SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Sustainable Economies 2024, 2(4), 301.
https://doi.org/10.62617/se.v2i4.301

Article

Detecting financial statements fraud: Evidence from listed companies in China

Yanmei Duan[1], **Guangshun Qiao**[2,*]

[1] School of Management, Wenzhou Business College, Wenzhou 325035, China
[2] School of Finance and Trade, Wenzhou Business College, Wenzhou 325035, China
*** Corresponding author:** Guangshun Qiao, qiaoguangshun@wzbc.edu.cn

CITATION

Duan Y, Qiao G. Detecting financial statements fraud: Evidence from listed companies in China. 2024; 2(4): 301.
https://doi.org/10.62617/se.v2i4.301

ARTICLE INFO

Received: 16 August 2024
Accepted: 19 September 2024
Available online: 4 November 2024

COPYRIGHT

Copyright © 2024 Author(s). *Sustainable Economies* is published by Sin-Chn Scientific Press Pte. Ltd (Singapore). This work is licensed under the Creative Commons Attribution (CC BY) license.
https://creativecommons.org/licenses/by/4.0/

Abstract: Financial statement fraud is the deliberate misrepresentation of a company's financial statements. Financial statement fraud has been a global concern since it not only harms the investors and creditors but also undermine the public confidence of the capital market. Based on the fact that a common incentive for companies to manipulate financial statement is a decline in the company's financial prospects, this paper applies the mixture hazard early warning model to identify the key impacting financial characteristics in detecting the financial statement fraud for listed companies in China. We find that in the construction industry the warning sign of suspcious business practices is a falling return on assets, while in the real estate industry the financial red flag is an increase in the inventory level. The estimation results indicate that the financial characteristics may have different implications in different industries in detecting financial statement fraud. This research has shed light on setting specific financial characteristics for fraud monitoring and detecting by the regulators.

Keywords: financial statement fraud; listed companies in china; mixture hazard model

1. Introduction

Financial fraud has been a global concern for the public, the press, and regulators because of its rapidly increasing adverse impacts not only on individual investors but the overall economic stability. The Federal Trade Commission reported that the U.S. consumers lost more than US$5.8 billion to fraud in 2021, an increase of more than 70% over 2020[1]. Among various types of financial frauds and scams, including Ponzi-schemes, credit card fraud, ecommerce fraud, imposter scam, and phishing, financial statement fraud in particular has received considerable attention from many organizations across industries and countries. Financial statement fraud directly harms the investors and creditors of the issuer of a fraudulent financial report, bringing related parties lost all or part of the investments if such fraud results in a bankruptcy or near failure. The Association of Certified Fraud Examiners (ACFE) reported that financial statement fraud had been the most costly form of occupational fraud, causing a median loss of around US$1 million [1]. Financial statement fraud may also impact heavily on the public trust in the integrity of the financial reporting mechanism, threaten to undermine the confidence of the capital market, and even lead up to massive devastation to the whole financial system. For example, the notorious Enron debacle[2] and Tyco International scandal[3], among others, significantly dampened the confidence of the investors and resulted in large-scale reputational damage in financial markets. In severe cases, such as the mortgage fraud of Fannie Mae and Freddie Mac[4] and the collapse of the Lehman Brothers[5], securities fraud affected not only the U.S. stock market but triggered the global financial crisis in 2008.

The American Institute of Certified Public Accountants concisely defines the financial statement fraud as the intentional misstatements or omissions in financial statements. Similarly, the ACFE defines fraud, in general, as any activity that relies on deception in order to achieve a gain, and financial statement fraud, in particular, as the deliberate misrepresentation of the financial condition of an enterprise accomplished through the intentional misstatement or omission of amounts or disclosures in the financial statements to deceive financial statement users. Despite minor variations in its definition, financial statement fraud in practice is a white-collar crime perpetrated by management insiders, usually involving overstating assets, revenues, and profits and understating liabilities, expenses, and losses. For instance, Enron's leadership cooked the books by using special purpose vehicles (SPVs) to hide its debts and liabilities from investors, creditors and regulators. When Enron declared bankruptcy in October 2001, its shareholders lost US$74 billion and its employees lost their jobs and billions of dollars in pension benefits, ranking it the largest bankruptcy reorganization at that time. Comparably, the Securities and Exchange Commission (SEC) charged six former top executives of Fannie Mae and Freddie Mac with securities fraud in December 2011, alleging they knew and approved misleading statements claiming the companies had minimal holdings of higher-risk mortgage loans, including subprime loans between December 2006 and August 2008. Actually, Fannie Mae's loan products included more than US$43 billion of Expanded Approval (EA) loans towards borrowers with weaker credit histories. Meanwhile, Freddie Mac's subprime exposure was about US$141 billion on 31 December 2006, accounting for 10% of its portfolio, and grew to approximately US$244 billion on 30 June 2008, or 14% of its portfolio[6]. In an extreme case, Lehman Brothers hided over US$50 billion of its bad assets by employing an accounting trick known as the "Repo 105"[7] to lower its leverage level and maintained the AAA status until it crashed on 15 September 2008.

