% \text{ soft assets}) + 7.654 \times (\text{CFO discretionary accruals AB}) \\ + 0.945 \times (\text{Actual issuance}) + 0.049 \times (\text{C score AR}) \\ -1.337 \times (\text{AB cash flow}) + 0.032 \times (\text{CORP})$$

Predicted Value =

$$-8.913 + 0.038 \times (79.341) + 7.654 \times (0.032) + 0.945 \times (1) + 0.049 \times (-0.318) \\ -1.337 \times (-0.014) + 0.032 \times (0.010)$$

Predicted Value = -4.705

Probability =  $e^{-4.705} / (1 + e^{-4.705})$

$e = 2.71828183$

Probability = 0.00897

Unconditional probability =  $229 / (47,709 + 229) = 0.00478$

$F\text{-Score} = 0.00897 / 0.00478 = 1.88$

OLYMPUS has an  $F\text{-Score}$  of 1.88. This suggests that OLYMPUS has almost twice the probability of engaging in financial fraud compared to a randomly selected firm from the population.

## (2) Accuracy of the prediction model

In order to test the validity of our  $F\text{-Score}$  in Models 1 and 2, we calculate the detection rate for accounting fraud firms as follows. We 1) sort all samples into five portfolios on the basis of the  $F\text{-Score}$  in ascending order, and 2) observe the frequency of the accounting fraud firms and non-accounting fraud firms in each quintile. If our models do a good job of identifying the accounting fraud firms, we expect the accounting fraud firms to be clustered in the fifth portfolio.

The results are reported in Table 12. For Model 1, 59.39% of the accounting fraud firms are in quintile 5, which is higher than the expected level of 20%. For Model 2, 90.83% of the accounting fraud firms are in quintile 5. Since the percentages in Dechow et al. (2011) are around 50% for all models, our model demonstrated superior performance. The cutoff to be included in quintile 5 (i.e., the minimum value) is 1.205 for Model 1 and 1.293 for Model 2, both above the 1.00 of the unconditional expectation.

## (3) Type I and Type II error rates of the prediction model

In this section, we discuss the Type I error and Type II error of our models. A Type I error means that the model incorrectly classifies a non-accounting fraud firm (non-FSF firm) as an accounting fraud firm (FSF firm), and a Type II error means the model incorrectly classifies a FSF firm as a non-FSF firm. We classify the observations based on the following procedure. We 1) define a firm with an  $F\text{-Score}$  of 1.00 or more as a predicted FSF firm (pred-FSF) and that with an  $F\text{-Score}$  less than 1.00 as a predicted non-FSF firm (pred-NonFSF); 2) we count the number of FSF and non-FSF firms among pred-FSF firms, and then 3) count the number of FSF and non-FSF firms among pred-NonFSF firms.

Table 13 shows the result of the classifications. The "correct classification" is the sum of the number of FSF firms among pred-FSF firms and the number of Non-FSF firms among pred-NonFSF firms, divided by the total firms. The result for Model 1 indicates that the correct

TABLE 12 VALIDATION OF PREDICTION MODEL ACCURACY

	Model1			Model2		
	<i>N</i>	<i>Min. F-score</i>		<i>N</i>	<i>Min. F-score</i>	
			% of total			% of total
<b>Quintile1</b>						
Accounting fraud firm-years	8	0.182	3.49%	4	0.008	1.75%
Non-accounting fraud firm-years	9,579	0.039	20.08%	9,583	0.001	20.09%
<b>Quintile2</b>						
Accounting fraud firm-years	9	0.342	3.93%	5	0.161	2.18%
Non-accounting fraud firm-years	9,579	0.331	20.08%	9,583	0.130	20.09%
<b>Quintile3</b>						
Accounting fraud firm-years	31	0.489	13.54%	1	0.361	0.44%
Non-accounting fraud firm-years	9,557	0.486	20.03%	9,587	0.315	20.09%
<b>Quintile4</b>						
Accounting fraud firm-years	45	0.720	19.65%	11	0.637	4.80%
Non-accounting fraud firm-years	9,543	0.708	20.00%	9,577	0.591	20.07%
<b>Quintile5</b>						
Accounting fraud firm-years	136	1.205	59.39%	208	1.293	90.83%
Non-accounting fraud firm-years	9,451	1.172	19.81%	9,379	1.179	19.66%

TABLE 13 VERIFICATION OF TYPE I ERRORS AND TYPE II ERRORS OF THE PREDICTION MODEL

Observed	Model1 predicted			Model2 predicted		
	Misstate	No-misstate		Misstate	No-misstate	
Misstate	152	77	229	212	17	229
No-misstate	11,876	35,833	47,709	11,326	36,383	47,709
	12,028	35,910	47,938	11,538	36,400	47,938
Misstate	66.38%	33.62%	0.48%	92.58%	7.42%	0.48%
No-misstate	24.89%	75.11%	99.52%	23.74%	76.26%	99.52%
Correct classification	75.07% (1)			76.34%		
Sensitivity	66.38% (2)			92.58%		
Type I errors	24.89% (3)			23.74%		
Type II errors	33.62% (4)			7.42%		

