



THE VALUE OF RESTATEMENT TO FRAUD PREDICTION

Shi QIU^{1,2} , Hong-Qu HE^{1*}, Yuan-sheng LUO³

¹*Business School, Central South University, Changsha, China*

²*School of Economics and Management, Changsha University, Changsha, China*

³*School of Computer & Communication Engineering,
Changsha University of Science & Technology, Changsha, China*

Received 18 December 2018; accepted 30 May 2019

Abstract. A financial report restatement reflects errors in the previous financial statement, and thus it increases investors' doubt about the credibility of the financial statement. The primary objective of this paper is to examine whether restatement announcements imply increased fraud risks in Chinese firms in the context that up to one quarter of listed companies have restated their financial reports in China, and explore the implications of the content, severity and reasons for restatements with respect to fraud. In this paper, firms with financial restatements prove to be more likely to be labeled as fraudulent by regulators in China. Second, the following results also are revealed: (1) financial statements, except balance sheet restatements, provide insights into the revelation of fraudulent behaviors, (2) the severity of restatements is positively correlated with future fraud disclosures, and (3) restatements due to negligence are positively correlated with future fraud occurrences. These results imply that restatement announcements and their different characteristics provide important information for detecting financial statement fraud.

Keywords: restatement, fraud, innocent error, negligence, severity of restatement, content of restatement, reasons of restatement, propensity score matching method, fraud risk.

JEL Classification: M41, M42, G34.

Introduction

According to the report of Association of Certified Fraud Examiners (2016), the total losses for all fraud cases exceeded \$6.3 billion, with 23.2% of cases causing losses of \$1 million or more. Financial statement fraud occurred in less than 10% of cases but caused a median loss of \$975,000 per scheme, and the median loss for single cases was \$150,000. Meanwhile, restatements increased by 101% from 2000 to 2011 in the US. Hennes, Leone, and Miller (2008) and Kim, Baik, and Cho (2016) detected financial statement fraud by distinguishing irregularities from errors in restatements. BenYoussef and Breton (2016) identified fraud by

*Corresponding author. E-mail: hehongqu@126.com

analyzing the content of corporate 8-K restatement reports. All of these studies are devoted to the problem that which kinds of restatements could be ascribed to fraud. However, research on the connection between restatements and firms' incentives to participate in fraud remains scarce. When internal controls and supervision systems are weak or have failed, an accounting restatement could be the first indication that accounting fraud has occurred (Land, 2010). Restatements are evidences of manipulation and internal control weaknesses within the company (Plumlee & Yohn, 2015). These indicate that firms' restatements can be early warning signs for investors to suspect fraud.

In China, the number of restatements increased by 210% from 2009 to 2013. The restatements have received scant attention from academics, in contrast to the high attention which American restatements have received from U.S. academics. Given the differences in institution and legislation between the U.S. and China, the findings in U.S. cannot be generalized to the context in China. Srinivasan, Wahid, and Yu (2014) found that U.S. listed foreign companies under a weak rule of native law are different from U.S. local firms in the characteristics of their restatements. The cost of accounting misstatements in China is much cheaper than that in a mature market (Wang & Wu, 2011). Accordingly, the purpose of this paper is to test whether the Chinese firms with a restatement are also more likely to be found guilty of corporate fraud in the future.

The paper begins with analyzing the restatements filed from 2009 to 2013 and the frauds dated from 2009 to 2016 in China to provide empirical evidence that restatement firms are more inclined to conduct fraudulent activities relative to the non-restatement sample. Furthermore, S. Scholz (2013) reported that fewer restatements involved serious issues, such as irregularities or fraud, while error restatements accounted for 65% of the total restatements. Christensen (2010) also stated that accounting should pay more attention to errors since errors were essential for the updating of beliefs. Therefore, this paper further examines how different kinds of restatements influence the revelation of future fraud.

Our study contributes to the literature in several ways. First, prior studies tend to focus on identifying the fraud in restatements under the proposition that the restatements use variants of the words "fraud" or "irregularity" (Hennes et al., 2008). However, Chinese enterprises have seldom utilized the words "fraud" and "irregularity" in their restatements, thus it is difficult to find the relationship between restatements and fraud. We fill this gap in order to assess the relationship of restatements with future fraud detection. Second, the intent of managerial restatements is always difficult to capture. Hennes et al. (2008) have classified restatements into errors and irregularities (fraud) according to managerial intent. In contrast to those studies, this paper seeks to better understand the characteristics of restatements in order to provide details for proving fraud intent. The paper also provides insights for regulators, researcher and auditors. The government always pays little attention to the implications of restatements. However, the findings in this paper suggest that financial restatements are red flags of future fraud, which will prompt regulators to pay more attention to financial restatements. In addition, researchers and auditors who are interested in fraud detecting should focus on the characteristics of restatements and their relationship with firms' future fraud. Furthermore, our studies also focus investors' attention on the implications of restatements and their benefits by identifying which restatement contents can predict fraud.

The rest of the paper is structured as follows. Section 1 reviews the related literature and develops our hypotheses. Section 2 defines the variables and explains our models. Section 3 reports the empirical findings and the last Section concludes.

1. Literature review

1.1. Restatements for fraud prediction

According to the Financial Reporting and Auditor Task web page of the SEC's website, "The financial statement restatements are areas susceptible to fraudulent financial reporting". As we all know, restatements can lead to a decrease in accounting credibility (Trompeter, Carpenter, Desai, Jones, & Riley, 2013; Hirschev, Smith, & Wilson, 2015). Recent studies suggest that non-restatement firms will be less exposed to managerial risk and bankruptcy risk compared to restatement firms (Dorminey, Fleming, Kranacher, & Riley, 2012). Hennes et al. (2008) and Kim et al. (2016) focused on how to distinguish fraud-related restatements from error-related restatements. Few had built a direct connection between financial restatements and future revelations of fraud, except for Fang, Huang, and Wang (2017), who insisted that errors discouraged fraud by lowering the firms' value or incentivized fraud by proving a camouflage effect. For this reason, the link between restatements and unrevealed fraud is explored in this paper. The first hypothesis is developed as follows:

Hypothesis 1: There is significant relationship between financial report restatements and future fraud revelations.

1.2. Characteristics of restatements and fraud prediction

Based on the above studies, we further question whether different characteristics of restatements are more indicative of future fraud than others. Palmrose, Richardson, and Scholz (2004) identified the characteristics of restatements by measuring the core earnings restatements, the number of accounts affected, the change in the net income to assets, and the number of years restated. Wei, Li, and Chen (2010) and Zhang, Huang, and Habib (2018) also studied market reactions to different types of Chinese firms' restatement, including the types, reasons, content and scope of restatement announcements, etc. This paper examines the effects of the content, severity and reasons for restatements on subsequent fraud revelations.

1.2.1. Content of restatement and fraud prediction

In view of the U.S. General Accounting Office (2003), Jiang, Cui, and Wang (2010) and Ma, Du, and Zhang (2018) on the classification of financial restatements, we partitioned the content of restatements into six categories: (1) core earnings, (2) noncore earnings, (3) reclassification of cash flow items, (4) reclassification of balance sheet items, (5) footnotes of financial statements, and (6) innocent errors. Through reading the numerous restatement disclosures, we further classify the footnotes of financial restatements into the following categories: related-party transactions; acquisitions and mergers; shareholders, executives, suppliers and customers; revenue and cost analysis; restricted assets and others.

Lee, Ingram, and Howard (1999) found that the difference between the adjusted net income and operating cash flow for fraudulent firms is significantly greater in the fraud pre-discovery years. Therefore, the earnings and cash flow relationship can be used as a signal of potential fraudulent financial reporting. Firth, Rui, and Wu (2011) also studied the relationship between restatements based on income statement and balance sheet corrections and fraud in China. None of these studies examined the individual economic consequences of innocent errors in restatements as a result of their small effects on markets and investors. Tan and Young (2015) found that little restatement firms have higher profitability and a lower likelihood of reporting misstatements compared with big restatement firms. However, Baucus (1994) argued that corporate illegality runs the gamut from unintentionally illegal corporate behavior to intentionally illegal corporate behavior to corporate crime. There is also few literature on the footnote restatements (Fang, Lobo, Zhang, & Zhao, 2018). All of the above discussion leads to the following hypotheses:

Hypothesis 2a: There is significant relationship between the restatements related to corrections of core earnings, noncore earnings, cash flow and balance sheet items and future fraud revelation.