Recently, the Theranos fraud represented a significant scandal in the world of technology startups, touching upon issues of technological innovation, business ethics, and regulation. Theranos, founded by Elizabeth Holmes in 2003, soared to a $10 billion valuation by promising revolutionary blood tests. However, these claims were exposed as fraudulent in 2015, triggering legal and financial turmoil. Holmes and former president Sunny Balwani faced SEC fraud charges in 2018, ultimately leading to Holmes's conviction and 11-year prison sentence in 2022. This is not an isolated case. FTX Trading Ltd., colloquially known as FTX, imploded spectacularly in 2022, brought down by the colossal fraud orchestrated by its founder, Sam Bankman-Fried. Established in 2019, FTX had risen to prominence as a leading cryptocurrency exchange and crypto hedge fund, boasting over a million users at its peak. Prior to its downfall, FTX stood as the third-largest cryptocurrency exchange globally, facilitating a staggering US$10 billion in active trading volume during 2021, underscoring its formidable position in the digital asset landscape.

The pervasiveness of financial statement fraud is not limited to the U.S., but a widespread problem all over the world. In Europe, the implosion of Germany's Wirecard[8] had turned out to be one of Europe's biggest financial scandals. In June 2020, Wirecard dramatically announced a "missing" of 1.9 billion euros in cash. The company committed a few days later that the 1.9 billion euros amount likely did not exist and filled for bankruptcy on 25 June 2020. In Japan, the Olympus scandal was surfaced by its former president and chief executive officer Michael Woodford on 14 October 2011, that the company had been implicated in loss-concealing arrangments for two decades. Six former executives of Olympus (except the whistleblower Michael Woodford) were charged in a fine of US$529 million and sentenced years in prison for their roles in a US$1.7 billion cover-up since the 1990s. Coincidentally, the Toshiba accounting scandal revealed in 2015 was tied to unscrupulous accounting practices, such as booking future profits early, pushing back losses, and pushing back charges. The company had overstated profits by US$1.2 billion over the previous seven years.

In the modern global economic system, no country is immune from financial state fraud, and China is no exception. Since the establishment of the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE) in the end of 1990, the domestic Chinese stock market[9] has expanded rapidly with China's runaway economic growth in the last three decades. Based on data from the China Securities Regulatory Commission (CSRC), **Figure 1** provides a development trend of the Chinese stock market in the recent years. As of the end of 2020, there were 4154 companies listed on the SSE and the SZSE, with an average increase of 8% year-on-year over the past 5 years. The total market capitalization combining the SSE and the SZSE reached US$12.2 trillion in 2020, ranking the Chinese mainland stock market the second largest stock market by country or region in the world, trailing only the U.S. equity markets. However, the development of China's stock market is accompanied by a proliferation of fraudulent activities. From the earliest case of Shenzhen Yuanye[10] in the 1990s, to the Yinguangxia[11] scandal in the 2000s, and the infamous LeEco[12] scam in recent years, the scale and scope of accounting frauds are both expanding fast with the expansion of China's capital market.

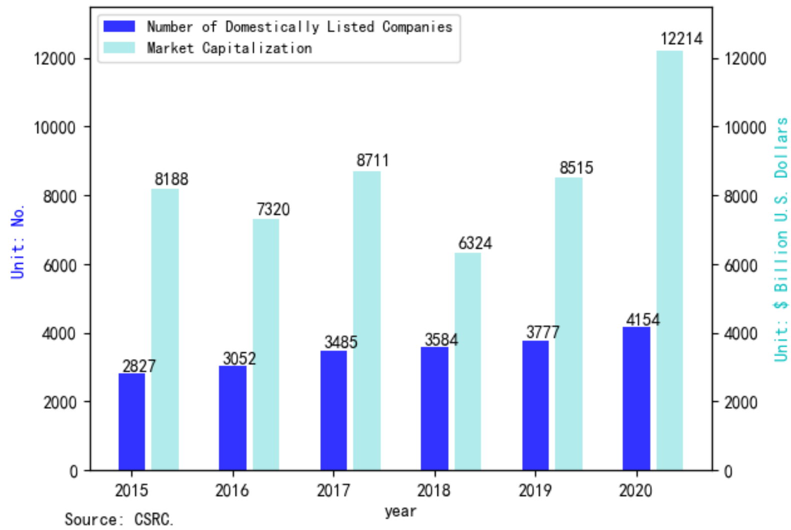

Figure 1. Market size of companies listed on the SSE and the SZSE (2015–2020).

With the prevalence of corporate fraud and malfeasance comes the public outcry for regulation and reforms. In the U.S., the financial reporting scandals of Enron and Worldcom[13] gave rise to the Sarbanes-Oxley (SOX) Act of 2002, a federal law that expanded reporting requirement for all the U.S. public company boards, management, and public accounting firms to ensure that companies report their financials honestly. In China, the CSRC, the main regulator of the securities industry, has been constantly committed to cracking down on stock market manipulation to protect investors, especially the medium and small shareholders. **Figure 2** presents the CSRC actions against securities violations in 2015–2020. Although the new investigation cases were approximately 300 cases in each year, the proportion of disclosure violation cases is steadily increasing per year. In 2020, the CSRC received 435 valid tips on illegal activities and misconducts, newly launched 353 investigations, and opened 282 cases. Among the 282 new investigations in 2020, 82 cases were related to material disclosure violations, which is a dramatic increase from the 42 cases of disclosure violations in 2015. It shed light on that information disclosure violation is becoming one of the most common types of financial statement fraud in China[14].

Earlier this year, China's regulatory authorities accused Evergrande and its founder, Xu Jiayin, of allegedly inflating revenues by a staggering US$78 billion, making it the center piece of the country's largest financial fraud case. Xu Jiayin, the founder and chairman of the Evergrande Group, was subsequently fined 47million yuan (approximately US$6.5 million) for the overstatement and other alleged infractions. Furthermore, PwC China, the auditing firm that had approved Evergrande's financial statements, has come under fire. Despite Evergrande's inflation of its mainland revenues by nearly 80 billion in the two years preceding its 2021 default, PwC China failed to uncover the discrepancies. In response, Chinese authorities imposed a six−month ban on PwC China and levied a fine of 441 million yuan (approximately US$62 million), citing staff members' concealment or even condoning of fraud in their audit failures related to the collapsed property developer.