(1) Correct classification =  $(152+35,833) / 47,938$

(2) Sensitivity =  $152 / 229$

(3) Type I errors =  $11,876 / 47,709$

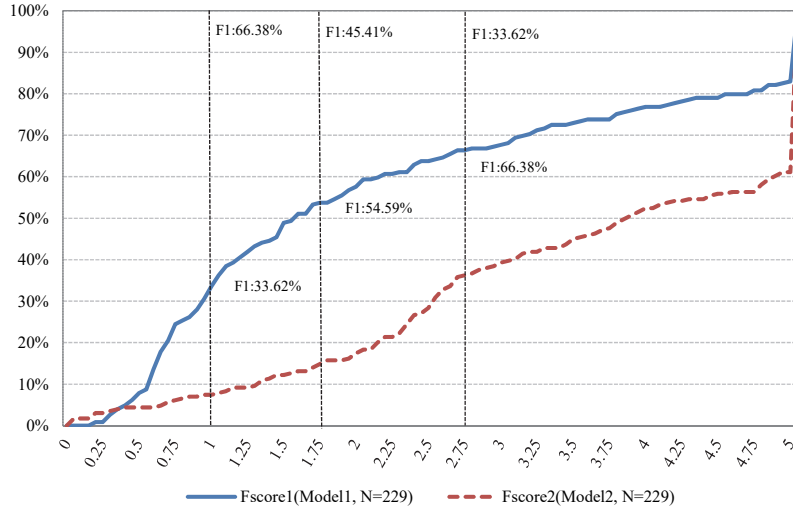
(4) Type II errors =  $77 / 229$

classification rate is 75.07%, and the sensitivity is 66.38%.

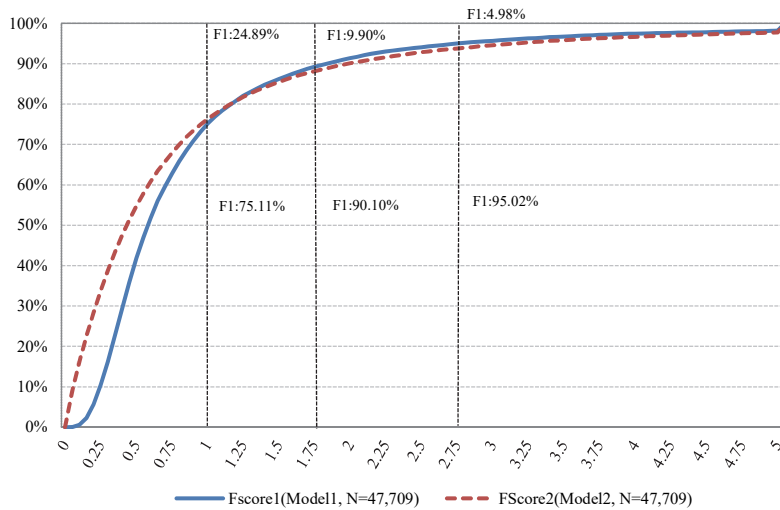
Since the Type I error is defined as the misclassifications of Non-FSF firms as FSF firms, the

FIGURE 1 CUMULATIVE DISTRIBUTION OF F-SCORES

Panel A : Accounting fraud firm-years



Panel B : Non-accounting fraud firm-years

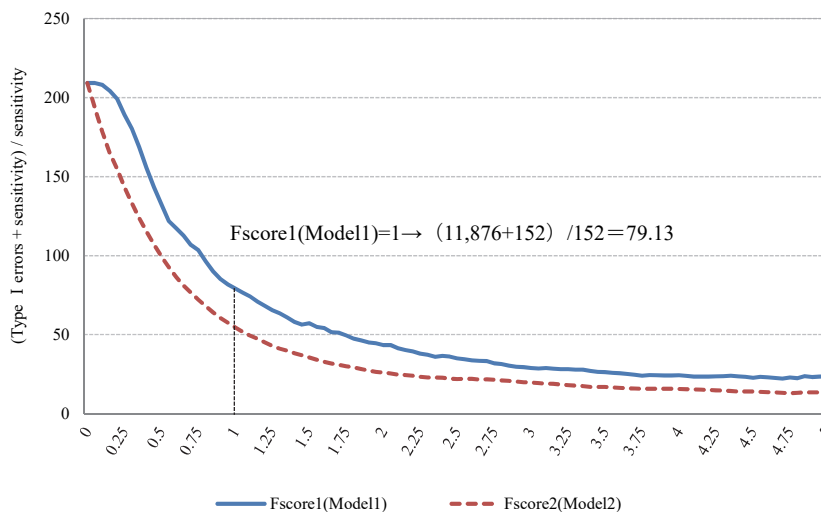


Type I error ratio is calculated as the number of pred-FSF firms among Non-FSF firms divided by Non-FSF firms; this is 24.89%. Furthermore, since the Type II error is defined as the misclassifications of FSF firms as Non-FSF firms, the Type II effort is defined as the number of pred-NonFSF firms among FSF firms divided by FSF firms; this is 33.62%.