Hypothesis 2b: There is significant relationship between innocent error restatements and future fraud revelations.

Hypothesis 2c: There is significant relationship between footnote restatements and future fraud revelations.

1.2.2. Severity of restatement and fraud prediction

The more severe the restatement is, the more serious the accounting problem and the higher the probability of a fraud revelation are. We use three measures for the severity of the restatement. The first is the Duration, which is measured by the number of quarters that the company restates (Badertscher, & Burks, 2011; Schmidt & Wilkins, 2013; BenYoussef & Khan, 2018). The second is the Revision Number, which is denoted by the number of revisions in one year (Wang & Wu, 2011; Srinivasan et al., 2014). The third is the Revision Accounts, which is calculated by the Sum of the account groups involved in the restatement (Palmrose et al., 2004; He, Sarath, & Wans, 2019). The account groups are revenues, the costs of sales, operating expenses, one-time items, non-operating expenses and other. All these proxies are indicators of the degree of accounting errors in restatements.

Singer and Zhang (2018) used the timeliness of misstatement discoveries as a proxy for reliable financial reports. The high detection speed of restatements is beneficial to improving the financial reporting reliability (Badertscher & Burks, 2011). Schmidt and Wilkins (2013) found that corporate restatements with fraud are more likely to delay their disclosure. Furthermore, Wang and Wu (2011) contended that the low frequent restatements reflected a high accounting quality, but Srinivasan et al. (2014) found an opposite relationship between the restatement frequency and accounting quality, indicating that a low rate of restatements represents represented opportunistic reporting. Palmrose et al. (2004) found that fraud restatements affected more accounts. He et al. (2019) argue that restatements with material-weakness disclosure of internal control would influence more accounts. Thus, the next hypothesis is as follows:

Hypothesis 3: The larger the number of restatements is, the longer the time periods are, and the more accounts that are involved in the restatements are, the higher the likelihood of fraud for a firm is.

1.2.3. Reasons for restatement and fraud prediction

It is well known that Chinese listed firms attributed their restatements to negligence, mistakes and unknown reasons in the first paragraphs of their restatement reports. Given that restatements generally reflect badly on management, managers may have incentives to act strategically and choose a less transparent restatement disclosure venue (Murphy & Dacin, 2011; Chung & Mccracken, 2014; Demirkan & Platt, 2018). According to the expanded fraud model presented by Dorminey et al. (2012), the fraudster as a decision maker will make efforts to conceal his behavior post-fraud. In addition, Corona and Randhawa (2018) find that high-quality companies more often admit mistakes than low quality companies. These studies indicate that restated companies have strategic motives regarding how they disclose restatements. The analyses can improve our understanding of the reasons for restatements.

Hypothesis 4: The restatement attributed to negligence is a positive predictor of future fraud.

2. Empirical analysis

2.1. Sample selection and data description

The sample data are composed of all A-shares companies listed on the Chinese Stock Exchanges during 2009–2016. The accounting data are extracted from the Wind and China Stock Market Accounting Research (CSMAR) databases. The study started from 2009 because financial statement fraud became prevalent after the global financial crisis in 2008. The restatement data is collected during the period from 2009 through 2013. Liou (2008) found an average time lag of around three years from the fraud to its detection. Therefore, We obtain the fraud data from 2009 to 2016, which is similar with the research (Donelson, Ege, & McInnis, 2017). For example, if a restatement occurred in 2013, the calculation of fraud sample would start from 2013 to 2016 which is released after the issuance of restatement. After deleting the financial service companies and observations with missing corporate governance, ownership and audit data, the final sample includes 10,464 firm-years for 2009 through 2013, 1,188 of which have fraudulent behavior and 1,588 have made restatement announcements respectively.

To identify instances of fraud, this paper gathers securities violations from CSMAR database, and selects cases such as false records, significant omissions, disclosure delays, improper accounting treatments, misleading statements, and asset or profit misstatements that were revealed from 2009 to 2016. We create the Fraud variable, which is equal to 1 if a fraud is published in three years after the disclosure date of the corporate restatements, otherwise is 0. In addition, we create the Fraud_no_reps variable, which is 1 if a fraud is released for the first firm-year in the three year window in order to avoid counting fraud cases repeatedly across multiple firm-years (e.g., the fraud revealed in 2011 may be both counted in 2009 and

2010 fraud sample, Fraud is set to 1 in both year 2009 and 2010, and Fraud_no_reps is set to 1 only in year 2009). Therefore, the sizes of our Fraud sample and Fraud_no_reps sample are 1188 and 458 respectively.

To obtain the restatement data, we collect our data from the Wind database by searching for the keywords “errors”, “irregularities”, “amend”, “revisions”, “supplements”, “retroactive adjustment” and “adjustments” in companies’ interim and annual disclosures. In China, restatements are triggered by a series of reasons, such as changes in accounting principles, revisions of estimates, and corrections of errors or irregularities. Nevertheless, the changes in accounting principles and revisions of estimates are less likely connected with earnings management (Ahmed & Goodwin, 2007). Therefore, this paper only studies the restatements related to the corrections of errors or irregularities. Overall, the final sample consists of 1588 firms’ restatements from January 1, 2009 to December 31, 2013. Table 1 provides the details of our sample’s construction.

Table 1. Sample construction by year

	2009	2010	2011	2012	2013	Total
	(n = 1522)	(n = 1953)	(n = 2223)	(n = 2376)	(n = 2390)	(n = 10464)
Fraud	213	230	226	266	253	1188
Fraud_no_reps	213	82	60	63	40	458
Restatement	204	205	230	336	613	1588

Note: This table presents the quantity of the fraud and restatement by year. The restatement sample is 2009 through 2013 and the fraud sample is 2009 to 2016. Fraud is equal to 1 if fraud is published in three years after the disclosure date of the corporate restatements. The Fraud_no_reps variable is 1 when the fraud is revealed for the first time in the three year window, in order to count each fraud case only one time across multiple firm-years. The total samples include all the fraud and non-fraud firms.

2.2. Research model

2.2.1. The effect of restatements on fraud prediction

To test Hypothesis 1, we build the following binary logistic regression.

$$\begin{aligned} \text{Prob}(\text{FRAUD}/\text{FRAUD_NO_REPS}=1) = & F(a_0 + a_1\text{RESTATE} + a_2\text{SIZE} + a_3\text{LRATIO} + \\ & a_4\text{LEVERAGE} + a_5\text{EPS} + a_6\text{ST} + a_7\text{CONCEN} + \\ & a_8\text{INS_HOLDER} + a_9\text{AUD_OPI} + a_{10}\text{BIG4} + a_{11}\text{ROA} + \\ & a_{12}\text{EBIT_ASSET} + a_{13}\text{MA_TURN} + a_{14}\text{SOE} + a_{15}\text{IND_DIRE} + \\ & a_{16}\text{MV} + \text{Industry and Year Dummies}. \end{aligned} \quad (1)$$

In Eq. (1), the dependent variable is the probability of corporate fraud. The primary independent variable is RESTATE, which is equal to 1 if the companies make a restatement. Hypothesis 1 predicts that the restatement firms are positively associated with future fraud revelations. Therefore, if the coefficient on RESTATE (a_1) is positive and significant, Hy-

pothesis 1 is supported by the empirical evidence. In addition, the firm characteristics that influence the relationship between RESTATE and FRAUD in prior research are controlled (Hong, Wang, Ping, & Information, 2014; H.-H. Zhang & L.-Y. Zhang, 2017). The notations provide the definitions of the variables.

