Stockholm School of Economics

Department of Accounting

Master Thesis in Accounting and Financial Management

Stockholm, 2017

Manipulation of Accruals and Real Activities for Earnings Management:

A Study of the Chinese Stock Market

Abstract: Since 2014, international investors gained unrestricted access to the Shanghai and Shenzhen Stock Exchanges for the first time since they were established. However, whether investors can seize the investment opportunities or not highly depends on the reliability of financial statements. Unfortunately, Chinese listed firms do not have a high reputation of transparency and honesty with their earnings. Even the Chinese government admits that accounting quality in mandatory disclosures is not satisfactory. At the same time, research about the reliability of reported earnings and the extent of earnings management does not reach consistent conclusion. This paper examines the extent of earnings management in a wide sample of listed Chinese companies. Both manipulation of accruals and of real activities are considered using models developed based on Jones (1991) and Roychowdhury (2006). There is evidence of significant earnings management, which is equivalent in size to 1.8-7.9 percent of total assets. The largest and most frequent manipulations occur in production costs. Furthermore, the paper discusses how alternative model specifications and estimation approaches affect the resulting estimated size of earnings management.

Keywords: Earnings Management, Chinese Listed Firms, Accruals Manipulation, Real Activities Earnings Management.

Authors: Biljana Ristevska, 40855@student.hhs.se; Yiming Wang, 40857@student.hhs.se

Tutor: Henrik Andersson

Acknowledgement

We would like to thank our thesis tutor, Henrik Andersson, Assistant Professor at the Department of Accounting at Stockholm School of Economics, for the challenging questions and valuable feedback. We would also like to thank Per-Olov Edlund, Associate Professor at the Center for Economic Statistics at Stockholm School of Economics, for insightful guidance and discussion. Last, but not least, we wish to thank our friends and family for all their love and support.

Biljana Ristevska, Yiming Wang

June, 2017

Table of Contents

1. Introduction

Foreign investors generally have a very complex view of the Chinese stock market. On the one hand, it is a part of the fastest sustainably expanding major economy in history. The market capital of the Shanghai and Shenzhen Stock Exchanges together counts nearly 6 trillion USD at the end of 2015. On the other hand, the reputation of “Made in China” is questionable due to the doubt towards its institutional background and accompanying transparency problem - “There’s no accounting for China’s accounting” (The Wall Street Journal, May. 2013). Not to mention even the Ministry of Finance once admitted that many listed companies were managing their earnings significantly which made their financial information less reliable for its users.

However, whether holding shares of firms listed in China is a wise idea was more a theoretical question than a practical one for foreign investors. For the majority of time since the opening of the Chinese stock markets in 1990, foreign investors did not have direct access. Until 2014, over 3000 A-shares in Chinese stock markets were only traded in Chinese Yuan (CNY), primarily among Chinese citizens and very limited number of qualified foreign institutional investors. Non-Chinese investors could freely invest only in cross-listed Chinese firms. The US SEC launched over 40 financial fraud investigation for those Chinese firms listed in the US, which further deepening the doubts.

In 2014, two new programmes launched, namely “Shanghai-Hong Kong Stock Connect” and “Shenzhen-Hong Kong Stock Connect”. These programmes allow investors in each market to trade shares on the other market using their local brokers and clearing houses. Therefore, theoretically, all international investors now have the access to trade eligible shares listed in mainland China. The Chinese stock market is more accessible than ever. Therefore, the quality of reported earnings of Chinese companies is critical for foreign investors to be able to understand investment opportunities in China.

The topic of earnings management has received some attention from researchers of the Chinese stock market. The scope of previous research is to apply models such as Dechow et al (1995) and Roychowdhury (2006) to Chinese data as the basis for the analysis of the size and frequency of earnings management. Those models were developed on data from the US, taking into consideration the relevant accounting practices. Researchers of the Chinese stock market have not contributed with significantly different models for estimating earnings management, presumably because they deem that the similarity between accounting systems and market characteristics is large enough to justify the application of those models.

We aim to do a comprehensive study of the size and frequency of Earnings Management in the Chinese market, with a particular focus on methodological choices appropriate for this task. Our approach differs from previous research in three aspects. Firstly, we question the use of the models which are specified for US conditions. The most significant similarity between US and China economy is the size, while in most other aspects, such as law and institution, including corporate governance, accounting standards they are very different.

Secondly, our study is based on a lager sample including observations from most recent available years. The data sample has a span between 2001 and 2015 and includes annual fundamentals of all non-financial firms listed in mainland China. China changes rapidly, and this study can give our readers a most updated understanding.

Thirdly, due to the focus on methodology, we show empirical results using alternative approaches in order to illustrate the differences and discuss the implications. These discussions can be useful for all studies of earnings management, not just those of the Chinese stock market.

The rest of the paper is organized as follows. Section 2 focuses on pertinent research and research questions. Section 3 describes the sample, EM measures and improved regression models. Section 4 presents empirical results and analysis. Section 5 concludes the paper.

2. Previous literature

Nowadays, accruals basis accounting is generally accepted and has been seen as a key to improve accounting information quality. Compared to cash basis accounting, accruals basis accounting provides more relevant information for financial decisions. For example, it matches and records the revenue, costs of goods sold and expense in an accounting period based on their causal relationship, thus gives more accurate information for performance measurement than just recording the cash flows when they occur. However, since accrual accounting usually requires estimation, for instance if the delivery of goods does not coincide with the payment, the necessity of the estimation provides space for manipulation of accounting numbers.

A widely-accepted definition of earnings management (EM) is given by Healy and Wahlen (1999): "EM occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers." This definition highlights the two main types of EM – manipulation of accruals, which depend on management's estimates, and manipulation of real activities, such as timing of expenses. Furthermore, Roychowdhury (2006) defines real-activities EM as operational decisions by managers that deviate from normal operational practices with the intention of achieving a desired level of earnings. Because EM is tied to managerial estimates and operating decisions, it is difficult for outside parties to see it. Hence, the detection of EM is the main challenge.

2.1Established regression models for measuring EM

EM cannot be measured directly. The most common method is to detect "abnormal" levels of accruals and expenses. Both are under the discretion of management judgement and thus are susceptible to manipulation.

To detect abnormal levels, one must know the normal levels of accruals and expenses. Then the abnormality i.e. manipulation is simply the difference between the actual level and the normal level. However, the normal levels also cannot be measured directly. Thus the EM estimation method involves two steps. Step one is to construct an estimation model which is based on accounting numbers that have a strong relationship with "normal" levels of the EM susceptible accounts. The point of the estimation model is to capture a sort of constant relationship between the observable accounting numbers and the unobservable or undistinguishable normal level of accruals and expenses. This relationship ideally is estimated during a period when there is no EM, called an estimation period. Otherwise, the estimated relationship would be between potentially manipulated levels of the EM susceptible accounts and accounting numbers. The coefficients generated in the estimation model are used in the second step of the EM estimation method.

Step two is to generate a prediction model. This model is applied in the event period, when one expects that there is EM. Thus the observed value of the EM measure is the actual level and the normal level is not known. The prediction model combines the coefficients from the estimation model with a company's specific circumstances in a particular year (captured in observable accounting numbers) to generate a normal level of the EM measure in the event year. Then the EM can be easily calculated.

This method is developed by Jones (1991). She tries to capture the relationship between total accruals and accounting numbers using company-specific OLS regression model. EM is measured as the abnormal component of total accruals. The underlying assumption is that normal accruals are not constant - they vary in accordance with the economic circumstances of the company; however, the normal level is approximately the OLS fitting line. Total accruals are defined as the non-cash component of working capital before income taxes payable and less the total depreciation expense. The estimation model and the prediction model are specified the same. Year-on-year change in revenues and gross PPE are used as proxies for economic circumstances of the firm. The regression equation is scaled by total assets to decrease heteroscedasticity (see Appendix A).

The OLS estimation distinguishes the estimation and prediction period by certain event that generates incentive of EM. Coefficients are estimated based on all observations before event period. The abnormal accruals are the residuals.

Two things should be noted. Firstly, the R-squared of the estimation model is 25 percent which means that the proxies that control for economic circumstances explain approximately 25 percent of the variation in scaled total accruals. Secondly, since change in revenues is used as a control variable, an implicit assumption is that this change is non-discretionary - this means that the model will not capture EM based on real activities, as far as it affects revenues rather than expenses.

Dechow et al (1995) build on the Jones (1991) model. The intention is to correct for the implicit assumption that revenues are nondiscretionary. The estimation of coefficients is exactly the same, while normal levels of total accruals are predicted by an adjusted regression. In other words, the estimation model is the same as Jones (1991) but the prediction model is not (see Appendix B). Instead of using changes in revenues, they use changes in revenue adjusted for the year-on-year change in accounts receivable as one of the proxies for economic circumstances. The assumption is that EM in revenues occurs with relation to sales on credit. The part of revenue which is due to suspicious sales should be excluded as its occurrence is questionable. This is done by subtracting the scaled increase in accounts receivables, where the 'bad' sales also appear.