We find that the correct classification rate of Model 2 is better. The sensitivity is greatly improved from 66.38% to 92.58%. Type I and Type II errors decrease more than in Model 1; Type II errors decrease drastically, from 33.62% to 7.42%.

We provide further insights into sensitivity and Type I and Type II error rates in Figure 1. Panel A provides the cumulative distribution of the *F-Score* for FSF firms, and Panel B presents

FIGURE 2 RELATIVE COST RATIO: AN ANALYSIS OF TYPE I ERRORS AND TYPE II ERRORS



the cumulative distribution for all Non-FSF firms. The upper part of the dotted line in Panel A indicates the sensitivity; the lower part indicates the Type II error, and the upper part of the dotted line in Panel B indicates the Type I error. For example, for an *F-Score* of 2.75, Panel A reveals that the sensitivity is 33.62% (the correct classification rate of FSF firms) and that the Type II error rate (the misclassification of Non-FSF firms as pred-FSF firms) is 66.38%, whereas the Type I error rate (the misclassification of FSF firms as pred-NonFSF firms) is 4.98% in Panel B. In general, the Type I error rate becomes smaller and the Type II error rate become greater as the *F-Score* increases. Similar results are found for Model 2, where the sensitivity increases in the higher *F-Score* region, and the Type II error rate is lower than in Model 1.

Figure 1 also provides helpful insights into the cost of these errors. The costs of Type I and Type II errors are not the same; the costs depend on who uses this information. For example, from an auditor's perspective, the cost of Type II errors seems to be far higher than that of Type I errors, since the former is the cost of overlooking an FSF firm. When an FSF firm goes undetected and is later revealed, the auditor is likely to be sanctioned by regulatory bodies and suffer a loss of reputation. Meanwhile, the cost of a Type I error is also not zero. It may result in lost fees, as the auditor may choose to drop a client. Because Type II errors are more costly to the auditor, an auditor is likely to prefer an *F-Score* cutoff that makes more Type I errors than Type II errors (Dechow et al. 2011, p. 62).

Figure 2 provides a tradeoff cost between Type I and Type II errors. The figure provides a ratio of the sum of the number of Type I errors (misclassification of FSF firms) and Sensitivity (correct classification of FSF and NonFSF firms) divided by the number of Sensitivity. We assume that the cost of investigating a firm is 1 yen (Y) and that an investigation always detects FSF firms if they exist. At an *F-Score* cutoff of 1.00, the relative cost ratio is Y79.13  $[(11,876+152)/152]$  for model 1. In other words, a cost of Y12,028 is incurred to avoid the 152 FSF firms. Therefore, if a cost of missing a FSF firm is Y79.13 or more, then an *F-Score* cutoff of 1.00 should be used by the auditor. If the cost is over Y230, then all firms should be investigated, because the *F-Score* cutoff is equal to zero. If the cost is less than Y5, it is cheaper not to do the

investigation and to just pay the extra cost of the FSF firms as they are identified. Because the relative cost ratio is about Y20, even the *F-Score* cutoff is around Y5.

#### (4) Marginal analysis of the prediction model

In this section, we evaluate the influence of each of the variables in the models for determining the magnitude of the *F-Score* (marginal effect analysis). We 1) calculate the value of the *F-Score* when all variables are held at their mean values, 2) recalculate the *F-Score* after moving one independent variable to its lower quartile value, holding all other variables at their mean values,<sup>18</sup> 3) recalculate the *F-Score*, moving the independent variable to its upper quartile value, 4) calculate the change in the *F-Score* across the interquartile range for that variable, and 5) repeat steps 2) through 4) for the next independent variable.

Table 14 summarizes the results. Panel A summarizes the descriptive statistics of each variable. Panel B shows the marginal effect of Model 1, revealing that the top three largest effect variables are *%Soft assets*, *CFO discretionary accruals AB*, and *Actual issuance*. It is noteworthy that the two variables relating to accruals have large marginal effects. Panel C shows the marginal effect of Model 2. We find that the top three largest effect variables are *%Soft assets*, *Actual issuance*, and *Book-to-market*. Among these, *Book-to-market* has an exceedingly high marginal effect. Finally, Panel D shows a matrix of the correlations among the variables. The correlations are all significant and are generally consistent with the results of the marginal effect analysis along with the above analysis. For example, the correlation coefficients in the Spearman's correlation in Model 1 increase in the order of *%Soft assets*, *Actual issuance*, and *CFO discretionary accruals AB*, and these three variables are the same as the three found in the marginal effect analysis.