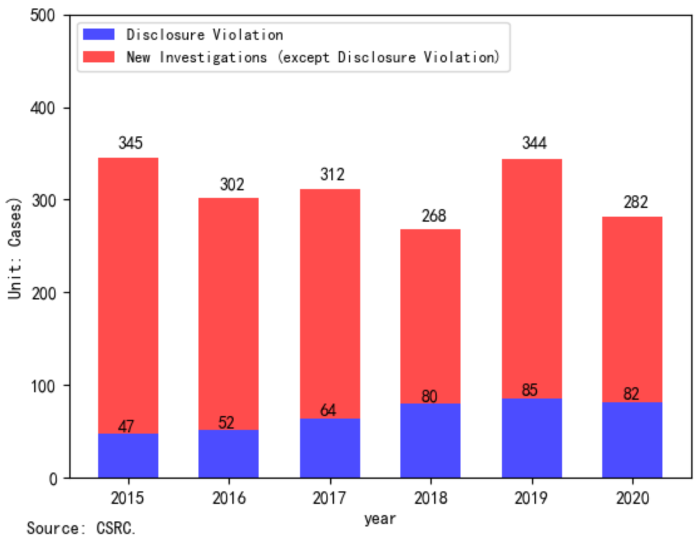

Figure 2. CSRC actions against securities violations (2015–2020).

The key to combat financial statement fraud is to predict it as early as possible, all the better to prevent it from ever happening. the on-site examinations are more costly USA status. A faster and efficient fraud detection strategy can substantially reduce the magnitude and loss of fraud. However, detecting financial statement fraud has long been a global challenge. Among existing data mining techniques of financial shenanigans, logistic regression is one of the most widely used methods to detect financial statement fraud [2], and somewhat surprisingly it also performs well relative to other more complex detection techniques [3]. Nevertheless, the logistic regression or other data mining models incorporate estimation errors which may affect the classification of the financial statement fraud. Alternatively, this paper aims to contribute on the topic of detecting financial statement fraud by applying an early warning model based on the mixture hazard model (MHM) of Farewell [4]. This early warning model approach is introduced by Almanidis and Sickles [5] which effectively combine a static model to identify companies with suspected fraud and a duration model to provide estimates of the probability of detecting such fraud by the regulators. The logistic regression is used as a benchmark to identify suspected financial shenanigans. The main assumption of this approach concerns the fact that companies with low performance as measured by the financial characteristics in their annual accounting reports will increase their probability of cooking the books.

The estimation results indicate that in detecting financial statement fraud, the financial characteristics may have different implications in different industries. For example, in the construction industry, manipulation of financial data is probably followed by a low level of the return on assets (ROA). In other words, a low level of the ROA implies a poor usage of its capital to generate profitability, hence provides a motivation for financial statement fraud. In the capital-intensive construction industry in China, profitability instead of growth is the key to survive and keep a company off the financial statement fraud. Alternatively, in the real estate industry, financial statement fraud are more closely linked to a high level of inventory. In other words, a low inventory turnover ratio will increase the company's finance cost, worsen its liquidity, and induce the management to cook the books. The intense financial pressure in the real estate industry in China, mainly from the high leverage, high debt and high turnover business mode, may motivate desperate management of a company to manipulate its financial data. We also find evidence from the MHM model that better financial conditions, which are measured by the Altman's Z-score [6], will reduce the risk of financial statement frauds.

The rest of the paper is organized as follows. Section 2 provides a literature view on financial statement fraud detection. Section 3 introduces the MHM model and its estimation method. Section 4 describes the dataset and defines the variables for this research. Section 5 gives the estimation results with a comparison of detecting financial statement fraud in both the construction industry and the real estate industry in China. Section 5 provides a summary.

2. Literature review

Research into financial statement fraud has traditionally focused on understanding its causes, consequences, and detection methods. Reurink [7] distinguishes financial statement fraud from financial scams and fraudulent financial mis-selling and finds that financial fraud is a complex phenomenon that can take very different forms, depending on the market segments in which it occurs, the financial instruments it pertains to, and the actors involved. Amiram et al. [8] pointed out that there is conflicting evidence in the literature concerning whether internal controls are effective in reducing financial misconduct within organizations. In contrast, Rashid et al. [9] investigate the literature concerning corporate fraud as well as financial crime from 2003 to 2018 and found that the internal control system, as a component of good governance, is the best approach to prevent and detect fraud. Ashtiani and Raahemi [10] systematically review and synthesize the existing literature on intelligent fraud detection in corporate financial statements, suggesting that in addition to structured data such as financial ratios, unstructured data in the financial fraud detection domain, such as textual content, deserves more attention to achieve remarkable results.

Effective detection of financial statement fraud is critical in mitigating its impacts. Traditional methods such as logistic regression have been widely utilized to identify potential fraud [11]. One notable advancement in fraud detection is the use of early warning systems as outlined by Almanidis and Sickles [5]. This approach integrates static models with duration models, offering a more nuanced and comprehensive method for identifying fraudulent activities. Hajek and Henriques [12] argue that a low relative frequency of negative words in annual reports may indicate non-fraudulent firms and it is necessary to use information extracted from both publicly available financial statements and analysts' forecasts of revenues and earnings. Kassem [13] finds that effective corporate governance can help reduce fraud risk, prevent fraud and detect fraud, particularly corporate fraud, insider fraud and asset diversion. Mandal [14] examines how auditors perceive the influence of crucial fraud prevention factors in deterring financial statement fraud within the corporate sector. Chen and Han [15] model the corporate financial data into a 3-Dimension data cube that contains both temporal and financial feature domain information, and find that the detection performance is dramatically improved when the accumulated number of quarters increased.