The Modified Jones model developed here is the starting point of numerous later studies which attempt to explain the determinants of accruals-based EM. The accrual widely used in later studies is redefined by Hribar and Collins (2002) as *the difference between net income and operating cash flow*. They argue that the former definition, where accruals are equal to the net of noncash current assets and of noncash current liabilities is not a good measure for EM estimation since "other non-operating events such as M&A impact the current assets and liability accounts with no earning impact".

In addition to the change in definition of accruals, there is also a change in the estimation approach. Later studies test whether management characteristics (age, gender) or firm characteristics (year of formation, CEO transition) can be associated with EM. These studies often analyse large samples. Instead of estimating the coefficients of 'economic circumstances' for each firm separately in an ex ante event period, many of these studies take an industry-year specific approach by running cross-sectional regression. In each year, the actual total accruals are regressed against proxies for economic circumstances using observations for all firms within an industry in a particular year. This yields the coefficients which are then used to estimate normal levels of accruals for each firm in the industry of the same year, based on its own changes in revenues, receivables and PPE. Then, the difference between actual accruals and predicted normal accruals is taken as a measurement of discretionary/abnormal accruals.

This industry-year process seems strange. The OLS estimation that leads to coefficients for normal accruals ensures that the residuals are as small as possible. And the size of residual is highly affected by the fitting level of regression that measured by R-squared. Yet, these same residuals are used as a measure of EM. However, the alternative firm-specific approach is problematic as well - it requires that some firm-years are assumed to be free of EM (and so these are used for coefficient estimation) and other years where EM is expected to be found. But, there is no straightforward way of partitioning the firm-years - management could be manipulating earnings for a variety of reasons in any and all periods, but to a different extent.

Roychowdhury (2006) introduces the measurement of real-activities EM. He identifies three types of real-activities EM – manipulation of cash flow from operations (CFO), discretionary expenses (SG&A, advertising, R&D) and production costs. Significantly lower discretionary expenses indicate upward EM – expenses are postponed to future periods. Significantly higher production costs also indicate upward EM - overproduction results in fixed costs being spread over more units, thus reducing the cost per unit and hence COGS. The effect of abnormal CFO depends on the circumstance - selling on discount will increase it while increasing production costs will decrease it, but both actions will result in upward EM.

Revenue, lagged revenue and change in revenue are used as explanatory variables. The regression equations are also scaled by total assets to decrease heteroscedasticity and applied to each industry-year.

Cohen and Zarowin (2010) combine insights from Roychowdhury (2006) and Zang (2012) (based on her working paper at the time) to aggregate the three measures of real activities manipulation in Roychowdhury (2006) into two measures. The first measure is the sum of abnormal production costs and discretionary expenses, where the latter are first multiplied by negative one. The second measure is the sum of abnormal cash flow from operations and abnormal discretionary expenses, both multiplied by negative one. For both measures, the higher the measure is, the more earnings are manipulated upward. By introducing these two variables, the authors can identify firms with extreme production costs and extreme aggregate real EM and then compare these extreme situations to the rest of sample.

2.2Thresholds and earnings distribution study

Burgstahler and Dichev (1997) are the first to develop a different approach to estimating the extent of EM directly, rather than estimate abnormal accruals or expenses. They analyse the cross-sectional distribution of reported earnings scaled by market values. The main premise is that earnings are managed up to avoid losses or to avoid decreases (negative year-on-year changes). The former would manifest itself in the distribution of *scaled earnings* with slightly negative earnings occurring more rarely than they should and slightly positive earnings occurring more often that they should. The latter would manifest itself in the distribution of *scaled changes in earnings* with slightly negative changes occurring relatively rarely and slightly positive changes occurring relatively often. They do find jumps in the distribution around the zero threshold (see Appendix C) consistent with the incentives for EM.

This logic is linked with contemporaneous research which shows that the capital market creates a key incentive for EM. Barth et al (1999) find that firms with a pattern of increasing earnings have higher price-earnings multiples than other firms, suggesting that the stock market puts a premium on sustained earnings performance. Furthermore, there is a significant decline in the multiple when the pattern is broken.

In order to test whether losses or decreases in earnings occur more frequently than expected, one needs to make an assumption about the expected distribution. Burgstahler and Dichev (1997) assume 'smoothness'. The number of observations within a certain interval of the distribution should be the average of the number of observations in the two adjacent intervals. The empirical analysis shows that the assumption of smoothness should be rejected - earnings levels and changes in earnings are not smooth, especially in the intervals before and after zero earnings. This provides empirical evidence that EM does occur for loss avoidance.

Burgstahler and Dichev (1997) then estimate the frequency of EM for loss avoidance. Because the previous assumption about the shape of the distribution is rejected, they make a new assumption based on symmetry and incentives. The results suggest that 8 percent to 12 percent of firms with small earnings decreases manipulate earnings to show an increase instead and 30 percent to 44 percent of firms with negative earnings use manipulation to show positive earnings.

Degeorge et al (1999) analyses of distributions around three thresholds: zero earnings, sustained past performance (positive increase from previous year) and meeting analysts' consensus forecast. As in Burgstahler and Dichev (1997), unmanipulated earnings are assumed to be smooth, with densities between intervals changing gradually. This assumption is used to test the significance of expected discontinuities in the sample reported earnings. While some variables are quite homogenous along price per share centiles, such as analyst forecast error and change in earnings per share, the earnings per share variable itself is not. Degeorge et al

(1999) deal with this by checking whether results from the entire sample hold also if looking at several specific quartiles.

Degeorge et al (1999) find strong evidence that the distribution of changes in earnings per share and of analysts' forecast error are significantly kinked at the zero threshold. This indicates that zero changes is earnings per share and just meeting analyst forecasts happens significantly more than slight negative changes in earnings per share and missing analysts forecast by a small amount. They also find strong evidence on loss avoidance - there is a kink at zero earnings per share, but even larger kink in slightly positive earnings per share. This is possibly because managers care more to show positive earnings per share, rather than just break even. We would add that it is also possible that this is due to 'overshooting' - managers cannot precisely anticipate the size of the necessary manipulation to break even, and err on the safe side.

Using conditional distributions of earnings, Degeorge et al (1999) are able to distinguish a hierarchy in the thresholds: the positive earnings threshold is significant regardless of whether other thresholds are met, the sustained performance threshold is only significant if the positive earnings one is met, and the analyst forecast threshold is only significant if both other thresholds are met.

Even though Burgstahler and Dichev (1997) and Degeorge et al (1999) analyze a very similar time period and market, their results are not completely consistent. Both studies show evidence that densities at the threshold intervals are not as high as one would expect - there is a sharp discontinuity in the distribution which is apparent from the histograms as well as t-statistics. While both studies find 'kinks' that are consistent with EM, the former study finds a pronounced dip just before zero earnings and zero change in earnings intervals, compared to both adjacent intervals. The latter study finds that while there is a sharp increase in density after the zero EPS threshold, the density at the threshold interval just before zero is nearly the same as the density of the intervals immediately preceding it. All intervals to the left of zero EPS have nearly the same density. The differences in these results are likely due to the difference in defining the earnings variable - earnings scaled by opening market value versus earnings per share. In addition, the sample in Degeorge et al (1999) is smaller due to the additional constraint of availability of analyst earnings forecasts.

2.3Development of the Chinese stock market.

The history of the stock market in mainland China started in early 1990s not only to facilitate economic reform, but also to provide financing for loss-making state-owned-enterprises via privatization. But firms were rarely completely privatized after their listing. The government still retained controlling ownership in the form of non-tradable shares in the 1990s. What is worse, the prices of the tradable and the non-tradable share of the same company were not the same. The prices of non-tradable shares were determined on an individual exchange basis as they were not traded on the stock markets. They were based on fundamentals. The prices of the tradable shares were determined on the stock market. Because of this characteristically Chinese share structure, the listed firms did not put much effort in improving their performance since they did not have any direct incentives from the stock market.

In late 1990s, instead of further reforming the ownership structure, the China Securities Regulatory Commission (CSRC) set profit requirements for listed firms as an incentive for improving operational performance. For example, if any listed firms reported losses in three consecutive years, it would be marked with *ST (standing for 'special treatment'). Although it would not be delisted immediately, all of its tradable share transactions would be stopped. If it still cannot make a profit in the following six months, it would then be delisted.

Indeed, this policy encouraged firms to improve their performance, but at same time, also became incentive for EM. Since the delisting policy issued in 2000 and till the end of 2015, less than 20 firms were delisted, which is a very small proportion of over 3000 listed firms in total. A lot of *ST firms restarted their transaction in stock market by restructuring, or more precisely "Illicit Asset Stripping[1]", which means revaluing the underperforming assets and liabilities; and transferring them into unlisted entities. By doing this, only the profitable part remains listed.