#### (5) Robustness tests

Finally, we conduct robustness tests to investigate the sensitivity of our models to the industry and the time period examined. In our analysis of the detection rate in Table 12, we use the same sample both to develop our prediction model and to evaluate its effectiveness. This may cause overestimates of the predictive ability due to the hindsight bias. To address the issue, we use data covering 2000 to 2006 to build the prediction models and data covering 2007 to 2010 to evaluate their predictive ability. The analyses conducted in the previous section are replicated based on the above out sample seen in Table 15. Panels A through C show the results of the parameter estimation (the same as is shown in Table 10), prediction ability (the same as is shown in Table 12), and Type I and Type II errors (the same as is shown in Table 13), respectively. Panel A summarizes the estimated parameters for the period from 2000 to 2006. There is little difference between the results of Table 10 and Table 15. We find that two variables, the change of cash flow margin (*Change in cash margin*) and the change of free cash flow (*Change in free cash flows*), are selected in Model 1, instead of abnormal operating cash flow (*AB cash flow*) and conservative score (*C score AR*). The panel also shows that, in Model 2, the abnormal operating cash flow is deselected.

Panel B shows the result for the predictive ability of our model for the sample from 2007 to 2010. The results show the same high capacity, whereby the fifth portfolio includes fraud firms at

<sup>18</sup> For indicator variables such as *Actual issuance*, we set the upper quartile value at 1 and the lower quartile value at 0. For variables with negative estimated parameters, we swap the values of the upper and lower quartiles to calculate the *F-Score*.

TABLE 14 MARGINAL EFFECT ANALYSIS

*Panel A: Descriptive statistics*

	% Soft assets	CFO discretionary accruals AB	Actual issuance	C score AR	AB cash flow	CORP	Book-to-market
Mean	66.302	0.032	0.266	0.107	0.000	2.234	1.237
Std Dev	14.401	0.035	0.442	0.970	0.087	7.523	0.873
Lower quartile	57.900	0.010	0.000	-0.007	-0.034	0.024	0.613
Median	67.626	0.021	0.000	0.075	0.001	0.059	1.054
Upper quartile	76.384	0.042	1.000	0.175	0.036	0.317	1.645
1%	25.166	0.000	0.000	-0.315	-0.269	0.003	0.000
99%	94.082	0.186	1.000	0.690	0.229	37.196	4.145

*Panel B: Marginal effect for Model1*

	% Soft assets	CFO discretionary accruals AB	Actual issuance	C score AR	AB cash flow	CORP
Coefficient estimate	0.038	7.654	0.945	0.049	-1.337	0.032
F <sub>score</sub> at upper quartile	0.938	0.686	1.272	0.640	0.667	0.600
F <sub>score</sub> at lower quartile	0.462	0.536	0.496	0.634	0.608	0.594
Interquartile marginal change in F <sub>score</sub>	0.476	0.149	0.776	0.006	0.059	0.006

*Panel C: Marginal effect for Model2*

	% Soft assets	CFO discretionary accruals AB	Actual issuance	AB cash flow	CORP	Book-to-market
Coefficient estimate	0.036	2.639	0.666	-1.690	0.024	-1.455
F <sub>score</sub> at upper quartile	0.541	0.387	0.614	0.399	0.360	0.933
F <sub>score</sub> at lower quartile	0.279	0.355	0.316	0.355	0.357	0.209
Interquartile marginal change in F <sub>score</sub>	0.261	0.031	0.298	0.044	0.003	0.724

*Panel D: Correlation between F<sub>score</sub> and its components*

	% Soft assets	CFO discretionary accruals AB	Actual issuance	C score AR	AB cash flow	CORP	Book-to-market
Spearman correlation							
F <sub>score</sub> (Model1)	0.674	0.324	0.560	0.083	-0.302	0.098	
F <sub>score</sub> (Model2)	0.284	0.179	0.457		-0.071	0.174	-0.854
Pearson Correlation							
F <sub>score</sub> (Model1)	0.294	0.437	0.355	0.118	-0.362	0.407	
F <sub>score</sub> (Model2)	0.131	0.195	0.159		-0.162	0.182	-0.286

Note: All variables are defined in Table 2.