Recent developments in fraud detection focus on big data techniques, especially data mining techniques [16]. Hooi et al. [17] propose a fraud detection algorithm named FRAUDAR, which provably bounds the amount of fraud adversaries can have, even in face of camouflage. Shin et al. [18] propose DenseStream, an incremental algorithm that maintains and updates a dense subtensor in a tensor stream, and DenseAlert, an incremental algorithm spotting the sudden appearances of dense subtensors, which successfully detects anomalies, including small-scale attacks, in real-world tensors. Craja et al. [19] employ a hierarchical attention network model, by extracting textual features from the annual reports and providing red-flag sentences, significantly enhances fraud detection capabilities by combining both content and context of managerial comments with financial ratios. Jiang et al. [20] propose a real-time fraud detection framework called SPADE, which enables developers to design their fraud semantics to detect fraudulent communities by providing the suspiciousness functions of edges and vertices. Jiang et al. [21] further refine the Spade fraud detection system to Spade+, which manages both edge additions and removals on dynamic graphs and efficiently handles batch updates and employs edge packing to diminish latency. Chen et al. [22] present RUSH, a novel system for efficiently detecting burst subgraphs in dynamic graphs, and offerr user-friendly APIs for customizable fraud detection across different scenarios.

3. The model

This paper follows the MHM approach of Almanidis and Sickles [5] to identify companies who cook the books. First, we define H_{it} as the financial health indicator of corporate i at time t. If the financial health H_{it} of the company i falls bellow a certain level, which can be represented by a threshold H_{it}^{*}, then the company is considered at risk of cooking the books. The difference between the threshold level H_{it}^{*} and H_{it} is denoted as $h_{it}* = H_{it}^{*} - H_{it}$, and is assumed to be dependent on company-specific financial metrics as

$$h_{it}^{*} = x_{it}'\beta + e_{it}, \tag{1}$$

where e_{it} represents the identically and independently distributed (iid) error term. Empirically, the financial threshold level $H_{it}*$ and the h_{it}^{*} are not observable. Instead, whether a company manipulates its financial record can be defined by a binary variable h_{it} as

$$h_{it} = \begin{cases} 1 & \text{if } h_{it}^{*} > 0 \\ 0 & \text{if } h_{it}^{*} \leq 0. \end{cases} \tag{2}$$

Hence, the probability that a company will become a $problem$ one is given by

$$P(h_{it} = 1) = P(h_{it}^{*} > 0) = P(e_{it} > -x_{it}\beta) = F_e(-x_{it}\beta), \tag{3}$$

where F_e is the cumulative distribution function of the random error e, which can be defined as normally distributed in the probit model or logistically distributed in the logit model. The main assumption of this approach concerns the fact that companies with low performance as shown in the accounting report will increase the probability to intentionally misstate their financial statement.

In practice, the rate at which regulatory authorities tend to misclassify healthy companies as fraudulent ones is close to zero. Hence the likelihood function for a company i can be expressed as

$$L_i(x, w) = [F_e(x\beta)\lambda_i^p(t; w_i)S^p(t; w_i)]^{d_i}[F_e(x_i\beta)S^p(t; w_i) + (1 - F_e(x_i\beta))]^{1-d_i}, \tag{4}$$

where λ^p represents the hazard rate or probability that such company will cheat during the next instant and S^p is a survivor function, which represents the probability that a problem company will not be detected for a period longer than t. A binary variable d_i takes on a value of 1 for observations of the fraud by regulators with a warning letter at time t and 0 otherwise. The variables of x and w are covariates associated with the probability of financial statement manipulation and the probability of been detected, respectively. For convenience of interpretation, we refer to the portion of the model that assesses the financial health of a company as the incidence component, and the portion of the model that assesses survivial times of financial statement frauds before been detected as the latency component. After some simplification of Equation (4), the sample likelihood for all the companies can be written as

$$L(x,w,d) = \prod_{i=1}^{n} L_i(x,w,d) = \prod_{i=1}^{n} F_e(x\beta)^{h_{it}}(1-F_e(x_i\beta))^{1-h_{it}}(\lambda_i(t;w_i))^{d_i h_{it}}(S_i(t;w_i))^{h_{it}}. \tag{5}$$

Following the approach of Kalbfleisch and Prentice [23], the discrete-time hazard function $S_i(t;w_i)$ in Equation (5) is given by

$$S_{ij}(t;w) = \prod_{j=1}^{t_i} \frac{1}{1+e^{w_{ij}\alpha}}, \tag{6}$$

and the discrete-time survivor function $\lambda_i(t;w_i)$ can be derived as

$$\lambda_{ij}(t;w) = 1 - \frac{S(t_{ij})}{S(t_{i,j-1})} = \frac{\exp(w_{it}\alpha)}{1+\exp(w_{it}\alpha)}, \tag{7}$$

where both the survivor function in Equation (6) and the hazard function in Equation (7) are assumed to be in discrete times of $j = 1, 2, ..., t_i$.