In Table 1, the fraud sample is only one tenth of the total sample. Fraud detection is similar to finding a needle in a haystack. Due to the relative rarity of fraudulent firms relative to the non-fraudulent control sample, fraud prediction is more difficult. The propensity score matching method excludes the non-restatement firms that are not matched with restatement firms. Therefore, we use the propensity score matching to empirically test Hypothesis 1. This approach matches the treatment firms (restatement) with the most similar control firms (without restatement) on several dimensions using the estimated likelihood of receiving treatment (Perols, Bowen, Zimmermann, & Samba, 2017; Shipman, Swanquist, & Whited, 2016). Finally, we establish the following logistic regression first, and then use Eq. (1) to predict fraud.

$$\begin{aligned} \text{Prob}(\text{RESTATE}=1) = & F(\beta_0 + \beta_1\text{SIZE} + \beta_2\text{LRATIO} + \beta_3\text{LEVERAGE} + \\ & \beta_4\text{EPS} + \beta_5\text{ST} + \beta_6\text{CONCEN} + \beta_7\text{INS_HOLDER} + \\ & \beta_8\text{AUD_OPI} + \beta_9\text{BIG4} + \beta_{10}\text{ROA} + \beta_{11}\text{EBIT_ASSET} + \\ & \beta_{12}\text{MA_TURN} + \beta_{13}\text{SOE} + \beta_{14}\text{IND_DIRE} + \\ & \beta_{15}\text{MV} + \text{Industry and Year Dummies}). \end{aligned} \quad (2)$$

2.2.2. *The effects of the characteristics of restatement on fraud prediction*

To test Hypothesis 2 to Hypothesis 4, we model the probability of fraud as a function of the characteristics of restatements by estimating the following logistic model. The notations provide the definitions of the variables.

$$\begin{aligned} \text{Prob}(\text{FRAUD}=1) = & F(\gamma_0 + \sum_{i=1}^m \delta_i \text{Variables} + \gamma_1\text{SIZE} + \gamma_2\text{LRATIO} + \\ & \gamma_3\text{LEVERAGE} + \gamma_4\text{EPS} + \gamma_5\text{ST} + \gamma_6\text{CONCEN} + \\ & \gamma_7\text{INS_HOLDER} + \gamma_8\text{AUD_OPI} + \gamma_9\text{BIG4} + \gamma_{10}\text{ROA} + \\ & \gamma_{11}\text{EBIT_ASSET} + \gamma_{12}\text{MA_TURN} + \gamma_{13}\text{SOE} + \gamma_{14}\text{IND_DIRE} + \\ & \gamma_{15}\text{MV} + \text{Industry and Year Dummies}). \end{aligned} \quad (3)$$

We take subsamples of the restatement firms to examine whether firms with different characteristics have different effects on the probability of future fraud. The variables in the model describe the different characteristics of restatements. They are CONTENT (core earnings, noncore earnings, cash flow items, balance sheet items, and innocent errors), SEVERITY (duration, restated numbers and accounts) and REASON (negligence, mistake and unknown). As in Eq.(1), the main control variables are SIZE, LRATIO, LEVERAGE, EPS, ST, CONCEN, INS HOLDER, AUD_OPI, BIG4, ROA, EBIT_ASSET, MA_TURN, SOE, IND_DIRE, and MV. The prior literature has found that the control variables that are included are

consistent with fraud prediction (Wang & Wu, 2011; Xin, Zhou, & Hu, 2018). After controlling the other determinants of fraud, we test how specific characteristics of restatements are correlated with the probability of fraud.

3. Empirical results

3.1. Descriptive statistics of restatements

Figure 1 presents the descriptive statistics for 1890 restatements of 1588 firms. The number of restatements increases dramatically from 241 restatements in 2009 to 746 restatements in 2013. In Figure 2, approximately 10% of the restatements affects one or more financial statement items, and 76% are related to the footnotes of financial restatements where shareholders occupy the highest proportion, and innocent errors account for 14%. These suggest that fewer restatements are quantitatively significant. In Figure 3, half the companies do not provide their restatement reason, one third of companies attribute their restatements to negligence, and only 20% of companies admit their mistakes publicly. Lastly, Figure 4 presents the sample

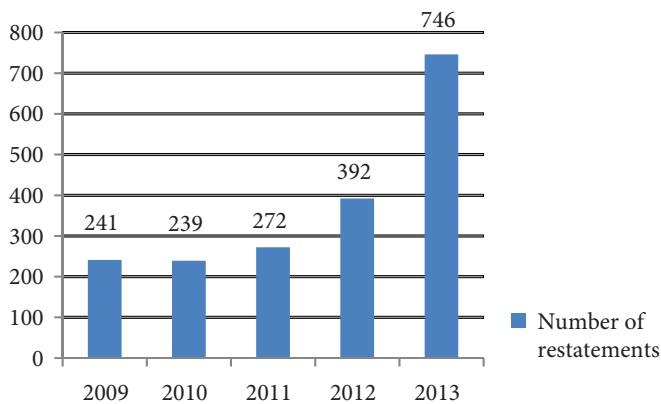


Figure 1. Numbers of restatements from 2009 to 2013

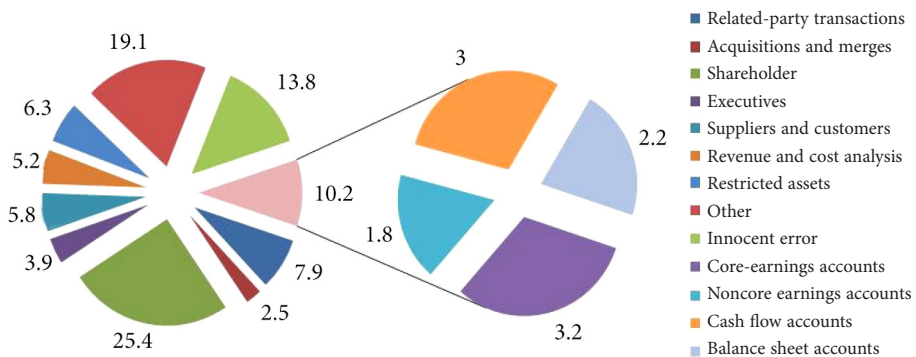


Figure 2. Frequency of types of restatements from 2009 to 2013 (%)

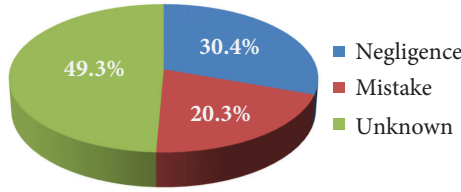


Figure 3. Frequency of reasons for restatements from 2009 to 2013

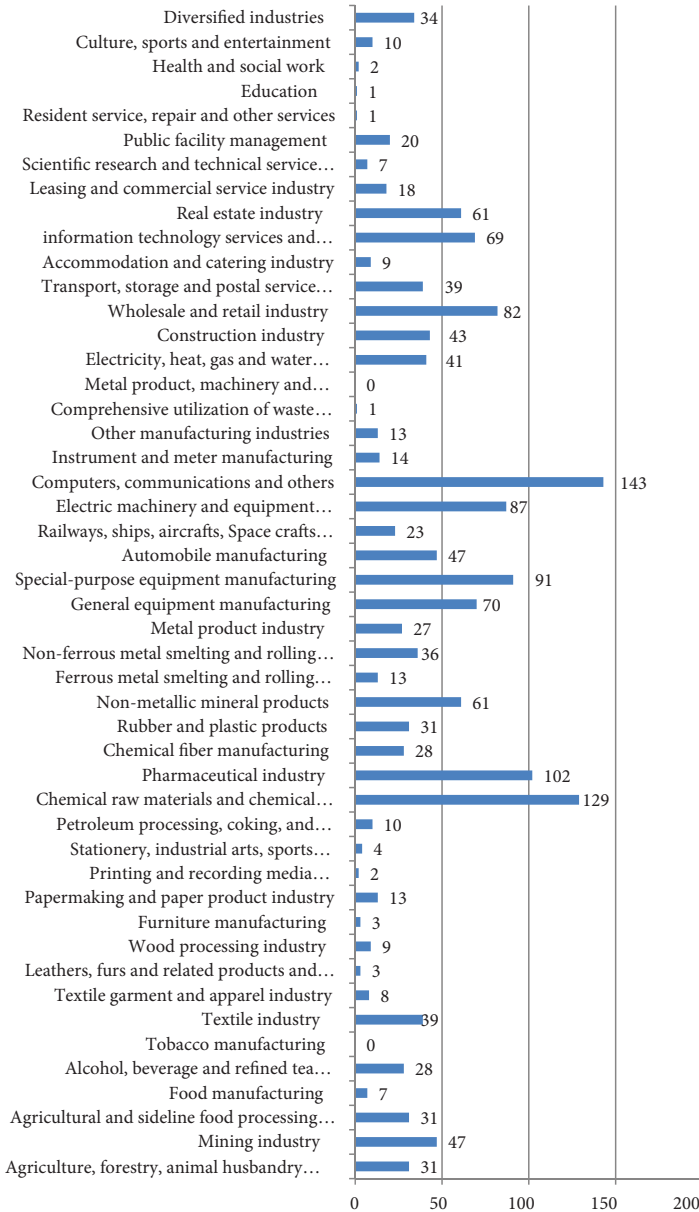


Figure 4. Restatement firms by industries from 2009 to 2013

distribution of the restatements by industry from 2009 to 2013 using Guidelines for the Industry Classification of Listed Companies (2012). Except for “Tobacco manufacturing” and “Metal product, machinery and equipment repair”, the other industries made restatements. Computers, communications and other industries make the most restatements.