Gradually realizing this share structure is not optimal for both growth and capital allocation, the government reduced its ownership starting in 2001. However, privatizing the non-tradable shares resulted in new shares suddenly flooding the capital market. Announcements of the conversion of non-tradable to tradable shares in 2001 caused large stock price drops for the company affected. Thus, the conversion was stopped until 2006, when a new split share structure reform (SSSR) of 2006 was proposed. In the SSSR, holders of non-tradable shares

[1] The term is translated from Chinese "资产剥离".

would transfer compensation to holders of tradable shares before the conversion occurs. This would significantly change the incentives of majority shareholders both in the transition period and afterwards.

Another remarkable event is in 2007 when new, substantially IFRS-convergent standards became mandatory for all listed firms. The government's efforts are to improve the accounting quality and to create a more transparent and internationally competitive market. One effort is issuing this IFRS-convergent accounting standard which hopefully can restrict the space for management's discretion in reported earnings.

2.4EM in the Chinese market

Previous EM research in the Chinese market is largely focused on the whether certain events mentioned in the above section would affect the size/trend of EM. However, it seems researchers only reach partial agreement when they apply the same models. Even if they study the same event based on data from same period, different models lead them to contradicting conclusions.

For example, Kuo et al (2014) study the effect of the SSSR on EM over the period 2002-2011 using models from Jones (1991) and Roychowdhury (2006). They conclude that Chinese companies engage in both accrual and real activities manipulation and that accrual manipulation decreases after the reform while real activities manipulation increases. The mean abnormal accruals are estimated as 0.1 percent of total assets. Conversely, Xiao (2015) analyse the SSSR over the same period using Dechow et al (1995) and conclude that accruals manipulation increased following the reform. The mean of abnormal total accruals is -0.36 percent of the assets.

Ho et al (2015) research the effects of new accounting standard adoption based on data from same period. Their modelling follows Dechow et al (1995) and Roychowdhury (2006). Similarly to Kuo et al (2014), the statistical tests would suggest that accrual manipulation is less prevalent and that there is an increase in the manipulation of real activities. However, their estimate of average abnormal accruals is actually closer to Xiao (2015) at -0.3 percent of assets.

Although Kuo et al (2014) and Ho et al (2015) reach the same conclusion about trends, the mean of abnormal total accruals differs from 0.1 percent to -0.3 percent of the assets, suggesting not only the different size but also the different directions of EM done in China. In the case of real activity manipulation, the average aggregated abnormal measures (adding up abnormal CFO, discretionary expense and production cost) are -0.3 percent and -0.56 percent of assets in respective studies.

Another topic interesting to scholars is EM in cross-listed companies. It is also accompanied with contradicting conclusions. Eng and Lin (2012) compare the EM of Chinese firms cross-listed in the US or Hong Kong with that of Chinese firms which are not cross-listed during 1993 to 2007, using Jones (1991) and Lang et al (2003) models. They find that while EM is present in both types of firms, the abnormal accruals of cross-listed firms are much larger. They report that the median of absolute abnormal accruals in cross-listed firms is 17.8 percent of lagged total assets versus 11.5 percent in non-cross listed firms.

Li et al (2014) analyse the effect of cross-listing using an alternative approach that focuses on the distributions of return on equity (ROE) around certain thresholds. a 3-year moving average ROE of 10 percent and ROE of at least 6 percent in the last three years was the preliminary requirement announced by CSRC for listing on a stock exchange. They compare the distributions of cross listed firms and non-cross listed firms with a wide sample spanning from 1990 to 2009. The conclusion is that EM in non-cross listed is more serious, as a result of more incentives generated with delisting policy introduced in the year 2000 in mainland China, which is contrary to Eng and Lin (2012).

In discussing earlier results, one issue requires extra attention: only in Ho et al (2015), authors mentioned the R-squared of the estimation models and it is surprising that the explanatory power of their models for the direction and magnitude of accruals is very weak (5 percent and 10 percent adjusted R-square respectively). Considering the nature of abnormal value reported by this regression approach, it is reasonable that the size of EM is overestimated because of the low predictive power for normal level.

2.5Research questions

Our aim is to find a reliable estimate of the size and frequency of EM in the Chinese market; thus the literatures reviewed raise several important questions for us.

Firstly, it is necessary to assess the existing models' fit for application in the Chinese market. Whether it is an accrual model or a real activities model, there has not been sufficient discussion so far whether these models (Jones, 1991; Roychowdhury, 2006) are appropriate in their original form.

Secondly, another important methodological question is regarding the estimation approach, which can be industry-specific or firm-specific. We show our initial results with each approach and discuss the implications. We then reach a decision about the more reliable approach, and proceed our further analysis with that one.

Thirdly, we need to consider how distribution analysis can be combined with the modelling of the size of EM. One shortcoming of the EM estimation as commonly done before is that the estimation models are partly based on manipulated observations. In other words, a model for detection is not constructed purely on unmanipulated observations, but on a mix of both manipulated and unmanipulated observations of the dependent variable. This design can, of course, jeopardize the ability of the model to detect manipulation. Some manipulation is normal to it! In order to overcome this shortcoming, we need to find a way to distinguish as well as possible between manipulated and unmanipulated observations, and then construct the estimation model only using unmanipulated observations. One way to distinguish between the two categories is by taking into consideration the incentives for EM, reflected in the thresholds of our distribution analysis.

The challenges and considerations as described above will guide our analysis throughout this paper.

3. Methodology

3.1Database

Our sample is based on the unique and comprehensive China Stock Market and Accounting Research (CSMAR) database. This database is jointly produced by GTA Information Technology Co. Ltd and some of the top universities in China to meet the needs of China, especially mainland, with respect to economic analysis and research. It consists of several parts, including Macroeconomics, Listed Firms' Fundamentals, Stock Market, Bond Market and Banking. For the listed firms fundamentals we use, the coverage is from 1990 (when China established its own stock market) onward. The data is organized by firm transaction code and fiscal year. The firm transaction code is the unique firm code by which the firm is identified on Chinese stock exchanges. The code is issued when a company becomes listed and it does not change as long as the company is listed. In addition to the company fundamentals database, there is another one which includes distinguishing information such as transaction code, address and industry code[2] and later used as our base to divide industry-year group.

We check the accuracy of the databases by randomly choosing some firm-year observations as well as some extreme observations (reported total assets equal to zero for instance) and comparing with the annual reports[3] of the firms. As far as we check, we do not find any error.

This database has limitations. Since it grabs fundamental data from financial statements in annual reports in China, the format of financial statement is uniform so sometimes the presented items are aggregated. For example, we are not able to get data such as gross PPE because Chinese listed firms do not report it: they report only net PPE obtained by subtracting all changes (depreciation, impairment) from gross PPE.

3.2Sample characteristics

Our sample consists of data from 2001 to 2015. Only firms listed on the main stock exchanges are included - the Shanghai Stock Exchange and the Shenzhen Main Board Stock Exchange (stock exchange code starting with 0 and 6 respectively). Firms listed in the Growth Enterprises

[2] This is coded in accordance with the Guidelines for the Industry Classification of Listed Companies issued by CSRC in 2012.

[3] Certain information can be found in http://www.cninfo.com.cn, which is the website appointed by CSRC for Chinese listed companies to disclose required information.

Market Board (code started with 3) are excluded because of different regulation requirements and delisting procedure. Financial institutions, brokers and insurance companies are also excluded because they have significantly different business models and regulation.

Moreover, we exclude some extreme observations. The criterion is extreme values of net income scaled by lagged assets. This ratio captures the influence of extreme events - events which would likely largely affect all accounting numbers used in the estimation of EM.

The challenge is to remove outliers so as to get a clear picture of the data but not to distort the data. We winsorize, i.e. remove the top 1 percent and bottom 1 percent of observations from the full sample of firm-years. With this approach, the 'extremeness' of observations is judged against what is 'normal' across the whole sample, not just across one accounting period. A single year with unusual events which affect the whole market could have more than 2 percent of extreme observations, and those would be removed with this approach.

Finally, we exclude some firm-years due to requirements of industry-specific estimation. Industry-years with fewer than ten observation should not be analysed, as there would be too few degrees of freedom, causing unreliable results.

After filtering the CSMAR full dataset in the above-described manner, the final sample covers 1,958 companies, which were listed over (all or part of) the period 2001-2015, resulting in a total of 18,066 unique firm-year observations. Firm observations from the year 2001 and 2002 are not included in this sample because at least one variable requires three consecutive years of data for computation, thus these are basis years. That variable is lagged changes in revenues scaled by lagged total assets. The first observation for that variable can be computed for the year 2003 and is based on inputs from 2001-2003. Thus, the possible number of observations per firm is 13. The median number of observations per firm is 11 and the mean number of observations per firm is 9.22.