TABLE 15 ROBUSTNESS TESTS

*Panel A: Logistic regressions*

	Model1			Model2		
	Coefficient estimate	Wald Chi-square	p-value	Coefficient estimate	Wald Chi-square	p-value
<i>Intercept</i>	-8.463	199.254	<.0001	-6.956	120.882	<.0001
<i>% Soft assets</i>	0.037	22.522	<.0001	0.042	28.576	<.0001
<i>CFO discretionary accruals AB</i>	8.140	17.275	<.0001	3.676	3.271	0.071
<i>Actual issuance</i>	0.808	12.746	<.0001	0.523	4.859	0.028
<i>C score AR</i>						
<i>AB cash flow</i>						
<i>CORP</i>	0.030	17.046	<.0001	0.016	4.584	0.032
<i>Book-to-market</i>				-2.493	48.592	<.0001
<i>Change in cash margin</i>	-0.001	4.511	0.034			
<i>Change in free cash flows</i>	-2.261	3.991	0.046			
AR			0.535			0.872
Accounting fraud firm-years			93			93
Non-accounting fraud firm-years			13,775			13,822

*Panel B: Validation of the prediction model accuracy*

	Model1			Model2		
	N	Min. F-score	% of total	N	Min. F-score	% of total
Quintile1						
Accounting fraud firm-years	2	0.220	2.06%	2	0.000	2.06%
Non-accounting fraud firm-years	2,844	0.037	20.12%	2,853	0.000	20.12%
Quintile2						
Accounting fraud firm-years	7	0.418	7.22%	0	0.000	0.00%
Non-accounting fraud firm-years	2,840	0.351	20.09%	2,856	0.023	20.14%
Quintile3						
Accounting fraud firm-years	13	0.525	13.40%	1	0.192	1.03%
Non-accounting fraud firm-years	2,833	0.513	20.04%	2,855	0.104	20.13%
Quintile4						
Accounting fraud firm-years	16	0.742	16.49%	1	0.504	1.03%
Non-accounting fraud firm-years	2,831	0.736	20.03%	2,855	0.284	20.13%
Quintile5						
Accounting fraud firm-years	59	1.2123	60.82%	93	0.798	95.88%
Non-accounting fraud firm-years	2,787	1.1992	19.72%	2,762	0.767	19.48%

*Panel C: Verification of the type I errors and type II errors of the prediction model*

Observed	Model1 predicted			Model2 predicted		
	Misstate	No-misstate		Misstate	No-misstate	
Misstate	66	31	97	92	5	97
No-misstate	3,675	10,460	14,135	2,147	12,034	14,181
	3,741	10,491	14,232	2,239	12,039	14,278
Misstate	68.04%	31.96%	0.68%	94.85%	5.15%	0.68%
No-misstate	26.00%	74.00%	99.32%	15.14%	84.86%	99.32%
Correct classification	73.96%			84.93%		
Sensitivity	68.04%			94.85%		
Type I errors	26.00%			15.14%		
Type II errors	31.96%			5.15%		

*Note:* All variables are defined in Table 2. Shaded variables in panel A are not used in the prediction model.

a rate of 60.82% in Model 1 and 95.88% in Model 2. Finally, Panel C reveals the results on Type I and Type II errors for the later period of 2007 to 2010. The Sensitivity is 68.04% for Model 1 and 94.85% for Model 2. These error rates are largely the same as the original results seen in Table 13. Thus, the results obtained using the subperiod samples show no significant differences from our main results.

Finally, to test the impact of industry differences on our results, we build models with industry indicator variables for *Electronics*, *Construction*, *Trading*, and *Services*, industries with relatively large numbers of firms. Although some industry indicator variables are selected as significant variables for both Model 1 and Model 2, the other selected variables are almost the same as those in the main results, and the models' predictive ability shows no significant differences from the main results (in untabulated results).

## 6. Conclusion

In this study, we investigate the relationship between accounting fraud and accounting information for Japanese firms. Our accounting fraud sample consists of two subsamples: 1) firms that have been accused or had administrative monetary penalties imposed by the SESC and 2) firms that have announced accounting fraud at timely disclosure disclosures. Specifically, we 1) explore the characteristics of accounting fraud firms by analyzing financial information obtained from the annual reports of Japanese firms and 2) develop a model to predict accounting fraud based on the characteristics of Japanese fraud firms.

To identify the characteristics of fraud firms, we focus on 38 variables for the eight factors of accruals quality, performance, nonfinancial measures, off-balance-sheet activities, market-related incentives, conservatism, real-activities manipulation, and Japanese-specific factors.

Through our univariate analysis and model building process, we find that accrual quality, market-related incentive, real activities manipulation, conservatism, and Japanese-specific factors are generally useful for detecting accounting fraud. Our analysis contributes to prior studies by clarifying the importance of focusing on the various dimensions of earnings quality and country-specific factors.



For example, the marginal effect analysis reveals that soft asset ratio, actual issuance, and the absolute value of discretionary accruals are useful when we focus on the accounting information in annual reports (i.e., the variables in Model 1). Further, when we add stock-related variables to the model, we find that the book-to-market has a higher marginal effect. In general, our prediction models for identifying accounting fraud have stronger predictive ability than do those developed by previous studies. Our models can be used widely in various aspects of accounting and finance practice.