3.2. Univariate results and correlation of fraud prediction model

Table 2 provides the results of the univariate analysis between the samples of the Restatement firms (RESTATE = 1) and the non-restatement firms (RESTATE = 0). The evidence indicates that Restatement firms have significantly more FRAUD (mean = 0.171) than non-restatement firms (mean = 0.103). In addition, the mean values of the control variables for non-restatement firms, such as SIZE, EPS, CONCEN, INS_HOLDER, AUD_OPI, BIG4 and MV, are all higher than those for Restatement firms, except for the SOE and ST, which has the opposite effect. Overall, these results suggest that non-restatement firms are healthier and have good corporate governance. They are larger, have higher profits and better audit quality and are also less likely to commit fraud. These univariate findings also provide preliminary evidence in support of Hypothesis 1.

Table 2. Univariate results

Variable	RESTATE=0			RESTATE=1			Mean Difference t-test
	n = 8875			n = 1589			
	Mean	SD	Median	Mean	SD	Median	
Fraud	0.103	0.304	0.000	0.171	0.376	0.000	-7.801***
SIZE	21.790	1.332	21.620	21.649	1.243	21.510	3.916***
LRATIO	3.007	6.123	1.560	2.758	4.051	1.490	1.561
LEVERAGE	0.482	0.883	0.450	0.477	0.743	0.450	0.209
EPS	0.387	0.616	0.300	0.314	0.591	0.240	4.369***
ST	0.027	0.163	0.000	0.036	0.186	0.000	-1.896 ⁺
CONCEN	0.365	0.157	0.350	0.351	0.150	0.330	3.122***
INS_HOLDER	7.403	17.617	0.460	5.021	14.698	0.390	5.081***
AUD_OPI	0.960	0.196	1.000	0.948	0.223	1.000	2.266**
BIG4	0.058	0.233	0.000	0.037	0.189	0.000	3.309***
ROA	0.050	0.261	0.040	0.043	0.283	0.040	1.032
EBIT_ASSET	0.074	1.181	0.060	0.054	0.332	0.050	0.692
MA_TURN	0.298	0.457	0.000	0.303	0.460	0.000	-0.462
SOE	0.665	0.472	1.000	0.739	0.439	1.000	-5.800***
IND_DIRE	0.367	0.056	0.330	0.368	0.057	0.330	-0.430
MV	15.266	0.940	15.110	15.191	0.912	15.000	2.943***

*, **, *** Indicate significance at $p < 0.10$, 0.05 , and 0.01 , respectively.

See Notations for variable definitions.

Table 3 displays the correlations for all variables (See Table 3 in the Appendix). The occurrence of fraud in the subsequent three years is significantly positively related with restatements (FRAUD and RESTATE, p -value <0.05). We also note that Restatement firms has significantly negative correlations with SIZE, EPS, CONCEN, INS_HOLDER, AUD_OPI, BIG4, SOE and MV, suggesting that restatement firms have smaller sizes, lower profits and share concentrations and are less likely to have a Big 4 auditor and a clean audit opinion. This may be viewed as prima facie evidence of the predictive ability of restatements regarding future fraud. Finally, the relatively small correlation coefficients among our control variables indicate that there is little evidence of a multicollinearity problem in our variables.

3.3. Multivariate results – restatement and future fraud revelation

Our first research question pertains to the ability of restatements to predict future fraud. To answer this question, we perform multiple regression tests. Table 4 reports our multinational logistic results of Model (1). The first and second columns use FRAUD as the dependent variable, and the coefficient on RESTATE is positive significantly (p -value less than 1%). The results support that firms with a restatement are prone to commit fraud within the next three years compared to non-restatement firms. Furthermore, the marginal effect from the logistical model (untabulated) predicts that the probability of fraud increases 5.18 percentage points from non-restatement firms to restatement firms¹.

Furthermore, the control variables of SIZE, LRATIO, EPS, CONCEN, AUD_OPI, BIG4, SOE, and MV are all negatively associated with fraud, suggesting that unhealthy firms are more related to the likelihood of fraud (Li & Yan, 2018). The empirical results provide support for Hypothesis 1, which posits that restatements are earlier warnings for future fraud revelation. We also perform a logistical regression to test the contagion effect of industry restatements on fraud (Ji, Kumar, Pei, & Xue, 2019). The results are consistent with our main findings, which suggest that industry restatements increase the firms' future fraud revelation (untabulated). They are also consistent with the view that accounting errors can incentivize fraud by providing camouflage effect on firms (Pfarrer, Smith, Bartol, Khanin, & Zhang, 2008; Fang et al., 2017).

The third and fourth columns display the results when the dependent variable is FRAUD_NO_REPS. There is a positive and statistically significant association between RESTATE and FRAUD_NO_REPS (p -values less than 5% and 1%, respectively), implying that the results are not affected by counting fraud repeatedly.

The results for Eq. (2) are tabulated in Table 5, Panel A. Through the probabilities of Eq. (2), we match the firms with restatements (treatment samples) to the firms without restatements (control samples) with the closest predicted probability of restatement. In Panel B, the mean tests between the treatment and control groups reveal no significant differences for all the control observations, showing that we have retained enough covariate balance. Panel C displays the results of Eq. (1) for the 3026 samples obtained by using the propensity score matching method. In the multiple regression model, RESTAT has a coefficient is 0.5634

¹ We estimate the marginal effect by using the average of the discrete or partial changes over all observations (Bartus, 2005)

Table 3. Variable correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 Fraud	1																
2 RESTATE	0.076**	1															
3 SIZE	-0.122**	-0.038**	1														
4 LRATIO	-0.042**	-0.015	-0.191**	1													
5 LEVERAGE	0.039**	-0.002	-0.016	-0.140**	1												
6 EPS	-0.121**	-0.043**	0.212**	0.090**	-0.106**	1											
7 ST	0.087**	0.019	-0.141**	-0.057**	0.133**	-0.108**	1										
8 CONCEN	-0.093**	-0.031**	0.293**	-0.014	-0.029**	0.1113**	-0.067**	1									
9 INS_HOLDER	-0.028**	-0.050**	0.062**	-0.026**	0.003	0.095**	-0.003	0.056**	1								
10 AUD_OPI	-0.154**	-0.022	0.203**	0.064**	-0.239**	0.178**	-0.288**	0.118**	0.017	1							
11 BIG4	-0.070**	-0.032**	0.382**	-0.065**	0.015	0.080**	-0.029**	0.136**	0.050**	0.027**	1						
12 ROA	-0.040**	-0.010	-0.017	0.038**	-0.114**	0.331**	-0.022*	0.007	0.015	0.016	0.004	1					
13 EBIT_ASSET	-0.014	-0.007	-0.041**	-0.009	-0.027**	0.167**	-0.006	-0.015	0.001	-0.037**	0.000	0.489**	1				
14 MA_TURN	0.028**	0.005	-0.048**	0.062**	0.010	-0.019*	0.061**	0.026**	0.015	-0.063**	-0.030**	-0.014	-0.012	1			
15 SOE	-0.050**	.057**	.223**	-.122**	0.007	-0.041**	-0.025*	0.1113**	-0.093**	0.017	0.066**	-0.032**	0.000	-0.066**	1		
16 IND_DIRE	-.015	.004	.038**	.022*	-0.001	-0.011	0.009	0.064**	-0.017	0.000	0.046**	0.008	0.012	0.009	-0.015	1	
17 MV	-.113**	-.029**	.635**	-.060**	-0.016	0.288**	-0.086**	0.233**	0.202**	0.125**	0.294**	0.049**	0.006	-0.021*	0.090**	0.046**	1

* Coefficients with a two-tailed p-value of 0.10, ** Coefficients with a two-tailed p-value of 0.05, ***Coefficients with a two-tailed p-value of 0.01. See Notations for variable definitions.

(p-value<0.01) across all specifications. These suggest firms with restatements are 5.54 percentage points more likely to have a fraud revelation during the next three years compared with firms without restatements.