Tables 1 and 2 show relevant descriptive statistics of the final sample. The mean revenue (*Rev*) is CNY 6.3 billion and the mean total assets (*A*) are CNY 9.2 billion. Some peculiar values of revenue and total assets require discussion. The minimum revenue figure suggests that total sales were negative for one firm-year. This figure is due to a firm-year observation when the firm was undergoing difficulties and restructuring. The surprising negative figure comes from returned goods and the corresponding adjustment to revenue.

Table 1. Firm characteristics (in CNY[4] million)

VARIABLES	Mean	Std. Dev.	Min	Max
Rev	6,261.246	25,882.110	-3.600	880,577.100
NI	313.026	1,588.806	-17,049.430	55,707.000
CFO	434.526	2,347.256	-22,369.900	72,864.000
TACC	-121.500	1,615.139	-37,000.670	27,875.170
A	9,163.711	32,815.090	0	1,074,905.000
EQ	3,473.772	10,522.270	-11,065.150	358,157.000
NetPPE	2,492.732	10,097.390	0	318,953.000
Inv	1,856.438	10,909.390	0	387,589.400

Variable definitions: Rev is net sales from the income statement; NI is net income from the income statement; CFO is cash flow from operations from the statement of cash flows; TACC is total accruals, which is the difference between NI and CFO; A and EQ are total assets and total equity respectively, from the balance sheet; NetPPE is the gross value of property, plant and equipment minus any depreciation and impairment; INV is inventory.

The minimum value of total assets of zero appears in three cases, where the firms are undergoing restructuring. But, it is reasonable to wonder - what is there to restructure if there are no assets? There is a question of whether after such a result, the new, restructured firm could be considered as a continuance of the old one. Formally, even though the book value of assets is reported as zero (in a single period only), and subsequent to the restructuring, the company changes its name and/or area of business, for legal purposes this is still the same company. Accordingly, we treat it as such. We do not find sufficient motivation to treat this case as bankruptcy of one company and creation of a new one, when the authorities, as well as the stock exchange, treat it as a single firm. In any case, this situation is very rare. Net PPE (*NetPPE*) of zero is also associated with observations right before a restructuring. Inventory (*Inv*) of zero occurs in service firms.

Table 2 shows relevant descriptive statistics related to the three types of EM. Discretionary expenses (*DISEXP*), production costs (*PRODC*) and total accruals (*TACC*) are defined analogously to Roychowdhury (2006):

[4] The exchange rate between CNY and USD is: CNY 1 = USD 0.14, May 10th, 2017

$$DISEXP_t = Administrative\ expenses_t + SG\&A_t + Marketing\ expenses_t + R\&D\ expenses_t$$

$$PRODC_t = COGS_t + \Delta INV_t$$

$$TACC_t = NI_t - CFO_t$$

In Table 2 there are also some extreme values that might be puzzling. For example, the minimum production costs scaled by lagged total assets are negative. This must be because production costs are negative, which can happen if the firm has a decrease in inventory that is larger than its COGS. This can only happen if the decrease in inventory is not due to sales but to some other reason, such as write-off of obsolete inventory. Thus, these values are unusual but not incorrect.

Table 2. Descriptive statistics

VARIABLES	Mean	Std. Dev.	Min	Max
Rev_t/A_{t-1}	0.800	0.745	-0.016	18.151
NI_t/A_{t-1}	0.039	0.069	-0.283	0.399
CFO_t/A_{t-1}	0.051	0.112	-2.075	1.364
$DISEXP_t/A_{t-1}$	0.097	0.089	-2.043	1.620
$PRODC_t/A_{t-1}$	0.676	0.734	-0.896	19.025
$TACC_t/A_{t-1}$	-0.012	0.111	-1.179	2.287

All variables are scaled by lagged assets.

In the case of discretionary expenses, it is even more puzzling to understand how discretionary expenses could be negative. We define expenses, as other studies do, to have a positive value if an expense occurs. In other words, if a firm spends CNY 10 million on administrative expenses, this would be recorded in the dataset as an expense of positive +10, not negative -10. The source of a negative expense in the dataset is due to a peculiarity of the Chinese accounting standards. In those standards, when a frim overestimates allowances for bad debt, in the subsequent period the adjustment is done to administrative expense by decreasing the expense. So, if the reversal of bad debt allowance is big enough, the firm-year could have a positive charge for administrative expenses on its income statement, as if it did not spend money on administrative expenses but actually made money. The source of the ‘income’ is the reversal of allowance for bad debt. This event would translate as a negative expense (a non-expense) in our dataset. Post 2007 and the application of IFRS-converging standards, such a reversal is reflected in a different account – impairment loss. In our sample, negative discretionary expenses appear in 30 firm-year observations due to negative administration expenses.

The initial information from Table 2 shows that production costs have the highest variability. This suggests that manipulation of production costs offers the most space for manipulation. It is harder for auditors to detect abnormalities if the variability of production costs is so large. Therefore, it could be easier for firms to manipulate without the manipulation being discovered by auditors.

3.3Existing estimation models

In the literature, the prevalent method of EM measurement relies on predicting normal levels of the EM measure (accruals or expenses), and then comparing these to actual levels, in order to determine when an abnormality (i.e. EM) occurs. The whole estimation of EM, therefore, is very sensitive to how accurately normal levels are predicted. For example, the prediction reasoning for normal levels of accruals is that they are determined by the economic circumstances faced by a firm in a particular industry and in a particular year. These circumstances are captured by PPE levels and year-on-year change in revenues. These explanatory variables are scaled by total assets.

After being introduced, these models are universally applied in future research, often without modification. However, an important question is “Are these models really good at predicting normal levels?” To be more specific, do revenues and PPE explain the variability of total accruals well? Are these variables good predictors of normal total accruals?

These questions are not addressed in research on the Chinese market, even though the fit on Chinese data seems like an important issue to consider. Table 3 illustrates this point, by applying Roychowdhury’s (2006) models on our final sample. We analyse the models’ predictive powers by applying panel data analysis with fixed effect. The panel variable is industry-year. That means that the data is organized in groups by industry-year and the model is applied separately to each group. The number of groups is equal to the total number of industry-years in the sample (which is 499). The average adjusted R-squared is part of the STATA output when running panel regressions with fixed effects.

Table 3. Roychowdhury (2006) models applied on final sample, industry-specific

VARIABLES	$DISEXP_t/A_{t-1}$	$PRODC_t/A_{t-1}$	$TACC_t/A_{t-1}$
$1/A_{t-1}$	2.164***	-0.277	0.503***
	(0.12)	(0.314)	(0.178)
Rev_t/A_{t-1}	0.0439***	0.935***	
	(0.000782)	(0.00253)	
$\Delta Rev_t/A_{t-1}$		0.0711***	0.0255***
		(0.00472)	(0.00223)
$\Delta Rev_{t-1}/A_{t-1}$		-0.00275	
		(0.00188)	
PPE_{t-1}/A_{t-1}			-0.0895***
			(0.00399)
Intercept	0.0606***	-0.0789***	0.0125***
	(0.000821)	(0.00219)	(0.00143)
Observations	18,066	18,066	18,066
Number of groups	499	499	499
Avg adj R^2	0.14	0.93	0.004
Avg adj R^2 reported in Roychowdhury (2006)	0.38	0.89	0.28

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

We see that the model proposed by Roychowdhury (2006) regarding discretionary expenses explains on average 38 percent of the variation of discretionary expenses for each industry-year in his original study of a US-based sample. This is not ideal if the whole inference of the size of earnings management relies on normal levels being accurately estimated. But the same model applied to the Chinese sample explains even less of the variability in discretionary expenses. If only 14 percent of the variation in normal levels is captured by the model, then how can any conclusions be based on the predicted normal levels of discretionary expenses?

In terms of the other measures, it seems that only in the case of production costs there is a sufficient predictability of normal levels. Both in the case of discretionary expenses and total accruals, the models are significantly less suited for Chinese data. A significant portion (over 50 percent!) of the normal variation in discretionary expenses or total accruals is not captured by the model and could then be misinterpreted as abnormal variation i.e. management manipulating earnings.

The main conclusion from the above discussion is that there is a need for a more accurate prediction of normal levels of EM measures better before anything can be said about the size of EM in the Chinese market.

3.4Proposed estimation and prediction models

In terms of model specification, our aim is to improve the models' predictive powers, except in the case of production costs where no changes are necessary. Ideally, the predictive power would be 80 percent or higher. It is worth mentioning that the same model changes are appropriate regardless of whether the industry-specific or firm-specific approach is used. In other words, we use the below models for predicting normal levels both when applying the industry-specific and firm-specific approach. Model changes are bolded in the formulas below.