In the real life, h_{it} is not observable. We follow the approach of the expectation–maximization (EM) algorithm by Dempster et al. [24] to deal with the unobservable latent variablesapply. The EM algorithm consists of two iterative steps: the expectation (E) step and the maximization (M) step. The E step creates a function for the expectation of the log-likelihood evaluated using the current or initial estimate for the parameters, and the M step computes parameters maximizing the expected log-likelihood found on the E step while treating the incomplete data as known. Iterating between these two steps yields estimates that under suitable regulatory conditions converge to the maximum likelihood estimates. In detail, in the E step, we compute h_{it} as

$$\begin{aligned} h_{it} &= \Pr(h_{it}=1 \mid t_i > T_i) \\ &= \begin{cases} \frac{F(x_i\beta)S_i(w_i)}{F(x_i\beta)S_i(w_i)+(1-F(x_i\beta))} & \text{if } d_i = 0 \\ 1 & \text{otherwise.} \end{cases} \end{aligned} \tag{8}$$

In the M step, we optimize the log-likelihood function of Equation (5) as

$$E(L) = h_{it}\ln(F(x_i\beta)) + (1-h_i)\ln(1-F(x_i\beta)) + h_i d_i \ln(\lambda_i) + h_i \ln(S_i). \tag{9}$$

In practice, the initial estimate for the parameter β is calculated by the logistic regression. Then we iterate between Equations (8) and (9) until convergence is reached.

4. The data

Our data for the listed companies in the Chinese stock market are from the iFinD database. As it is well known that the real estate sector is among the most crucial sectors of the Chinese economy, we choose companies in the construction industry and the real estate industry during 2015–2020 to study the features of the financial statement fraud in the real estate market. Hence the sample includes 504 and 626 observations from the construction industry and the real estate industry, respectively. Based on the annual revenue in 2015–2020, the market size of the sample in these two industries is around 20% of all the listed companies in China's stock market.

The dependent variable is a binary variable, labeling if the company received a warning letter of misstatement from the CSRC. Following Summers and Sweeney [25], we select four financial statement characteristics, including financial condition, financial performance, growth, and changes in inventory, as explanatory variables to model the propensity to commit fraud. The advantages of choosing these four financial statement characteristics as explanatory variables are twofold. First, these characteristics are public data available in the annual report of the listed companies, avoiding the missing data or biased data issues if the data are collected from unofficial channels. Second, it provides a criteria in detecting unethical misstatement based on key information from the financial statement.

The measurement of financial condition is proxied by Altman's Z-score, which takes into account profitability, leverage, liquidity, solvency, and activity ratios to predict whether a company has a high probability of becoming insolvent [6]. In general, a Z-score closer to or higher than 3 implies a solid financial situation, while a Z-score below 1.8 implies a poor financial condition which might motivate unethical financial reporting. The financial performance is measured by the ROA, which is a bang-for-the-buck financial ratio calculated by dividing a company's net income over its total assets. A low level of ROA indicates that the management could not maintain or improve the profitability of the company, which might trigger financial statement fraud. The GROWTH is computed as the simple average of the percentage growth in sales over a three year window ending in the current year (t). Rapid growth is expected to be associated with the incidence of fraud, because unethical managers may be induced to misstate financial statements when growth slows or reverses in order to maintain the appearance of consistent growth. Therefore, the usage of a three year average growth rate provides a measure of sustained growth of a company, avoiding the short-run fluctuations by the measurement of one year growth rate. The changes in inventory, ΔI, is defined as the changes in the ratio of inventory (I) to sales (S) between the current year (t) and the previous year ($t-1$) that $\Delta I = \frac{I_t}{S_t} - \frac{I_{t-1}}{S_{t-1}}$. Since overstating ending inventory has the effect of reporting higher profit, management may be motivated to inflate income through the misstatement of inventory records.

Table 1 shows the descriptive statistics of the variables. Among the 504 observations from the construction industry in 2015–2020, there are 11 detected fraudulent financial statement cases, while among the 626 observations from the real estate industry in 2015–2020, there are 6 financial statement fraud cases been detected. The distribution of the explanatary variables, such as the Z-score, the GROWTH, and the ΔI are heavily right-skewed, implying that a few leading companies outperform the others in the market, especially in the real estate industry. On the other hand, the similar distribution of the ROA in the construction industry and the real estate industry is consistent with the fact that both industries are highly capital intensive. The next section will further discuss how the changes in these financial characteristics may give clues in identifying the financial statement fraud in both industries.

Table 1. Descriptive statistics.

	Min	Q1	Median	Mean	Q3	Max	N
Construction Industry							
FRAUD	0.000	0.000	0.000	0.022	0.000	1.000	504
Z-score	−5.639	1.238	1.678	2.247	2.530	27.536	504
ROA	−2.509	0.017	0.031	0.030	0.053	0.252	504
GROWTH	−35.343	4.820	11.590	15.314	21.961	235.742	504
ΔI	−2.766	−0.083	−0.003	−0.029	0.034	5.733	504
Reak Estate Industry							
FRAUD	0.000	0.000	0.000	0.010	0.000	1.000	626
Z-score	−9.776	0.972	1.333	3.892	2.329	248.760	626
ROA	−0.833	0.013	0.027	0.027	0.047	0.345	626
GROWTH	−64.161	1.119	16.714	32.684	32.286	600.800	626
ΔI	−78.115	−0.646	−0.008	0.019	0.566	99.235	626

5. Estimation and inference

Since financial statement fraud is a rare but deliberate action, although there are many attempts to study the detecting methods of the financial statement fraud, it can be very hard to predict. This study chooses three econometric approaches, including the ordinary least squares (OLS) regression, the logistic regression, and the MHM model, to compare the key affecting financial factors of fraudulent misstatement cases in the construction industry and the real estate industry in China. The OLS is the most widely used linear regression method. However, in this study the dependent variable is a binary outcome. Although there are cases to model binary outcomes using the OLS regresion, in general it is not recommended because of the interpretation difficulties and the issue of inconstant conditional variances. The logistic regression is an althernative approach for binary outcomes, but it still depends on distributional assumptions of the unknown error term. As discussed in Section 2, the MHM early warning model is a preferred approach in this study, since it combines a static model to identify suspected fraud companies, and a duration model to estimate the probability of detecting such frauds by the regulators.