Table 4. Multiple regression fraud prediction results

Independent Variables	FRAUD		FRAUD_NO_REPS	
	①	②	③	④
RESTATE	0.5226***	0.5370***	0.2766**	0.4255***
	(0.078)	(0.079)	(0.129)	(0.132)
SIZE	-0.1086***	-0.0905**	-0.2728***	-0.1609**
	(0.036)	(0.038)	(0.057)	(0.064)
LRATIO	-0.0507***	-0.0513***	-0.0358**	-0.0146
	(0.011)	(0.012)	(0.014)	(0.012)
LEVERAGE	-0.0519	-0.0471	-0.0461	-0.0429
	(0.036)	(0.034)	(0.041)	(0.039)
EPS	-0.4860***	-0.4698***	-0.5226***	-0.3792**
	(0.078)	(0.077)	(0.114)	(0.124)
ST	0.2584*	0.2467	0.1172	0.0278
	(0.154)	(0.154)	(0.239)	(0.240)
CONCEN	-0.9310***	-1.040***	-0.5435	-0.9034**
	(0.229)	(0.232)	(0.350)	(0.362)
INS_HOLDER	-0.0020	-0.0030	-0.0110***	-0.0172***
	(0.002)	(0.002)	(0.004)	(0.004)
AUD_OPI	-0.8401***	-0.8163***	-0.4801**	-0.4438**
	(0.128)	(0.128)	(0.207)	(0.211)
BIG4	-1.2457***	-1.241***	-1.2481**	-1.2952**
	(0.300)	(0.301)	(0.513)	(0.516)
ROA	0.0226	0.0228	-0.1233	-0.1743
	(0.167)	(0.171)	(0.293)	(0.369)
EBIT_ASSET	-0.0128	-0.0143	0.0107	-0.0068
	(0.079)	(0.084)	(0.074)	(0.234)
MA_TURN	0.1058	0.1054	0.4279***	0.4123***
	(0.069)	(0.069)	(0.101)	(0.103)
SOE	-0.3108***	-0.2511***	-0.8214**	-0.3197***
	(0.069)	(0.075)	(0.102)	(0.109)
IND_DIRE	-0.6397	-0.5950	-2.0335**	-0.9209
	(0.607)	(0.613)	(0.943)	(0.951)
MV	-0.1522***	-0.1664***	0.1107	-0.0677
	(0.051)	(0.052)	(0.077)	(0.090)

End of Table 4

Independent Variables	FRAUD		FRAUD_NO_REPS	
	①	②	③	④
Year fixed effects?	No	Yes	No	Yes
Industry fixed effects?	No	Yes	No	Yes
Obs. where Fraud=1	1188	1188	458	458
NO. of observations	10464	10464	9734	9734
Area under the ROC Curve	0.689	0.698	0.713	0.767

* Coefficients with a two-tailed p-value of 0.10, ** Coefficients with a two-tailed p-value of 0.05, ***Coefficients with a two-tailed p-value of 0.01.

Notations provide variable definitions.

Table 5. Propensity score matched results
Panel A: First-stage Regression Model

Dependent Variable is RESTATE	
Independent Variables	①
SIZE	-0.0938*** (0.030)
LRATIO	-0.0106* (0.006)
LEVERAGE	-0.0502 (0.051)
EPS	-0.1346** (0.059)
ST	0.1237 (0.161)
CONCEN	-0.3816** (0.189)
INS_HOLDER	-0.0079*** (0.002)
AUD_OPI	-0.0689 (0.142)
BIG4	-0.2931* (0.150)

Dependent Variable is RESTATE	
Independent Variables	①
ROA	0.1567 (0.338)
EBIT_ASSET	-0.1295 (0.297)
MA_TURN	0.0423 (0.060)
SOE	0.3824*** (0.064)
IND_DIRE	0.3835 (0.500)
MV	0.0571 (0.040)
Obs. where RESTATE=1	1588
NO. of observations	10464
Area under the ROC Curve	0.757

Panel B: Covariate Balance

Variable	RESTATE=1(Treatment)		RESTATE=0(Control)		Mean Difference t-test
	Group		Group		
	Mean	SD	Mean	SD	
SIZE	21.648	1.242	21.663	1.215	0.319
LRATIO	2.761	4.052	2.895	4.458	0.869
LEVERAGE	0.477	0.744	0.456	0.347	-0.986
EPS	0.316	0.588	0.318	0.474	0.109
ST	0.036	0.186	0.032	0.176	-0.598
CONCEN	0.351	0.151	0.351	0.149	-0.062
INS_HOLDER	5.027	14.706	5.706	15.396	1.240
AUD_OPI	0.948	0.221	0.958	0.202	1.201
BIG4	0.037	0.189	0.042	0.200	0.638
ROA	0.043	0.282	0.041	0.074	-0.267
EBIT_ASSET	0.055	0.325	0.053	0.087	-0.259
MA_TURN	0.303	0.460	0.297	0.457	-0.339
SOE	0.739	0.440	0.741	0.438	0.143
IND_DIRE	0.370	0.056	0.370	0.055	-0.024
MV	15.189	0.911	15.192	0.913	0.095

Panel C: Second stage regression models

Dependent Variable is FRAUD			
Independent Variables		①	P-value
RESTATE	+	0.5634	0.000***
		(0.116)	
SIZE	-	-0.0958	0.162
		(0.069)	
LRATIO	-	-0.0658	0.009***
		(0.025)	
LEVERAGE	+	0.2956	0.197
		(0.229)	
EPS	-	-0.4406	0.005***
		(0.156)	
ST	+	0.1340	0.614
		(0.265)	
CONCEN	-	-0.7958	0.052*
		(0.410)	
INS_HOLDER	-	-0.0047	0.307
		(0.005)	

End of Panel C

Dependent Variable is FRAUD			
Independent Variables		①	P-value
AUD_OPI	–	–0.8442	0.000***
		(0.221)	
BIG4	–	–1.3903	0.021**
		(0.603)	
ROA	+	0.7184	0.664
		(1.656)	
EBIT_ASSET	–	–1.2677	0.362
		(1.391)	
MA_TURN	+	0.2420	0.039**
		(0.117)	
SOE	–	–0.2086	0.133
		(0.139)	
IND_DIRE	–	0.2857	0.784
		(1.044)	
MV	–	–0.2302	0.012**
		(0.092)	
Year fixed effects?		Yes	
Industry fixed effects?		Yes	
Obs. where Fraud=1		424	
NO. of observations		3026	
Area under the ROC Curve		0.718	

* Coefficients with a two-tailed p-value of 0.10, ** Coefficients with a two-tailed p-value of 0.05, *** Coefficients with a two-tailed p-value of 0.01.

Notations provide variable definitions.

3.4. Multivariate results – characteristics of restatement and future fraud revelation

This part examines the association between the characteristics of restatements and future fraud revelations. The results for Model (3) are tabulated in Table 6. Consistent with Hypothesis 1, a significant positive association exists between the characteristics of restatements and future fraud revelations. Table 6, Panel A1 provides partial support for Hypothesis 2. In support of Hypothesis 2a, except for balance sheet items, corrections of core earnings, noncore earnings and cash flow items are all associated with a significant increase in the likelihood of firms' future fraud (p-values of <0.01, 0.05 and 0.01, respectively). These results indicate that investors shed more light on firms' profits and cash flows than their balance sheet items such that listed companies are more inclined to falsify their profit and cash flow data. This is also consistent with the findings in the existing fraud

researches (Lee et al., 1999; Firth et al., 2011). Likewise, firms with only innocent errors are less likely to engage in fraud, thus providing support for Hypothesis 2b. The finding is also consistent with that of Tan and Young (2015), who suggested that firms with little errors have better financial performance and are less likely to have a material weakness. We also test the relationship between the content of footnote restatements and future fraud revelations in Panel A2, Table 6. Certain types of footnotes could also be indicative of firms' future fraud. The related-party transactions and acquisitions and mergers are positively related to the probability of future fraud ($p\text{-value} < 0.05$). The results are consistent with the existing studies (Erickson, Hanlon, Maydew, 2006; Wang, 2011; Fang, Lobo, Zhang, and Zhao, 2018). They found that acquisitions and mergers are associated with a higher likelihood of fraud, and are often regarded as red flags of fraud. Our evidence also supports Hypothesis 2c, implying that restatements related to footnotes also provide information on detecting fraud. Overall, these results suggest that the content of restatements affect future fraud revelations.