Our estimation models are:

$$DISEXP_t/A_{t-1} = \beta_0 + \beta_1(1/A_{t-1}) + \beta_2(Rev_t/A_{t-1}) + \boldsymbol{\beta_3}(\boldsymbol{DISEXP_{t-1}/A_{t-1}}) + \varepsilon_t$$

$$\begin{aligned} PRODC_t/A_{t-1} \\ &= \beta_0 + \beta_1(1/A_{t-1}) + \beta_2(Rev_t/A_{t-1}) + \beta_3(\Delta Rev_t/A_{t-1}) \\ &+ \beta_4(\Delta Rev_{t-1}/A_{t-1}) + \varepsilon_t \end{aligned}$$

$$\begin{aligned} TACC_t/A_{t-1} &= \beta_0 + \beta_1(1/A_{t-1}) + \beta_2(\Delta Rev_t/A_{t-1}) + \beta_3(NetPPE_t/A_{t-1}) \\ &+ \boldsymbol{\beta_4}(\boldsymbol{\Delta TACC_t/A_{t-1}}) + \boldsymbol{\beta_5}(\boldsymbol{CFO_t/A_{t-1}}) + \varepsilon_t \end{aligned}$$

The specification of production costs is the same as in Roychowdhury (2006). This specification works well enough to accurately predict normal production costs. For the other two measures, we propose some changes in specification (shown in bold above).

Regarding discretionary expenses, following Dechow et al (1998) and Roychowdhury (2006) we assume that there is a linear relationship between contemporaneous revenues and expenses. This assumption means that, for example, when marketing expenses are higher than last period, also revenues are expected to be higher. We add one independent variable - lagged discretionary expenses. This is not driven by any causality but because the discretionary expense is usually comparable from one period to another. Thus, this period's revenues and last period's level of expenses should be indicative of this period's discretionary expenses.

Regarding total accruals, following Jones (1991), change in revenues and net PPE and used as proxies for growth and economic circumstances of the firm. We propose that two variables should be added. The addition of change in total accruals is motivated by the understanding that for certain kind of business, there are always benchmarks for the composition of assets. The accruals-to-assets ratio should be comparable and relatively stable over time. The addition of cash flow from operations is inspired by Dechow and Dichev (2002), who showed that saying the objective of accruals is to correct temporary matching problems with firm's underlying cash flows, then nondiscretionary accruals should be negatively correlated with contemporaneous cash flows and positively correlated with previous and following cash flows

Our prediction models are:

$$Normal\ DISEXP_t/A_{t-1} = \hat{\beta}_0 + \hat{\beta}_1(1/A_{t-1}) + \hat{\beta}_2(Rev_t/A_{t-1}) + \hat{\beta}_3(DISEXP_{t-1}/A_{t-1})$$

$$\begin{aligned} Normal\, PRODC_t/A_{t-1} &= \hat{\beta}_0 + \hat{\beta}_1(1/A_{t-1}) + \hat{\beta}_2(Rev_t/A_{t-1}) + \hat{\beta}_3(\Delta Rev_t/A_{t-1}) \\ &\quad + \hat{\beta}_4(\Delta Rev_{t-1}/A_{t-1}) \end{aligned}$$

$$\begin{aligned} Normal\ TACC_t/A_{t-1} &= \hat{\beta}_0 + \hat{\beta}_1(1/A_{t-1}) + \hat{\beta}_2(\Delta Rev_t/A_{t-1}) + \hat{\beta}_3(NetPPE_t/A_{t-1}) \\ &\quad + \hat{\beta}_4(\Delta TACC_t/A_{t-1}) + \hat{\beta}_5(CFO_t/A_{t-1}) \end{aligned}$$

Although these models *are specified the same as the estimation models,* the applications are different. The estimation models are used for obtaining the coefficients of the independent variables. In order to get the coefficients, we use actual observed values both for the dependent and independent variables. These coefficients are then used in a second step. The second step is to predict values for the dependent variables. These predictions are based on actual observed values of the independent variables only and multiplied with the previously obtained coefficients. Notice that the prediction models result in generating normal levels of each measure, for each firm-year. There are no error terms. There is a residual when the actual value is compared to the predicted value.

3.5Industry versus firm approach

The literature suggests two different estimation approaches: industry-specific and firm-specific. The choice of approach primarily affects how the estimation of coefficients is carried out. It does not affect how the prediction is done – the prediction per firm-year is always done by multiplying the obtained coefficient with contemporaneous values of the independent variables. As the two approaches have significant differences, it can be expected that they will generate significantly different conclusions about EM. That is why we discuss them in great detail in this study.

In the industry-specific approach, different sets of coefficients $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5)$ are generated for each industry-year and for each EM measure. In the final sample, there are 499 industry-years. This implies that we would have 499 different sets of coefficients for one EM measure, or 1,497 sets in total for all three measures. Each set of coefficients is obtained by running the estimation model on all firm-year observations within a certain industry-year. In the second step, the obtained coefficients are applied to each firm's specific numbers for that year (expenses, revenues, assets etc) to obtain the predicted normal level of the EM measure. The coefficients used for predicting for a particular firm change over time – every year a new set of coefficients is used. This implies that the relationships between the dependent and independent variables are assumed to vary over time but remain constant over economic and business conditions (which is characteristic for the industry-year).

Conversely, in the firm-specific approach, one set of coefficients per each measure is used for the prediction of the firm's normal level in all years it is in the sample. There are 1,246 firms in the sample and three EM measures, which implies that there are 3,738 sets of coefficients with this approach. While a firm's specific numbers (expenses, revenues, assets etc) change every year, the coefficients they are multiplied by are the same over time, for each measure. This implies that the relationships between the dependent variables and independent variables are expected to stay the same over time.

Here is a comparison of how discretionary expenses would be predicted for one firm with each approach, starting with industry-specific. First, the estimation model is applied to all firm observations in that industry-year (including the firm-year we are predicting for). Then, the coefficients from that regression are combined with the firm's revenues and lagged discretionary expenses for that year to obtain the predicted normal level.

With the firm-specific approach, the estimation model is applied on all observations from that firm. Whereas the industry-specific mode is applied in this study in the same way as in other studies, the firm-specific model is modified in that future observations are used as well. The logic is that the firm-specific relationship between EM measures and their predictors is as relevant at *t-1* as at *t+1* for the purpose of drawing inferences at *t*.

For example, if we need to find the discretionary expenses for 2005 of a firm which exists in the dataset during 2001-2009, then we estimate the relationship between scaled discretionary expenses, scaled revenues and scaled lagged discretionary expenses in the periods 2003-2009. The coefficients from that regression are then combined with the firm's revenues and lagged discretionary expenses for the prediction year to obtain the predicted normal discretionary expenses.

3.6Possible estimation bias

We expect to improve the predictability of discretionary expense and total accruals estimation regressions by adding lagged levels of these EM measures themselves as independent variables. This can possibly exacerbate the problem of bias due to estimation models being applied to manipulated observations. In previous studies, without these added variables, there is still bias because potentially manipulated observations of the dependent variable are used in estimation models. as discussed above, the estimation period and event period overlap with both estimation approaches i.e. the estimation of coefficients is done on all firm-years.

Now, with the addition of these variables, the result is that some observations of the independent variables are also manipulated. This introduces even more noise in the estimation of a constant relationship between normal levels of EM measures and the independent variables. We are aware of this problem, but decide that overall the gain from increased predictability is larger than the loss in accuracy from added noise. As will be shown in later sections, the ability of an estimation model to accurately predict normal levels has a very large effect on the obtained residuals. This is completely reasonable due to the nature of OLS regression.

3.7The relationship between EM measure and earnings

Before moving on to the estimates of abnormal discretionary expenses, production costs and total accruals an important question here is how to interpret the relationship between the abnormality and the effect on earnings.

In the case of discretionary expenses, there is an inverse relationship between the direction of the abnormality/residual and the impact on earnings. A negative residual appears when the actual amount of discretionary expenses is lower than the predicted amount, which happens when the firm spends less than usual i.e. postpones spending for another time. This results in upward earnings management.

$$Abnormal\ EM\ measure = Actual\ value - Predicted\ value$$

In the case of production costs, there is a direct relationship between the direction of the residual and the impact on earnings. Positive residuals imply upward earnings management. When actual production costs are larger than predicted, then the result is a positive residual. In this case, the firm has larger production costs for the purpose of spreading fixed costs over a larger number of produced units. The total amount of this additional expenditure can be seen on the cash flow statement, but not all of it will appear in the period's operating expense. The result is a lower COGS and higher net income.

Total accruals are defined as the difference between net income and cash flow from operations and it is clear that if the cash flow is unchanged then total accruals and net income will change in the same direction. This can be explained by not only arithmetic relationship, but also the current accounting conceptual framework that defines income as a result of increase in the assets or decrease in liabilities and in opposite the loss as a result of decrease in assets or increase in liabilities.