Our results have several important implications for business practice. For example, auditors and regulatory bodies could use the model to reduce their investigation costs and improve the accuracy of their judgments. Furthermore, capital providers such as stock investors and banks could also improve their decision-making by estimating the probability of accounting fraud using our prediction model. Although investment strategies on the basis of accrual quality such as accrual anomalies have already become popular, stock investors in the Japanese market might be able to consider their investment strategies based on the *F-Score*. Furthermore, creditors could also use the *F-Score* to evaluate the credit risk in lending or doubtful accounts receivable.

Finally, this study has several limitations. Our main accounting fraud samples are based on firms that have been accused or had administrative monetary penalties imposed by the (SESC). We cannot deny the possibility that the SESC has misclassified an accounting fraud firm as a non-fraud firm because its investigatory power is limited. To address this issue, we supplemented the observations by using timely disclosure information and conducted additional analyses based on the extended sample. However, we cannot entirely rule out the sample selection issue.

Furthermore, we do not examine the effect of corporate governance on accounting fraud except for two Japanese-specific factors.<sup>19</sup> As Dechow et al. (2011) also argue, this is because the availability of data on corporate governance are limited and costly, reducing the practicability of our prediction models. It is important to examine the relationship between accounting fraud and corporate governance in future research. Finally, it might be useful to focus on the information concerning internal control such as significant deficiency and material weakness for predicting accounting fraud.

## REFERENCES

- ALLEN, E., C. LARSON, AND R. SLOAN. 2013. Accrual reversals, earnings and stock returns. *Journal of Accounting and Economics* 56 (1): 113-129.
- BARTON, J., AND P. SIMKO. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* 77 (Supplement): 1-27.
- BASU, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24 (1): 3-37.
- BENEISH, M. D. 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16 (3): 271-309.

---

<sup>19</sup> For example, we can point out board independence, executive compensation contracts, and other ownership structure elements (such as ownership by foreign investors or management).

- BENEISH, M. D. 1999a. The detection of earnings manipulation. *Financial Analysts Journal* 55 (5): 24–36.
- BENEISH, M. D. 1999b. Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74 (4): 425–457.
- BRAZEL, J. F., K. L. JONES, AND M. F. ZIMBELMAN. 2009. Using nonfinancial measures to assess fraud risk. *Journal of Accounting Research* 47 (5): 1135–66.
- DECHOW, P. M., R. G. SLOAN, AND A. P. SWEENEY. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–226.
- DECHOW, P. M., R. G. SLOAN, AND A. P. SWEENEY. 1996. Causes and consequences of earnings misstatement: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (1): 1–36.
- DECHOW, P. M., GE, W., AND C. SCHRAND. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50 (2-3): 344–401.
- DECHOW, P. M., GE, W., LARSON, C. R. AND R.G. SLOAN. 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28: 17–82.
- ETTREDGE, M., L. SUN, P. LEE, AND A. ANANDARAJAN. 2008. Is earnings fraud associated with high deferred tax and/or book minus tax levels? *A Journal of Practice and Theory* 27 (1): 1–33.
- GRAHAM, J., C. HARVEY, AND S. RAJGOPAL. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40 (1-3): 3–73.
- GRUBBS, F. E. 1969. Procedures for detecting outlying observations in samples. *Technometrics* 11 (1): 13–14.
- KASZNIK, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37 (1): 57–81.
- KHAN, M. N., AND R. L. WATTS. 2009. Estimation and empirical properties of a firm-year measure of accounting conservatism. *Journal of Accounting and Economics* 48 (2-3): 132–150.
- KOTHARI, S. P., A. LEONE, AND C. WASLEY. 2005. Performance-matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.
- LEE, T., INGRAM, R., AND T. HOWARD. 1999. The difference between earnings and operating cash flow as an indicator of financial reporting fraud. *Contemporary Accounting Research* 16(4): 749–786
- OKUMURA, M. 2014. *Misstatement of Earnings Information and the Stock Market*. : Chuo-Keizaisha. (in Japanese).
- OSHIRO, N. 2014. Research on early detection of financial statement fraud, *FSA Institute Discussion Paper Series*, DP 2014-6, Financial Research Center, Financial Services Agency Government of Japan (in Japanese).
- RAJGOPAL, S., T. SHEVLIN, AND M. VENKATACHALAM. 2003. Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. *Review of Accounting Studies* 8 (4): 461–492.
- RICHARDSON, S., R. SLOAN, M. SOLIMAN, AND I. TUNA. 2005. Accrual reliability, earnings persistence, and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.
- ROYCHOWDHURY, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335–370.
- SHUTO, A. 2010. *Earnings Management of Japanese Firms—Theories and Practices*. : Chuo-Keizaisha (in Japanese).
- STEFANSKY, W. 1972. Rejecting outliers in factorial designs. *Technometrics* 14 (4): 469–479.