Table 6, Panel B shows the results of an association test between the severity of restatement and fraud revelations. The revision number, duration and revision accounts are also positively associated with the probability of fraud at least 95% significance level. These results clearly support Hypothesis 3 that companies with high severity of restatements have a higher probability of fraud. Contrary to Srinivasan et al. (2014), our results show that the high frequency of restatements is a signal of future fraud. Our results are also confirmed by Badertscher and Burks (2011), who found that lengthy restated-period are concentrated in restatements involving fraud. Finally, Hypothesis 4 is totally supported. Panel C reports the regression results. As expected, the coefficient on NEGLIGENCE is significantly positive at the 0.01 level, suggesting that companies that attribute restatements to their negligence are more likely to have a revelation of fraud. This is consistent with the theory developed by Corona and Randhawa (2018), who maintained that low-quality agents have an incentive to hide their mistakes.

Table 6. Characteristics of restatement and future fraud prediction
Panel A1: Content of financial statement restatements to fraud prediction

Independent Variables	Dependent Variable is FRAUD	
	(1)	(2)
CORE-EARNING		1.0972 ^{***}
		(0.308)
NONCORE EARNING		0.8961 ^{**}
		(0.412)
CASH FLOW ITEMS		0.9249 ^{***}
		(0.335)
BALANCE SHEET ITEMS		-0.3615
		(0.520)

End of Panel A1

Dependent Variable is FRAUD		
Independent Variables	(1)	(2)
INNOCENT ERRORS	-0.4101*	
	(0.217)	
SIZE	0.0102	0.0090
	(0.086)	(0.087)
LRATIO	-0.0379	-0.0417
	(0.025)	(0.026)
LEVERAGE	0.0130	-0.0155
	(0.150)	(0.156)
EPS	-0.5852***	-0.5900***
	(0.177)	(0.176)
ST	0.2796	0.3256
	(0.342)	(0.343)
CONCEN	-0.7973	-0.7503
	(0.532)	(0.539)
INS_HOLDER	-0.0100	-0.0101
	(0.009)	(0.009)
AUD_OPI	-1.2270***	-1.1566***
	(0.281)	(0.284)
BIG4	-2.1402**	-2.0401**
	(1.034)	(1.033)
EBIT_ASSET	-0.0423	-0.0369
	(0.270)	(0.285)
MA_TURN	0.2163	0.2546*
	(0.153)	(0.154)
SOE	-0.1966	-0.1856
	(0.197)	(0.200)
IND_DIRE	-0.5654	-0.6975
	(1.350)	(1.366)
MV	-0.3705***	-0.3442***
	(0.121)	(0.121)
Year fixed effects?	Yes	Yes
Industry fixed effects?	Yes	Yes
Obs. where Fraud=1	271	271
NO. of observations	1588	1588
Area under the ROC Curve	0.732	0.744

Panel A2: Content of Footnote Restatements to Fraud prediction

Dependent Variable: FRAUD		Independent Variables									
Related-party transaction		Acquisition and merger	Shareholder	Executives	Suppliers and customers	Revenues and Cost analysis	Restricted asset	SIZE	LRATIO	LEVE-RAGE	EPS
0.5496**	(0.241)	0.7304**	-0.1291	-0.0130	-0.1062	-0.0611	-0.2028	0.0046	-0.0341	0.0061	-0.5804***
		(0.358)	(0.172)	(0.422)	(0.308)	(0.319)	(0.309)	(0.086)	(0.025)	(0.147)	(0.179)
Dependent Variable: FRAUD		Independent Variables									
ST		CONCEN	INS_HOLDER	AUD_OPI	BIG4	EBIT_ASSET	MA_TURN	SOE	IND_DIRE	MV	
0.3164	(0.342)	-0.8181	-0.0098	-1.2437***	-2.0649**	-0.0736	0.2262	-0.2126	-0.4263	-0.3762***	
		(0.535)	(0.009)	(0.282)	(1.034)	(0.309)	(0.154)	(0.198)	(1.359)	(0.120)	

Panel B: Severity of Restatement to Fraud Prediction

Dependent Variable is FRAUD			
Independent Variables	(1)	(2)	(3)
REVISION NUMBER	0.4285***		
	(0.158)		
DURATION		0.4451**	
		(0.213)	
REVISION ACCOUNTS			0.4026***
			(0.087)
SIZE	-0.0010	0.0077	0.0230
	(0.086)	(0.086)	(0.087)
LRATIO	-0.0394	-0.0381	-0.0401
	(0.026)	(0.025)	(0.026)
LEVERAGE	0.0218	0.0133	-0.0218
	(0.146)	(0.150)	(0.180)
EPS	-0.5622***	-0.5587***	-0.5834***
	(0.175)	(0.176)	(0.176)
ST	0.2361	0.2747	0.2983
	(0.342)	(0.341)	(0.343)
CONCEN	-0.8117	-0.8056	-0.7517
	(0.532)	(0.531)	(0.535)
INS_HOLDER	-0.0097	-0.0098	-0.0106
	(0.009)	(0.009)	(0.009)
AUD_OPI	-1.1683***	-1.2086***	-1.1985***
	(0.282)	(0.281)	(0.283)
BIG4	-2.0834**	-2.1205**	-2.0415**
	(1.033)	(1.034)	(1.034)
EBIT_ASSET	-0.0268	-0.0480	-0.0560
	(0.269)	(0.270)	(0.297)
MA_TURN	0.2272	0.2246	0.2408
	(0.153)	(0.153)	(0.154)
SOE	-0.1968	-0.2015	-0.2267
	(0.197)	(0.196)	(0.198)
IND_DIRE	-0.6445	-0.5388	-0.6361
	(1.352)	(1.343)	(1.359)
MV	-0.3415***	-0.3597***	-0.3683***
	(0.121)	(0.120)	(0.122)
Year fixed effects?	Yes	Yes	Yes

End of Panel B

Dependent Variable is FRAUD			
Independent Variables	(1)	(2)	(3)
Industry fixed effects?	Yes	Yes	Yes
Obs. where Fraud=1	271	271	271
NO. of observations	1588	1588	1588
Area under the ROC Curve	0.734	0.732	0.742

Panel C: Reasons of Restatement to Fraud Prediction
Dependent Variable is FRAUD

Independent Variables	(1)	(2)
NEGLIGENCE	0.4402***	
	(0.169)	
MISTAKE	0.2416	
	(0.198)	
UNKNOWN		-0.3661**
		(0.154)
SIZE	0.0024	0.0069
	(0.087)	(0.086)
LRATIO	-0.0406	-0.0401
	(0.026)	(0.026)
LEVERAGE	0.0185	0.0184
	(0.170)	(0.164)
EPS	-0.6144***	-0.6121***
	(0.177)	(0.177)
ST	0.3160	0.3075
	(0.344)	(0.344)
CONCEN	-0.8076	-0.8323
	(0.534)	(0.533)
INS_HOLDER	-0.0094	-0.0094
	(0.009)	(0.009)
AUD_OPI	-1.2164***	-1.2079***
	(0.282)	(0.282)
BIG4	-2.1332**	-2.1316**
	(1.034)	(1.033)
EBIT_ASSET	-0.0084	-0.0078
	(0.266)	(0.265)

End of Panel C

Independent Variables	(1)	(2)
MA_TURN	0.2439 (0.154)	0.2407 (0.154)
SOE	-0.2084 (0.197)	-0.2007 (0.197)
IND_DIRE	-0.4212 (1.353)	-0.4559 (1.350)
MV	-0.3360*** (0.121)	-0.3443*** (0.121)
Year fixed effects?	Yes	Yes
Industry fixed effects?	Yes	Yes
Obs. where Fraud=1	271	271
NO. of observations	1588	1588
Area under the ROC Curve	0.739	0.737

* Coefficients with a two-tailed p-value of 0.10, ** Coefficients with a two-tailed p-value of 0.05, *** Coefficients with a two-tailed p-value of 0.01.

See Notations and Appendix A for variable definitions.

Conclusions

This study reveals a significant relationship between financial report restatements and future fraud revelations. Firms with a restatement are more likely to commit fraud in subsequent periods. It is found in this paper that the content, severity and reasons for restatements provide warning signs for regulators and investors to suspect fraud. The results in this paper highlight the important economic implications of restatements. Based on the analysis of the restatements, on the one hand, regulators and investors should be alert to profit and cash flow restatements; restatements about related-party transactions, acquisitions and mergers; high severity of restatements and restatements attributed to negligence. In addition, the findings support that the technical error restatements have a negative correlation with future fraud revelations. This finding can help to identify firms with low fraud risks. On the other hand, academics should be aware of the potential implications of restatements for firms' tendency to commit fraud.