Positive total accruals result when net income is larger than cash flow from operations. For instance, if the company generates a lot of sales but most customers pay with credit, then the result is lower cash flow from sales but large accounts receivable. Positive total accruals should be viewed as assets and hence they should be associated with upward earnings management.

When cash flow from operations is larger than net income, this can only be due to negative accruals, such as allowances for bad sales, depreciation and others or due to sufficiently large interest expenses[5]. Almost always, it is a combination of both.

However, for individual accrual accounts, the relationship can be opposite, depending on whether the account is income-increasing or income-decreasing. For example, a depreciation charge which is larger than predicted would result both in a positive residual and in lower earnings i.e. downward earnings management.

[5] Cash flow from operations in our dataset is the outcome of applying indirect method.

4. Empirical analysis

4.1Proposed Models Fit - Alternative Approaches

Panel A of Table 4 shows the regression results from using our models on the final sample, using the industry-specific approach. Considering the degree of freedom and in order to decrease the randomness of the estimations, we require a minimum number of observations per industry-year. Thus, we exclude from our sample industry-years which have fewer than ten observations i.e. industries which consist of fewer than 10 firms in a particular year, as mentioned in Section 3.2. The industry approach is used on that final sample. The number of groups in Panel A of Table 4 shows that this panel analysis was done for 499 groups, which is to say 499 industry-years since the regression is industry-specific.

Table 4. Our models applied to final sample, two alternative estimation approaches

	Panel A. Industry-specific			Panel B. Firm-specific		
VARIABLES	$DISEXP_t/A_{t-1}$	$PRODC_t/A_{t-1}$	$TACC_t/A_{t-1}$	$DISEXP_t/A_{t-1}$	$PRODC_t/A_{t-1}$	$TACC_t/A_{t-1}$
$1/A_{t-1}$	-0.665***	-0.277	0.641***	0.397***	-0.970**	0.821***
	(0.11)	(0.314)	(0.105)	(0.1)	(0.409)	(0.131)
Rev_t/A_{t-1}	0.0342***	0.935***		0.0476***	0.935***	
	(0.000691)	(0.00253)		(0.000778)	(0.00473)	
$\Delta Rev_t/A_{t-1}$		0.0711***	0.0452***		0.0658***	0.0432***
		(0.00472)	(0.00129)		(0.00569)	(0.00135)
$\Delta Rev_{t-1}/A_{t-1}$		-0.00275			-0.00243	
		(0.00188)			(0.00186)	
$DISEXP_{t-1}/A_{t-1}$	0.364***			0.170***		
	(0.00481)			(0.00405)		
PPE_{t-1}/A_{t-1}			-0.0132***			0.00488*
			(0.00233)			(0.00275)
$\Delta TACC,t/A_{t-1}$			0.0274***			0.0313***
			(0)			(0.00151)

CFO,t/A,t-1			-0.792***			-0.861***
			(0.00447)			(0.00477)
Intercept	0.0382***	-0.0789***	0.0274***	0.0423***	-0.0720***	0.0216***
	(0.000772)	(0.00219)	(0.00083)	(0.000745)	(0.00359)	(0.000946)
Observations	18,064	18,064	18,064	15,501	15,501	15,501
Number of groups	499	499	499	1,246	1,246	1,246
Avg adj R^2	0.350	0.931	0.671	0.247	0.849	0.716

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Panel B of Table 4 shows the specification and predictive ability of our proposed estimation models with firm-specific approach. The firm-specific sample does not include firms which have fewer than ten observations, for the same estimation requirements as in the industry-specific approach. The number of groups in Table 4 Panel B shows that this panel analysis was done for 1,246 groups, which is to say 1,246 firms instead of the 1,958 firms included in the final sample. The difference is due to the exclusion of firms which have less than ten observations.

The predictability of our models is still not ideal, but at least it is close to or higher than the respective level in Roychowdhury (2006). Thus, the gap due to data and goodness of fit has been bridged. But the overall gap to high (over 80 percent) prediction power is still elusive. Especially in the case of discretionary expenses, until further improvements are found, our solution is to interpret the results based on this EM measure very cautiously. Regarding total accruals, the improved model has significantly higher prediction power than the Jones (1991) model applied to US data.

In addition, the signs of the coefficients of the added variables are reasonable. We observe a negative coefficient on cash flow from operations, as Dechow and Dichev (2002) predict. Regarding lagged discretionary expenses and change in total accruals, it is reasonable that the coefficients are positive because there is a natural increasing progression in the size of expenses and accruals as the business grows.

4.2 Size of EM in Chinese sample

Table 5 shows some initial results about the size of earnings management in the Chinese market. Results with both estimation approaches are presented for the sake of comparison, with the aim to select a final approach as a result. The mean values of the EM measures in Table 5 show the average size of the detected abnormalities in relation to lagged assets. The size of EM is defined in that way because the EM measures themselves are expressed in relation to total assets. Combining Table 2 and Table 5, we could say that the amount of discretionary expenses is equivalent in size to 9.74 percent of lagged assets and that the abnormality/manipulation in discretionary expenses is equivalent in size to 0.000000000416 percent of lagged assets, on average.

Judging only by the mean size of detected manipulations suggests that, on average, EM in the Chinese stock market is negligible, regardless of the chosen approach. However, this fails to take into account the dispersion of the measures and more importantly, the fact that large detected manipulations in opposite directions would result in a mean close to zero while still EM would be significant. To get a better insight into the size of the average manipulation, we look at the mean of the absolute values of the detected abnormalities. The absolute mean does not tell us the direction of the manipulation, but it does tell us the average size. On average, discretionary expenses equivalent to the size of 1.8 percent of lagged assets are either overstated or postponed. From this information, we cannot tell what percentage of discretionary expenses are manipulated.

Table 5. The size of EM in final sample, two alternative estimation approaches

	Panel A. Industry-specific			Panel B. Firm-specific		
	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$
Mean	-4.68E-12	-3.9E-12	-3.4E-12	3.34E-12	2.48E-11	-6.83E-12
Std Dev	0.031	0.137	0.042	0.019	0.084	0.026
Absolute mean	0.018	0.079	0.030	0.010	0.044	0.017

What is striking is that even with the firm-specific approach, which has firm-tailored coefficients, the average EM is significant, ranging between 1.0 percent and 4.4 percent of lagged assets. Focusing on the industry-approach, the range of average manipulation is 1.8 percent to 7.9 percent of lagged assets. This is before the aggregation of EM measures. It is important to keep in mind that a firm can use any one of the EM channels to manage earnings, as well as a combination of them.

At this point in the analysis, we need to choose which approach is more appropriate for further analysis. The choice is necessary because the difference between the approaches is material. Comparing the residuals generated with both approaches, these approaches do not even generate residuals with the same sign 37 percent of the time. The use of the firm-specific approach is troubling for a number of reasons. Firstly, the issue of degrees of freedom cannot be solved easily. The degrees of freedom are important for coefficient predictions - few degrees of freedom can result in coefficients which are not sensible but which generate low or no residuals. Even if the sample period is extended, it is not obvious that a firm's business model can remain the same over 20+ years of operations, or similar enough so that normal levels can be estimated correctly.

It is far easier to organize firms by industry such that the firms are similar enough but also the number of firm-years is larger than 20. Industry-years naturally have more observations. It is more reasonable to expect that there is a stable relationship between the dependent and independent variables within companies that certainly operate in the same area. This certainty is lacking with the firm approach, as a firm can change its operations significantly over time. In addition, the industry approach estimates new coefficients each year. Within one year, macroeconomic variations should not be significant, but even if they are - the whole industry would be experiencing them. The firm-approach does not control for macroeconomic variations over time.