- SUDA, K. 2000. *Positive Theory of Financial Accounting*. : Hakuto Shobo (in Japanese).
- TEOH, S. H., I. WELCH, AND T. J. WONG. 1998. Earnings management and the long-run market performance of initial public offerings. *The Journal of Finance* 53 (6): 1935–1974.
- WATTS, R. L. 2003. Conservatism in accounting part I: Explanations and implications. *Accounting Horizons* 17 (3): 207–221.
- XU, M., AND C. ZHANG. 2009. Bankruptcy prediction: the case of Japanese listed companies. *Review of Accounting Studies* 14 (4): 534–558.

## APPENDIX

## A. Measurement of discretionary accruals

We use three types of modified versions of Jones' (1991) model to measure discretionary accruals. Specifically, we use the following three models: the modified Jones model in Dechow et al. (1995; *MJ discretionary accrual*), the performance-matched Jones model in Kothari et al. (2005; *PM discretionary accruals*), and the CFO modified Jones model in Kasznik (1999; *CFO discretionary accruals*). The methods of calculation are as follows:

(1) The modified Jones model

$$WC\ accruals_t = \alpha + \beta_1(\Delta Sales_t - \Delta Receivables_t) + \beta_2 PPE_t + \varepsilon_t$$

where,

*WC accruals* = working capital accruals,

$\Delta Sales$  = change in sales,

$\Delta Receivables$  = change in account receivables, and

*PPE* = net property, plant and equipment.

(2) The performance-matched Jones model

$$WC\ accruals_t = \alpha + \beta_1(\Delta Sales_t - \Delta Receivables_t) + \beta_2 PPE_t + \beta_3 ROA_t + \varepsilon_t$$

where,

*ROA* = return on assets; net earnings / total assets.

(3) CFO modified Jones model

$$Total\ accruals_t = \alpha + \beta_1(\Delta Sales_t - \Delta Receivables_t) + \beta_2 PPE_t + \beta_3 CFO_t + \varepsilon_t$$

where,

*Total accruals* = total accruals as the net current profit after tax minus cash flow from operation,

*CFO* = cash flow from operations.

$\Delta$  is a change in the value of a variable, and  $t$  stands for fiscal year. All variables are scaled by lagged total assets. The model is estimated cross-sectionally for each industry in a given fiscal year according to the Nikkei Industry Classification Code (*Nikkei gyousyu chu-bunrui*). We combined some industries in order to include more than 10 firms at least. Using the estimated coefficients of the model, we measured non-discretionary accruals (*NDA*). The difference between total accruals and measured non-discretionary accruals is a proxy for discretionary accruals (*DA*).

For the performance-matched Jones model in (2), we estimate two models based on two types of *ROA*, at current year ( $ROA_t$ ) or at previous year ( $ROA_{t-1}$ ), following Kothari et al. (2005). Although the estimation results of the two models are similar to each other, we employed the model with  $ROA_t$  because the explanatory power ( $Adj.R^2$ ) of the model with  $ROA_t$  is higher than that with  $ROA_{t-1}$ . The definition of accruals in CFO modified Jones model is based on the most standard definition used in empirical research on Japanese firms (Shuto 2010).

The untabulated results indicate that all independent variables of all models are consistent with the expected signs on average. The explanatory power of the CFO modified Jones model has the highest value among the models, consistent with the previous study (Shuto 2010).

## B. Measurement of conservatism

We estimate accounting conservatism following the method of Khan and Watts (2009). Khan and Watts (2009) extended the cross-sectional measurement model of conditional conservatism proposed by Basu (1997) in order to estimate conservatism for an individual firm. The model of Basu (1997) is as follows:

$$X_t = \beta_1 + \beta_2 D_t + \beta_3 R_t + \beta_4 D_t R_t + \varepsilon_t \quad (1)$$

where,

$X$  = net earnings divided by the market value at beginning of fiscal year,

$R$  = annual return,

$D$  = a dummy variable equal to 1 when  $R < 0$  and equal to 0 otherwise,

where  $\beta_3$  is the good news timeliness measure, and  $\beta_4$  is conditional conservatism. We assume that the timeliness and conditional conservatism are the liner functions of three firm-specific characteristics: firms size, book-to-market, and leverage.

$$Gscore = \beta_3 = \mu_1 + \mu_2 Size_t + \mu_3 M / B_t + \mu_4 Lev_t \quad (2)$$

$$Cscore = \beta_4 = \lambda_1 + \lambda_2 Size_t + \lambda_3 M / B_t + \lambda_4 Lev_t \quad (3)$$

$$\begin{aligned} X_t = & \beta_1 + \beta_2 D_t + R_t(\mu_1 + \mu_2 Size_t + \mu_3 M / B_t + \mu_4 Lev_t) \\ & + D_t R_t(\lambda_1 + \lambda_2 Size_t + \lambda_3 M / B_t + \lambda_4 Lev_t) \\ & + (\delta_1 Size_t + \delta_2 M / B_t + \delta_3 Lev_t + \delta_4 D_t Size_t + \delta_5 D_t M / B_t + \delta_6 D_t Lev_t) \\ & + \varepsilon_t \end{aligned}$$

We substitute Eqs. (2) and (3) into regression (1), and perform annual cross-sectional regression as follows:

(4)

where,

$Size$  = natural logarithm of market capital,

$M/B$  = book-to-market ratio,

$Lev$  = leverage, net interesting bearing debt divided by total asset.