This study also calls attention to the disclosure of more detailed financial restatements. Chinese firms' restatements provided few details about their revisions, and even have presented inaccurate or incomplete explanations in regard to the underlying causes of the restatements. Regulators should establish some incentives and rules to prompt companies to provide clear and transparent disclosures, which will help quell investors' concerns. The studies provide a valuable insight about the value of restatements for fraud prediction.

This research is subject to some limitations. First, the characteristics of the restatements were collected and analyzed manually, thus there may exist inconsistency due to the devia-

tion in understanding. Second, our fraud samples depend on the fraud revelation by the regulators. As a result, the undiscovered fraud firms may influence the fraud detection of the restatements. This is due to the fact that the managerial intent for the restatements is hardly to be measured. So, future studies could aim to develop advanced tools to explore more characteristics of the restatements, and examine whether these characteristics are related with future fraud revelation. The analysis of the linguistic feature of the restatements could help to understand the intent of corporate restatements more deeply in the future work.

Funding

Please supply all details required by any funding and grant-awarding bodies as Funding in a separate paragraph as follows:

For multiple agency grants

This work was supported by the <National Natural Science Foundation of China> under Grant [number 61303043]; < Scientific Research Project of Education Department of Hunan Provincial > under Grant [number18C0761,17C0147]; < Changsha science Project> under Grant [number K1705091]; < Hunan Provincial Education Science 12th Five-year Planning Subject > under Grant [number XJK015QJG001]; <Natural Science Foundation of Hunan Province > under Grant [number2016JJ3017]; <Key projects of scientific research of Education Department of Hunan Provincial > under Grant [number15A022]; <Social Science Foundation of Hunan Province > under Grant [number16YBQ002]; < Project of the Achievement Evaluation Committee of Social Sciences in Hunan Province > under Grant [numberXSP18Y-BZ019].

Author contributions

Shi Qiu designed the research model, carried out the data regression analysis and wrote the draft of this paper. Hong-qu He was responsible for supervising the analysis work and revising the paper. Yuan-sheng Luo gave suggestions on related work and model.

Disclosure statement

We declare that we have no competing financial, professional, or personal interests from other parties.

References

- ACFE. (2016). *ACFE report to the nations on occupational fraud and abuse*. Technical Report, Association of Certified Fraud Examiners, Austin, TX. Retrieved from <https://www.acfe.com/rtn2016/docs/2016-report-to-the-nations.pdf>
- Ahmed, K., & Goodwin, J. (2007). An empirical investigation of earnings restatements by Australian firms. *Accounting and Finance*, 47(1), 1-22. <https://doi.org/10.1111/j.1467-629X.2006.00196.x>

- Badertscher, B. A., & Burks, J. J. (2011). Accounting restatements and the timeliness of disclosures. *Accounting Horizons*, 25(4), 609-629. <https://doi.org/10.2308/acch-50026>
- Bartus, T. (2005). Estimation of marginal effects using margeff. *Stata Journal*, 5(3), 309-329. <https://doi.org/10.1177/1536867X0500500303>
- Baucus, M. (1994). Pressure, opportunity and predisposition: A multivariate model of corporate illegality. *Journal of Management*, 20(4), 699-721. [https://doi.org/10.1016/0149-2063\(94\)90026-4](https://doi.org/10.1016/0149-2063(94)90026-4)
- BenYoussef, N., & Breton, G. (2016). *Restatement announcements: sincerity analysis through information manipulation theory*. Retrieved from <https://ssrn.com/abstract=1985680>
- BenYoussef, N., & Khan, S. (2018). Timing of earnings restatements: CEO equity compensation and market reaction. *Accounting and Finance*, 58(2), 341-365. <https://doi.org/10.1111/acfi.12220>
- China securities regulatory commission. (2012). *Guidelines for the industry classification of listed companies*. Retrieved from http://www.csrc.gov.cn/pub/newsite/flb/flfg/bmgf/zh/gfxwj/tj/201310/20131016_236281.html
- Christensen, J. (2010). Accounting errors and errors of accounting. *Accounting Review*, 85(6), 1827-1838. <https://doi.org/10.2308/accr.2010.85.6.1827>
- Chung, J., & Mccracken, S. (2014). Understanding the restatement process. *Accounting Perspectives*, 13(4), 253-281. <https://doi.org/10.1111/1911-3838.12036>
- Corona, C., & Randhawa, R. S. (2018). The value of confession: admitting mistakes to build reputation. *The Accounting Review*, 93(3), 133-161. <https://doi.org/10.2308/accr-51893>
- Cuadrado-Ballesteros, B., Garcia-Sanchez, I. M., & Ferrero, J. M. (2016). How are corporate disclosures related to the cost of capital? The fundamental role of information asymmetry. *Management Decision*, 54(7), 1669-1701. <https://doi.org/10.1108/MD-10-2015-0454>
- Demirkan, S., & Platt, H. D. (2018). Differential investors' response to restatement announcements: an empirical investigation. *Journal of Economics and Financial Analysis*, 2(2), 29-59.
- Donelson, D. C., Ege, M. S., & McInnis, J. M. (2017). Internal control weaknesses and financial reporting fraud. *Auditing: A Journal of Practice & Theory*, 36(3), 45-69. <https://doi.org/10.2308/ajpt-51608>
- Dorminey, J., Fleming, A. S., Kranacher, M. J., & Riley Jr, R. A. (2012). The evolution of fraud theory. *Issues in Accounting Education*, 27(2), 555-579. <https://doi.org/10.2308/iace-50131>
- Erickson, M., Hanlon, M., & Maydew, E. L. (2006). Is there a link between executive equity incentives and accounting fraud? *Journal of Accounting Research*, 44(1), 113-143. <https://doi.org/10.1111/j.1475-679X.2006.00194.x>
- Fang, J., Lobo, G. J., Zhang, Y., & Zhao, Y. (2018). Auditing related party transactions: evidence from audit opinions and restatements. *Auditing: A Journal of Practice & Theory*, 37(2), 73-106. <https://doi.org/10.2308/ajpt-51768>
- Fang, V. W., Huang, A. H., & Wang, W. (2017). Imperfect accounting and reporting bias. *Journal of Accounting Research*, 55(4), 919-962. <https://doi.org/10.1111/1475-679X.12170>
- Firth, M., Rui, O. M., & Wu, W. (2011). Cooking the books: recipes and costs of falsified financial statements in China. *Journal of Corporate Finance*, 17(2), 371-390. <https://doi.org/10.1016/j.jcorpfin.2010.09.002>
- He, L., Sarath, B., & Wans, N. (2019). Material weakness disclosures and restatement announcements: The joint and order effects. *Journal of Business Finance & Accounting*, 46(1-2), 68-104. <https://doi.org/10.1111/jbfa.12363>
- Hennes, K. M., Leone, A. J., & Miller, B. P. (2008). The importance of distinguishing errors from irregularities in restatement research: the case of restatements and CEO/CFO turnover. *The Accounting Review*, 83(6), 1487-1519. <https://doi.org/10.2308/accr.2008.83.6.1487>
- Hirschey, M., Smith, K. R., & Wilson, W. M. (2015). The timeliness of restatement disclosures and financial reporting credibility. *Journal of Business Finance and Accounting*, 42(7-8), 826-859. <https://doi.org/10.1111/jbfa.12125>