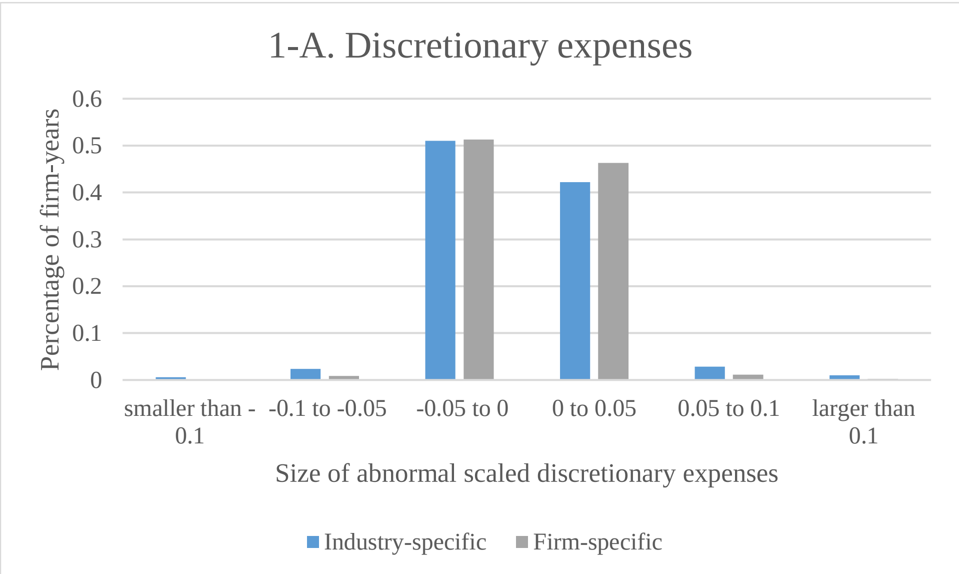
Ultimately, the firm-specific approach raises significant risk that the prediction models are overfitted. For instance, using this model on a company with frequent EM would likely result in small residuals - because the dependent variable observations which are themselves manipulated are used in the estimation of the coefficients. The industry-specific approach does suffer from a similar problem, but to a much smaller degree. It is less likely that a whole industry would manipulate earnings in the same way and in the same year than it is that a company would do that over time.

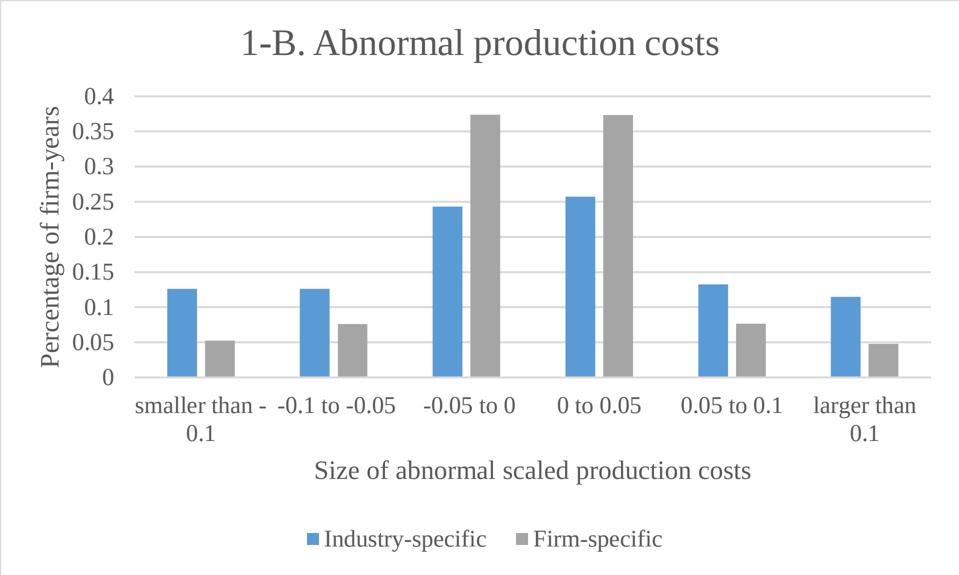
We proceed with the industry approach for further analysis. The first question is regarding the direction of earnings management. Are overstatements of earnings more common than understatements? Because of the significant difference between the mean and absolute mean in Table 5, it can be deduced that negative and positive residuals cancel each other off on average. While Table 5 is useful in showing the average size of the manipulation, it does not show the direction. The direction can be seen from the distributions of the abnormal measures which are shown in Figure 1.

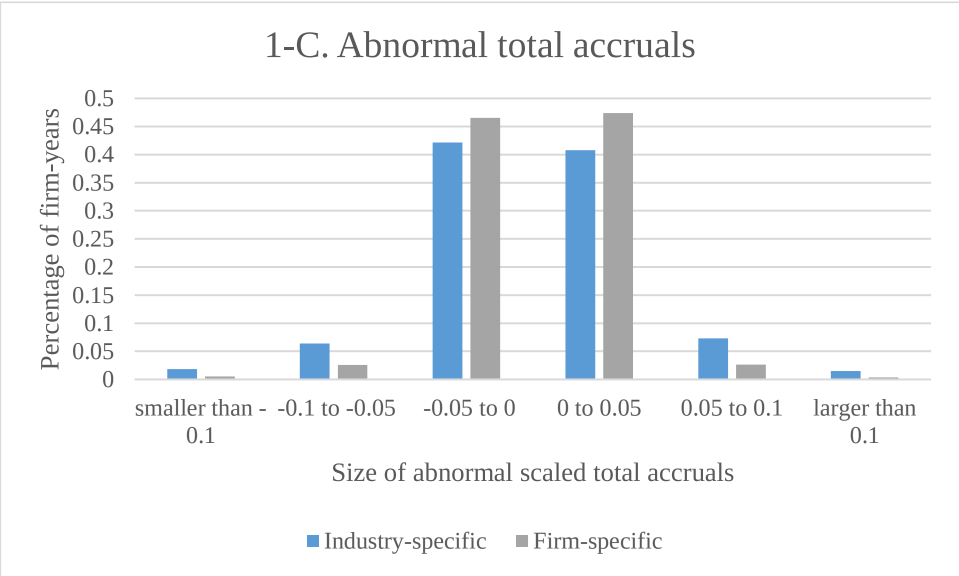
This figure shows the frequency of large abnormalities for the three measures of EM. It is important to notice that the interval widths are constructed so as to show the frequency of extreme large understatements and overstatements. Not all intervals are of the same width. For example, the 6th interval shown in Figure 1-A shows the frequency of overstatement of discretionary expenses which are equivalent in size to 10 percent of lagged total assets or more. The number of firm-years with such overstatements is about 180 and counts for nearly 1 percent of the whole final sample of observations.

Looking at the industry-approach, it is interesting to note that upward EM is almost equally likely as downward EM. This is consistent across measures, with the exception of discretionary expense manipulations larger than 10 percent of scaled assets. In that case, we see that large understatements are almost twice as likely as overstatements, regardless of approach.

Figure. 1. Distributions of abnormal EM measures







The above analysis raises several questions. Firstly, whether there is a way to ameliorate the shortcoming of coefficients being based on manipulated values of the dependent variables. This issue is addressed in the next two sections. Second, do the models proposed here generate significantly different results compared to Roychowdhury (2006), when applied to Chinese data. This question are addressed in Table 6 in Section 4.4. Finally, we analyse what the aggregation of EM measures implies for the size and frequency of EM in the Chinese market in Section 4.5. All of these analyses are done using the industry-specific approach.

4.3Distribution analysis

In order to decrease the estimation bias from using manipulated observations of dependent variables in the estimation process, it is necessary to categorize firm-years in two groups: one group which is less likely to manipulate earnings and one group which is more likely to so. As Burgstahler and Dichev (1997) and Degeorge et al (1999) show, the distributions of scaled earnings and scaled changes in earnings can be used for making the group distinction.

If those distributions are used to discuss EM and loss avoidance, then the choices regarding interval widths and outliers are crucial. This is because ‘suspect’ firm-years are those that are found in a specific interval - the first interval to the right of the zero threshold. Those firm-years came very close to reporting zero earnings or slightly negative earnings. The size of the interval puts a number on ‘very close’. When the difference between loss and profit is small, managers have both the incentives and opportunity to use managerial discretion to get over the zero threshold. The dependence of the result on such choices is undoubtedly the biggest shortcoming to this approach. Degeorge et al (1999) use the following formula to determine the interval width:

$$interval\ width = 2 * (Interquartile\ Range) * (n.of\ observations^{-1/3})$$

This formula applied to our sample results in an interval width of 0.0043. This would imply that a ‘small’ loss for our purposes is defined as earnings scaled by lagged assets of -0.43 percent to zero. By comparison, the bin width in Roychowdhury (2006) is 0.005 or small loss equivalent to 0.5 percent of assets. Burgstahler and Dichev (1997) use a bin width of 0.0025 while they have almost 4 times more observations than us.