In estimating the regressions, we winzorized these variables due to outliers. Specifically, four variables,  $X$ ,  $Size$ ,  $M/B$ , and  $Lev$ , are winsorized at 1% and 99%. Furthermore, we adopt the Smilnov-Grabbs test<sup>20</sup> to detect outliers.

In our regression models, we use two types of returns: the raw returns and the market-adjusted returns. The regression results reveal that the coefficients on the independent variables of both models to estimate the Cscore have the expected signs. We also find that the explanatory power of the model with market-adjusted returns is higher than is that of the model with raw returns.

### C. Measurement of real-activities manipulation

<sup>20</sup> For more details on the Smilnov-Grabbs test, please refer to Grubbs (1969) and Stefansky (1972).

We measure real-activities manipulation in accordance with the method of Roychowdhury (2006). As discussed above, we expect that managers may perform manipulations and discretionary activities such as control of their sales, reducing discretionary costs/expenses, and overproduction. In order to capture these manipulations, we measure three variables: (1) abnormal cash flow from operation (*AB cash flow*), (2) discretionary expenses (*Discretionary expense*), and (3) abnormal production cost (*AB product cost*). The estimation methods for each variable are as follows:

(1) Abnormal cash flow from operation

$$CFO_t / A_{t-1} = \alpha + \beta_1(1 / A_{t-1}) + \beta_2(Sales_t / A_{t-1}) + \beta_3(\Delta Sales_t / A_{t-1}) + \varepsilon_t$$

where,

*CFO* = cash flow from operation,

*A* = total assets, and

*Sales* = amount of sales.

(2) Discretionary expense

$$DE_t / A_{t-1} = \alpha + \beta_1(1 / A_{t-1}) + \beta_2(Sales_{t-1} / A_{t-1}) + \varepsilon_t$$

where,

*DE* = discretionary expenses: the sum of following costs/expenses, research and development, advertisement, other costs of sales/marketing, board members' compensation, labors cost/welfare expense.

(3) Abnormal production cost

$$PD_t / A_{t-1} = \alpha + \beta_1(1 / A_{t-1}) + \beta_2(Sales_t / A_{t-1}) + \beta_3(\Delta Sales_t / A_{t-1}) + \beta_4(\Delta Sales_{t-1} / A_{t-1}) + \varepsilon_t$$

where,

*PD* = cost of production; the sum of cost of goods sold, inventory at the end of fiscal year minus inventory at the beginning of fiscal year.

We estimate the parameters by a cross-sectional industry regression in year and industry, the same procedure as for discretionary accruals. By calculating the residual in each model, we estimate three real-activities manipulations as the difference between the actual and estimated value (residual).

#### D. Calculation method of AR

The AR is a variable used to measure the accuracy of prediction models. First, we explain the CAP curve in order to understand the AR. The CAP curve is the result of a plot on the basis of the *F-Score* in descending order (i.e., the *F-Score* becomes higher in origin), of which the X-axis is the ratio of total (accumulated) samples, and the Y-axis is the ratio of accounting fraud samples only (see the "CAP curve of estimated model" in the figure below).

If the prediction model can discriminate (capture) among all firms in terms of accounting fraud (generally referred to as a "perfect model"), the CAP curve corresponds to the "CAP curve in perfect model" in the figure.

On the other hand, if the prediction model cannot capture the firms in terms of accounting fraud (generally referred to as a "random model"), the situation is the same as when we discriminate among fraud samples randomly (i.e., when we use dice to determine the probability of accounting fraud instead of using a prediction model). For the random model, the CAP curve is "CAP curve in random model (45 degree)."

In general, the CAP curves of the prediction models are positioned between two curves (i.e., the perfect model and the random model). It shows higher accuracy when the CAP curve in the estimated model swells out widely. The AR is the ratio of the area between the estimated model and the random model (Region B) divided by the sum of the area between the perfect model and the estimated model (Region A) and Region B as follows:

$$AR = \text{Region B} / (\text{Region A} + \text{Region B})$$

In the perfect model, AR becomes 1 because Region A is 0. In the random model, AR becomes 0 because Region B is 0. Therefore, the larger AR closer to 1 (i.e., closer to the perfect model) indicates a better prediction result.