- Hong, W. Z., Wang, X. X., Ping, H. Q., & Information, S. O. (2014). Research on the financial report fraud detection of listed companies based on logistic regression model. *Chinese Journal of Management Science*, 22(11), 351-356.
- Ji, A. E., Kumar, K. R., Pei, H., & Xue, Y. (2019). Does the market value auditors' industry specializations? Evidence from the contagion effects of restatements. *Accounting Horizons*, 33(1), 125-152. <https://doi.org/10.2308/acch-52350>
- Jiang, Y. B., Cui, G., & Wang, Y. W. (2010). The trend and characteristics of Financial Statement Restatements: 2004–2008[J]. *Journal of Accounting and Economics*, 24(2), 17-27.
- Kim, Y. J., Baik, B., & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems with Applications*, 62, 32-43. <https://doi.org/10.1016/j.eswa.2016.06.016>
- Land, J. K. (2010). CEO turnover around earnings restatements and fraud. *Pacific Accounting Review*, 22(3), 180-198. <https://doi.org/10.1108/01140581011091666>
- Lee, T. A., Ingram, R. W., & Howard, T. P. (1999). The difference between earnings and operating cash flow as an indicator of financial reporting fraud. *Contemporary Accounting Research*, 16(4), 749-786. <https://doi.org/10.1111/j.1911-3846.1999.tb00603.x>
- Li, Q., & Yan, S. G. (2018). Comparative study of logistic regression accounting fraud identification model and accounting fraud index. *Journal of Northeast Normal University (Philosophy and Social Sciences)*, 294(4), 181-187.
- Liou, F. (2008). Fraudulent financial reporting detection and business failure prediction models: a comparison. *Managerial Auditing Journal*, 23(7), 650-662. <https://doi.org/10.1108/02686900810890625>
- Ma, C., Du, H., & Zhang, J. (2018). Chinese accounting restatement and the timeliness of annual report. *Applied Economics*, 50(50), 5436-5453. <https://doi.org/10.1080/00036846.2018.1486995>
- Murphy, P. R., & Dacin, M. T. (2011). Psychological pathways to fraud: understanding and preventing fraud in organizations. *Journal of Business Ethics*, 101(4), 601-618. <https://doi.org/10.1007/s10551-011-0741-0>
- Office USGA. (2003). *Financial statement restatements: trends, market impacts, regulatory responses, and remaining challenges*. Government Accountability Office Reports. Retrieved from <https://www.gao.gov/assets/240/236067.pdf>
- Palmrose, Z. V., Richardson, V. J., & Scholz, S. (2004). Determinants of market reactions to restatement announcements. *Journal of Accounting & Economics*, 37(1), 59-89. <https://doi.org/10.1016/j.jacceco.2003.06.003>
- Perols, J., Bowen, R., Zimmermann, C., & Samba, B. (2017). Finding needles in a haystack: using data analytics to improve fraud prediction. *The Accounting Review*, 92(2), 221-245. <https://doi.org/10.2308/accr-51562>
- Pfarrer, M. D., Smith, K. G., Bartol, K. M., Khanin, D. M., & Zhang, X. (2008). Coming forward: the effects of social and regulatory forces on the voluntary restatement of earnings subsequent to wrongdoing. *Organization Science*, 19(3), 386-403. <https://doi.org/10.1287/orsc.1070.0323>
- Plumlee, M., & Yohn, T. L. (2015). An examination of managements' regulatory filing choices surrounding restatements. *Journal of Management Accounting Research*, 27(2), 121-144. <https://doi.org/10.2308/jmar-50744>
- Schmidt, J. J., & Wilkins, M. S. (2013). Bringing darkness to light: the influence of auditor quality and audit committee expertise on the timeliness of financial statement restatement disclosures. *Auditing: A Journal of Practice & Theory*, 32(1), 221-244. <https://doi.org/10.2308/ajpt-50307>
- Scholz, S. (2013). Financial restatement trends in the United States: 2003–2013. Washington, DC: Center for Audit Quality. Retrieved from <https://www.theqaq.org/financial-restatement-trends-united-states-2003-2012>

- Securities and Exchange Commission (SEC). (2014). *Implementing Dodd-Frank Wall Street reform and consumer protection act – pending action*. Washington DC: SEC. Retrieved from <http://www.sec.gov/spotlight/dodd-frank/dfactivity-upcoming.shtml>
- Shipman, J. E., Swanquist, Q. T., & Whited, R. L. (2016). Propensity score matching in accounting research. *The Accounting Review*, 92(1), 213-224. <https://doi.org/10.2308/accr-51449>
- Singer, Z., & Zhang, J. (2018). Auditor tenure and the timeliness of misstatement discovery. *The Accounting Review*, 93(2), 315-338. <https://doi.org/10.2308/accr-51871>
- Srinivasan, S., Wahid, A. S., & Yu, G. (2014). Admitting mistakes: home country effect on the reliability of restatement reporting. *The Accounting Review*, 90(3), 1201-1240. <https://doi.org/10.2308/accr-50887>
- Tan, C. E., & Young, S. M. (2015). An analysis of “Little r” restatements. *Accounting Horizons*, 29(3), 667-693. <https://doi.org/10.2308/acch-51104>
- Trompeter, G. M., Carpenter, T. D., Desai, N., Jones, K. L., & Riley, R. A. (2013). A synthesis of fraud-related research. *Auditing: A Journal of Practice & Theory*, 32(Supplement 1), 287-321. <https://doi.org/10.2308/ajpt-50360>
- Wang, T. Y. (2011). Corporate securities fraud: Insights from a new empirical framework. *The Journal of Law, Economics & Organization*, 29(3), 535-568. <https://doi.org/10.1093/jleo/ewr009>
- Wei, Z. H., Li, C. Q., & Chen, T. Y. (2010). Determinants of restatement announcements: evidence from China. *Journal of Business Economics*, 22(4), 75-82.
- Wang, X., & Wu, M. (2011). The quality of financial reporting in China: An examination from an accounting restatement perspective. *China Journal of Accounting Research*, 4(4), 167-196. <https://doi.org/10.1016/j.cjar.2011.09.001>
- Xin, Q., Zhou, J., & Hu, F. (2018). The economic consequences of financial fraud: evidence from the product market in China. *China Journal of Accounting Studies*, 6(1), 1-23. <https://doi.org/10.1080/21697213.2018.1480005>
- Zhang, H., Huang, H. J., & Habib, A. (2018). The effect of tournament incentives on financial restatements: evidence from China. *The International Journal of Accounting*, 53(2), 118-135. <https://doi.org/10.1016/j.intacc.2018.05.002>
- Zhang, H. H., & Zhang, L. Y. (2017). Managerial tournament incentive and financial fraud: evidence from Chinese listed companies. *Business Management Journal*, 39(4), 176-194.

APPENDIX

Appendix A. Categories of restatement content

Category	Description
Core earnings	This category includes misstatements regarding revenues; the costs of goods sold; selling, general and administration fees; business taxes and surcharges. The core earnings are pretax.
Noncore earnings	This category involves a series of special, one-time transactions such as asset impairments, investments, fund-raising, and the disposal of assets, extraordinary items and discontinued operations. The noncore earnings are subject to income taxes.
Cash flow items	Restatements are related to the improper reclassification of items on cash flow statements.
Balance sheet items	These are restatements due to improperly classified accounting items on balance sheets.
Footnotes	These are restatements due to errors in the footnotes of financial statements.
Innocent errors	These are restatements due to calculation mistakes.

Notations

Variables and functions

Fraud – Indicator variable equal to 1 if fraud is announced within three years of the filing date of firms' restated reports for year t and 0 otherwise. Fraud includes false records, significant omissions, disclosure delays, improper accounting treatments, misleading statements, and asset or profit misstatements.

Fraud_no_reps – An indicator variable equal to 1 if fraud is announced within three years of the filing date of firms' restated reports for year t and 0 otherwise. For firm-years from fraud firms, only the first firm-year within the three-year window is kept. All firm-years other than the first firm-year that are within the three-year window are excluded.

RESTATE – An indicator variable equal to 1 if the company issues a restatement of their financial reports in year t .

SIZE – Natural logarithm of the total assets of the firm.

LRATIO – The ratio of liquid assets to total assets.

LEVERAGE – The ratio of total debt to total assets.

EPS – The earnings per share.

ST – An indicator variable equal to 1 if a firm is classed as an "ST" firm by stock exchange.

CONCEN – Holding percentage of the largest shareholder.

INS_HOLDER – The percentage of institutional shareholders with respect to all shareholders.

AUD_OPI – An indicator variable equal to 1 if the company receives a standard audit opinion.

BIG4 – An indicator variable is equal to 1 if the company's auditor is one of the largest four accounting firms. The largest four accounting firms include Ernst & Young, KPMG, Deloitte and PricewaterhouseCoopers.

ROA – The ratio of net profits to total assets.

EBIT_ASSET – The ratio of earnings before interest and taxes to total assets.

MA_TURN – An indicator variable is equal to 1 if the company's manager has changed.

SOE – An indicator variable equal to 1 if the largest shareholder is the government.

IND_DIRE – The percentage of independent directors on the board.

MV – The natural logarithm of the market value of the firm in year t .

Duration – The sum of the restated periods, where a fiscal year = 1 and each additional quarter = 0.25.

Revision Number – The number of revisions or supplements are in one year.

Revision Accounts – The number of account groups involved in the restatement. The account groups are revenues, the costs of sales, operating expenses, one-time items, non-operating expenses and other.