Table 2 provides the estimation results of all the three models for the listed companies in the construction industry in China. Among the four key financial characteristics of the Z-score, the ROA, the GROWTH, and the ΔI, the ROA is significantly negatively correlated to detected misstatement in all the three models. As the return on capital slows down in China in the past decades, the capital-intensive construction industry is struggling in a world of painfully low margins. A high level of ROA could give the company in this industry a wide economic moat against the risk of financial statement fraud. At the same time, though the relationship between the detected misstatement and the variables of the Z-score, the GROWTH, and the ΔI are vague in the OLS and the logit models, the unambiguous negative impacts of the Z-score and the ΔI to the financial statement fraud are shown in the incidence portion of the MHM model. A higher Z-score indicates a solid financial situation, while a higher inventory level implies potential business expansion in the near future, both will reduce the threat of financial statement fraud. On the other hand, it is worth paying attention to the positive effect of the GROWTH to the fraudulent misstatement, which might suggest that the increasement of revenue, if not accompanies by the increasement of profit, would not be sound or robust for a company but to trigger fraudulent dealings.

Table 2. Estimates for the construction industry.

Variables	**OLS**	**Logit**	**MHM Model**	
			Latency	**Incidence**
Intercept	0.027** (0.010)	−3.294*** (0.765)		−1.736 *** (0.064)
Z-score	−0.001 (0.003)	−0.416 (0.447)	0.018 (0.082)	−0.172 *** (0.023)
ROA	−0.417*** (0.066)	−13.655 ** (5.502)	−0.297 (1.478)	−6.635 *** (0.252)
GROWTH	0.053 (0.032)	1.638 (1.424)	−0.121 (0.392)	0.569 * (0.241)
ΔI	−0.024 (0.016)	−0.633 (0.573)	−0.062 (0.166)	−0.268 *** (0.058)
N	504	504	504	504

NOTE: $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Table 3 demonstrates the estimates for the listed companies in the real estate industry in China. Among the four key financial characteristics of the Z-score, the ROA, the GROWTH, and the ΔI, only the ΔI demonstrate s significant and positive effects to detected misstatement in all the three models. As is well known, the real estate industry is highly capital intensive in China. With the raipd urbanization in the last decades, property enterprises in China expand quickly via the high leverage, high debt and high turnover business mode. Hence to control the level of inventory becomes critical for companies in the real estate industry to reduce finance cost and improve liquidity. Since a increase of ΔI will cause increased financial burden of a company, the management might have the motivation to cook the books in urgent cases of financial difficulties. Concurrently, the impacts of the Z-score and the GROWTH to the detected misstatement are positive and visual, similar as the results shown in **Table 2** for the construction industry.

Table 3. Estimates for the real estate industry.

Variables	OLS	Logit	MHM Model	
			Latency	Incidence
Intercept	0.009* (0.004)	-4.831^{***} (0.53)		-2.402 *** (0.012)
Z-score	-0.000 (0.000)	-0.023 (0.092)	0.003 (0.002)	-0.009 *** (0.001)
ROA	-0.079 (0.058)	-3.407 (3.269)	-0.015 (0.042)	-1.643 *** (0.017)
GROWTH	0.010 * (0.005)	0.239 (0.363)	0.026 (0.016)	0.098 (0.016)
ΔI	0.003 *** (0.000)	0.050 *** (0.017)	-0.001*** (0.000)	0.025 *** (0.000)
N	626	626	626	626

NOTE: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Furthermore, the estimation results in **Tables 2** and **3** provide evidence that the MHM model delivers more information than the OLS regression or the logistic regression in detecting the financial statement fraud for the listed companies. Specifically, the incidence portion of the MHM model, which assesses the financial health of a company, is analogous to the OLS regression or the logistic regression, while the latency portion of the MHM model provides additional information of the survivial times of financial statement frauds before been detected. However, the estimation results of the latency portion of the MHM model are statistically not significant for either the construction industry or the real estate industry, which also support the facts that the fraudulent misstatements are hard to be detected.

6. Summary and conclusions

Financial statement frauds or accounting scandals arising from intentional manipulation of financial statements are ubiquitous among listed companies in China. Since the disclosure of financial misdeeds are from trusted executives of the public corporations, the fraudulent financial statements are difficult to be detected. Based on the fact that a common incentive for companies to manipulate financial statement is a decline in the company's financial prospects, this study apply the MHM early warning model to identify the key impacting financial characteristics in detecting the financial statement fraud. The estimation results indicate that in the construction industry the warning sign of suspecious business practices is a falling ROA, while in the real estate industry the financial red flag is an increase in the inventory level.

The policy implications of this study are twofold. First, the financial characteristics may have different implications in different industries in detecting financial statement fraud. The financial pressures for the management to commit fraud, which are constrained by the industry heterogeneity such as capital intensity, resource intensity, and technology intensity, will be manifested in different financial characteristics in the company's accounting report. Hence, setting specific financial characteristics by peculiarities of the industry, such as the ROA for the construction industry, and the inventory ratio for the real estate industry, would be recommended for fraud monitoring and detecting by the regulators. In the aftermath of the Evergrande crisis, China's Real Estate Market has emerged as a prominent source of economic and financial risk, prompting the government to prioritize comprehensive measures aimed at addressing the underlying issues in the sector to effectively prevent and mitigate systemic financial risks. In response to the rising inventory ratio in the real estate market in the post-COVID period, Chinese government encourages real estate enterprises to transform from traditional development and sales models towards diversified and integrated business models. This includes developing long-term rental apartments, senior living communities, tourism real estate, and strengthening property management and community services to enhance profitability and market competitiveness.

Second, it is vital to a company, instead of the market regulators, to have a fraud prevention plan in place, as preventing fraud is much easier than recovering the losses after a fraud has been committed. In practice, internal accounting controls, including segregation of duties, the usage of an external auditor, and performing accounting reconciliations on a regular basis, can help to reduce the risk of fraud from occurring. In the Chinese real estate market, while not all companies may publicly disclose their detailed fraud prevention plans, many have implemented a range of measures to safeguard against fraud and ensure market fairness and transparency. By strengthening internal controls, enhancing transparency, refining contract and after-sales services, overseeing advertising, and rigorously managing partners, real estate companies can effectively mitigate fraud risks, uphold market order, and protect consumer interests. Moreover, with advancements in technology and regulatory oversight, real estate firms are increasingly leveraging digital tools like big data and AI to monitor sales data, promptly detect and address anomalies, further enhancing market fairness and transparency.

At the same time, the principle of the regulators should be grounded in strengthening legal foundation, non-intervention in the market, and zero tolerance against violations. As China's domestic economy experiences a slowdown compared with the frantic pace of growth of the past decades, an efficient, stable, and well-designed capital markets become more critical than ever to China's sustainable growth. On the one hand, China is deepening the reform of the capital markets constantly. The Financial Stability and Development Committee was established in June 2017 to further collaboration and alignment among the various regulatory bodies in China. The new securities law came into effect on March 1^{st} 2020 to allow for registration-based systems for IPOs. Additionally, disclosure requirements and fines for violations have been increased to improve investor protection. On the other hand, China is taking steps to further opening-up the Chinese financial markets to foreign institutional investors. The increased connectivity between domestic and foreign capital markets will bring new opportunities and challenges for both the companies, the investors, and the regulators. Detecting financial statement fraud will remain a long-term task with the further reform and opening-up of the Chinese stock market.

Author contributions: Conceptualization, YD and GQ; methodology, GQ; software, GQ; validation, YD; formal analysis, YD and GQ; investigation, GQ; resources, YD; data curation, YD and GQ; writing—original draft preparation, GQ; writing—review and editing, YD; project administration, GQ; funding acquisition, YD. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Notes

1 `ftc.gov/business-guidance/blog/2022/02/ftc-2021-data-just-facts`

2 Enron Corporation (NYSE: ENE), an American energy company based in Houston, Texas, was discovered in 2001 that the company had been using accounting loopholes to hide billions of dollars of bad debt.

3 Tyco International Ltd. (NYSE: TYC) was discovered in 2002 that its CEO, Dennis Kozlowski, and CFO, Mark Swartz, had stolen over US$150 million from the company and had inflated the company's earnings by over US$500 million in their reports.

4 The Federal National Mortgage Association, commonly known as Fannie Mae, and the Federal Home Loan Mortgage Corporation, better known as Freddie Mac, are federally backed home mortgage companies created by the U.S. Congress. They purchase, package, and sell home loans in the form of mortgage-backed securites (MBS) to provide liquidity to the U.S. martgage finance system.

5 Lehman Brothers (NYSE: LEH) was the fourth-largest investment banks in the U.S. with 25,000 employees worldwide before filling for bankruptcy in the climax of the subprime mortgage crisis on 15 September 2008.

6 `sec.gov/news/press/2011/2011-267.htm`

7 Repo 105 is Lehman Brothers' name for an accounting maneuver that it used where a short-term repurchase agreement is classified as a sale.

8 Wirecard was a German fintech company providing payment processing and financial services. The firm was a constituent of the DAX index between 2018 and 2020 before its collapse.

9 The domestic Chinese stock market discussed in this paper mainly comprises three independently operated stock exchanges: the Shanghai Stock Exchange, the Shenzhen Stock Exchange, and the newly founded Beijing Stock Exchange. The Hongkong Stock Exchange and the Taiwan Stock Exchange are not included.

10 Shenzhen Yuanye Textile Co., Ltd., established in 1987, was China's first listed Sino-foreign joint venture in 1990. The company is ordered by the SZSE to suspend trading on 7 July 1992, which is the first listed company fraud case in China's stock market history. Its stock resumed trading on 3 January 1994, after reorganization and changing the company's name to "Shenzhen Century Xingyuan Group Co., Ltd.", also the first listed company to be restructured.

11 In 2001, China Securities Regulatory Commission unearthed non-existent profits at Guangxia (Yinchuan) Industry Co., Ltd. of 178 million yuan in 1999 and 567 million yuan in 2000.

12 LeEco, formerly known as Letv, the first Chinese video website listed company, was delisted from the SZSE on July 20, 2020. The CSRC announced the details of the LeEco financial fraud on 13 April 2021, that LeEco has inflated revenue by 1.87 billion yuan, and inflated profit by 1.73 billion yuan in 10 consecutive years from 2007–2016. The CSRC fined LeEco a total of 241 million yuan and its founder Jia Yueting a total fine of 241 million yuan.

13 Worldcom was the second largest U.S. telecommunication company, after the AT&T in the 1990s. The senior executives at Worldcom inflated revenue by over US$3.8 billion from 1999 to 2001 to maintain its stock price. After the scandal broke, the company filed for bankruptcy protection on 21 July 2002 and was eventually acquired by Verizon in January 2006

14 `csrc.gov.cn/csrc_en/c102063/c1606114/content.shtml`

References

1. Association of Certified Fraud Examiners (ACFE). Report to the nations on occupational fraud and abuse global fraud study 2020. ACFE; 2020.
2. Zhou W, & Kapoor G. Detecting evolutionary financial statement fraud. Decision support systems. 2011; 50(3): 570–575.doi:10.1016/j.dss.2010.08.007
3. Perols J. Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing: A Journal of Practice & Theory. 2011; 30(2): 19–50. doi:10.2308/ajpt-50009
4. Farewell VT. A model for a binary variable with time-censored observations. Biometrika. 1977; 64(1): 43–46. doi:10.1093/biomet/64.1.43
5. Almanidis P, & Sickles RC. Banking crises, early warning models, and efficiency. In: Advances in Efficiency and Productivity. Springer, Cham. 2016; pp. 331–364. doi:10.1007/978-3-319-48461-7_14
6. Altman EI. Predicting financial distress of companies: revisiting the Z-score and ZETA® models. In Handbook of research methods and applications in empirical finance. Edward Elgar Publishing; 2013.doi:10.4337/9780857936080.00027
7. Reurink A. Financial fraud: A literature review. Journal of Economic Surveys. 2018; 32(5): 1292–1325.doi:10.

1111/joes.12294

8. Amiram D, Bozanic Z, Cox JD, et al. Financial reporting fraud and other forms of misconduct: a multidisciplinary review of the literature. Rev Account Stud. 2018; 23: 732-–783. doi:10.1007/s11142-017-9435-x
9. Rashid MA, Al-Mamun A, Roudaki H, & Yasser QR. An overview of corporate fraud and its prevention approach. Australasian Accounting, Business and Finance Journal. 2022; 16(1): 101–118.doi:10.14453/aabfj.v16i1.7
10. Ashtiani MN, & Raahemi B. Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review. Ieee Access. 2021; 10: 72504–72525. doi:10.1109/ACCESS.2021.3096799.
11. Albashrawi M. Detecting financial fraud using data mining techniques: a decade review from 2004 to 2015. Journal of Data Science. 2016; 14(3): 553–569. doi:10.6339/JDS.201607_14(3).0010
12. Hajek P, & Henriques R. Mining corporate annual reports for intelligent detection of financial statement fraud–A comparative study of machine learning methods. Knowledge-Based Systems. 2017; 128: 139–152. doi:10.1016/j.knosys.2017.05.001
13. Kassem R. Elucidating corporate governance's impact and role in countering fraud. Corporate Governance: The International Journal of Business in Society. 2022; 22(7): 1523–1546. doi:10.1108/CG-08-2021-0279
14. Mandal A. Preventing financial statement fraud in the corporate sector: insights from auditors. Journal of Financial Reporting and Accounting; 2023. doi:10.1108/JFRA-02-2023-0101
15. Chen ZY, & Han D. Detecting corporate financial fraud via two-stage mapping in joint temporal and financial feature domain. Expert Systems with Applications. 2023; 217: 119559. doi:10.1016/j.eswa.2023.119559
16. Gepp A, Linnenluecke MK, O'Neill TJ, & Smith T. Big data techniques in auditing research and practice: Current trends and future opportunities. Journal of Accounting Literature. 2018; 40(1): 102–115. doi:10.1016/j.acclit.2017.05.003
17. Hooi B, Song HA, Beute A, et al. Fraudar: Bounding graph fraud in the face of camouflage. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining; 13–17 August 2016; Francisco, CA, USA. pp. 895–904.doi:10.1145/2939672.2939747
18. Shin K, Hooi B, Kim J, & Faloutsos C. Densealert: Incremental dense-subtensor detection in tensor streams. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 13–17 Augus 2017; New York, NY USA. pp. 1057–1066.doi:10.1145/3097983.3098087
19. Craja P, Kim A, & Lessmann S. Deep learning for detecting financial statement fraud. Decision Support Systems. 2020; 139: 113421. doi:10.1016/j.dss.2020.113421
20. Jiang J, Li Y, He B, et al. Spade: A real-time fraud detection framework on evolving graphs. Proceedings of the VLDB Endowment. 2022; 16(3): 461–469. doi:10.14778/3570690.3570696
21. Jiang J, Chen Y, He B, et al. Spade+: A Generic Real-Time Fraud Detection Framework on Dynamic Graphs. IEEE Transactions on Knowledge and Data Engineering; 2024. doi:10.1109/TKDE.2024.3394155
22. Chen Y, Jiang J, Sun S, et al. RUSH: Real-Time Burst Subgraph Detection in Dynamic Graphs. Proceedings of the VLDB Endowment. 2024; 17(11): 3657–3665.
23. Kalbfleisch JD, & Prentice RL. The statistical analysis of failure time data. John Wiley & Sons; 2011.
24. Dempster AP, Laird NM, & Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological). 1977; 39(1): 1–22. doi:10.1111/j.2517-6161.1977.tb01600.x
25. Summers SL, & Sweeney JT. Fraudulently misstated financial statements and insider trading: An empirical analysis. Accounting Review. 1998; 131–146.