Figure 2. Histogram of scaled net income with 0.0043 bin width using our final sample

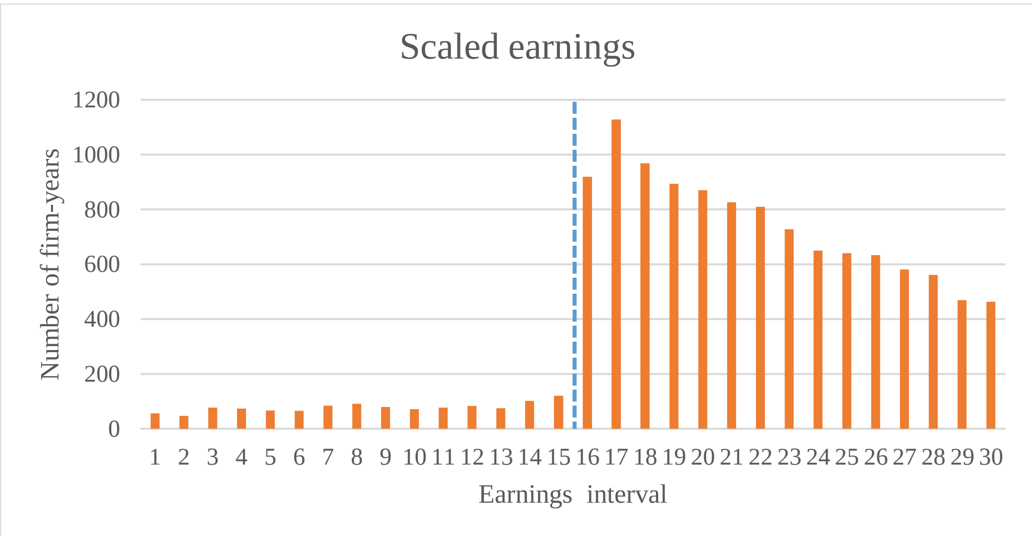


Figure 2 shows the distribution of net income scaled by lagged total assets. For clarity, it is truncated in the range of -0.0645 to 0.0645. Within this range, we construct 30 intervals with width of 0.0043 or equivalently 0.43 percent of total assets. The intervals are centred on the zero threshold, which is indicated with the dashed line. Interval 15 contains firm-years with scaled earnings in the range of -0.43 percent of lagged assets *up to but not including* 0.00 percent of lagged assets. Interval 16 contains firm-years with scaled earnings in the range of 0.00 percent of assets to 0.43 percent of assets, and so forth.

The result is more similar to Degeorge et al (1999) than to Burgstahler and Dichev (1997) in that there is no dip in the distribution at the 15th interval but there is a significant relative increase in the density of the 16th interval. However, unlike Degeorge et al (1999), here the distribution of scaled earnings is strongly weighted toward positive earnings. This result is largely due to the delisting regulation, which has a direct and an indirect effect on the distribution. The direct effect is that unprofitable firms are filtered out of the sample. The survivorship bias can be expected to be more significant compared to other markets where such regulation does not exist. The indirect effect is related to EM. In this market, firms have a significant incentive to avoid losses due to the delisting regulation. Therefore, small losses are more infrequent than small gains.

The dual effect of delisting regulation is unique for the Chinese market. In Burgstahler and Dichev (1997) and Degeorge et al (1999), it is assumed that a kink or discontinuity in the distribution of scaled earnings can only be caused by EM. Here, such an assumption is not appropriate. Therefore, their procedure for testing the existence of a discontinuity cannot be applied here either. Their testing procedure is based on the assumption of smoothness of earnings, where the actual density at the interval to the right of zero (interval 16 in Figure 2) is compared to the implied density given that earnings are smooth. The implied density is the average of the densities of the preceding and of the succeeding intervals (intervals 15 and 17 in Figure 2). Here such an approach is clearly not appropriate, since earnings cannot be expected to be smooth, not only because of EM, but also because of another cause.

This presents a problem. On the one hand, even though there is an apparent discontinuity in the distribution, it cannot be attributed certainly to EM. However, there is a strong reason to believe that firm-years with small profits have used EM – they have the strongest incentive to do so. On the other hand, suspect firm-years need to be identified in some way, so that the estimation of coefficients and subsequently the normal levels of EM measures are not based on manipulated observations, as much as possible.

One solution to this problem is to treat firm-years in the 16th interval as suspicious, even though a formal test has not been conducted to suggest that they are suspicious. But, the result should not be taken blindly. Rather, that result can be compared with an alternative procedure. If two different procedures yield similar conclusions about EM in the Chinese market, the results are more reliable.

The alternative procedure is to choose suspicious firm-years based on the distribution of scaled *changes* in earnings, and run the models accordingly. The incentive for avoiding negative changes in earnings is the possible stock market reaction.

Figure 3. Histogram of scaled change in net income with 0.0024 bin width using our final sample

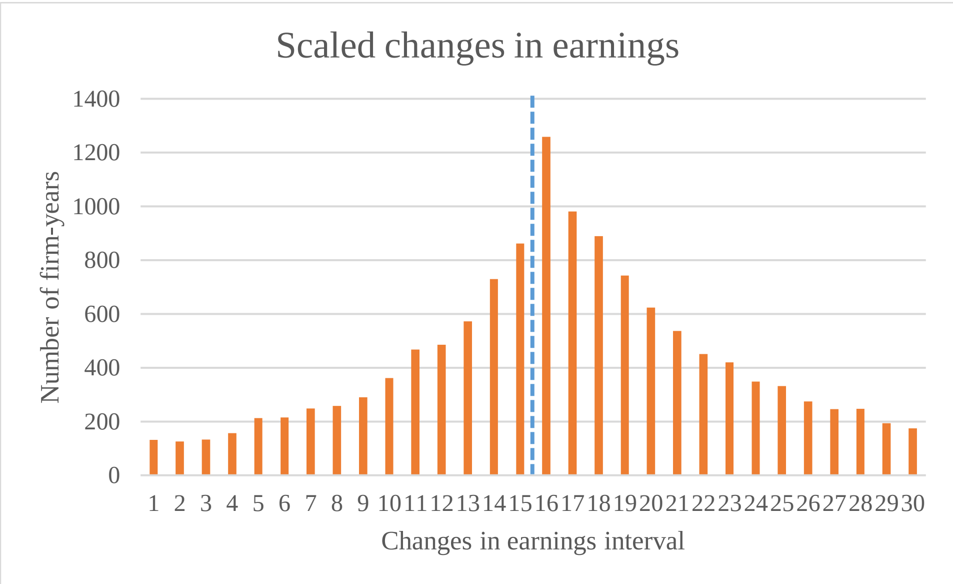


Figure 3 shows the distribution of change in net income scaled by lagged assets, truncated within the range from -0.0375 to 0.0375. Similarly as in Figure 2, we construct 30 intervals centred on the zero point, each with the equal width of 0.0025 or 0.25 percent of total assets. There is no pronounced dip at the 15th interval as Burgstahler and Dichev (1997) find in their study, but there is a pronounced jump at the 15th interval. We see that it is much more symmetrical than the distribution of scaled earnings. This suggests that the direct effect of the delisting policy is less pervasive. However, the shape of the distribution is still affected by two factors to some extent – EM and delisting. Thus, again, testing for smoothness at the zero threshold would not give a reliable result.

4.4Size of EM with suspect firm-years

Table 6 shows the manipulation as a result of applying different regression models with defined suspicious years. Firm-years with the net income between 0-0.42 percent of lagged assets and define them as "suspicious" because they are close to reporting a small loss or zero earnings. That is the kind of situation where the incentives for EM are significant. We run the estimation models for each industry-year using those "unsuspicious" observations. Then, the estimated coefficients are used in determining the normal levels for both suspicious and unsuspicious firm-years. In panel A, the estimation and prediction models are specified as in section 3.4. In panel B, the models from Roychowdhury (2006) are used. Both panels show the size of EM detected in the final sample and follow industry-specific estimation.

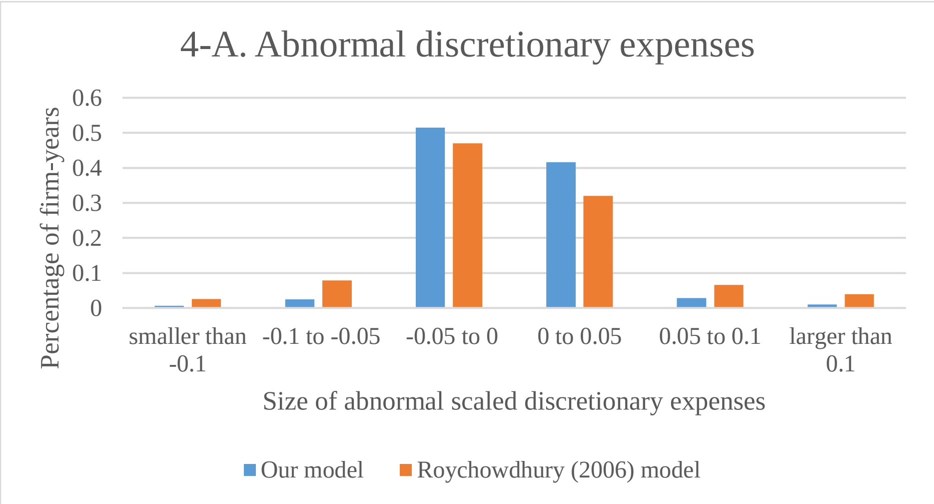
By comparing the results showing in table 6, it is clear that the estimated size of EM is very sensitive to how accurately normal levels are predicted. Due to their lower prediction power to the Chinese market data, applying models of Roychowdhury (2006) shows much more significant manipulation while also the standard deviations are larger.

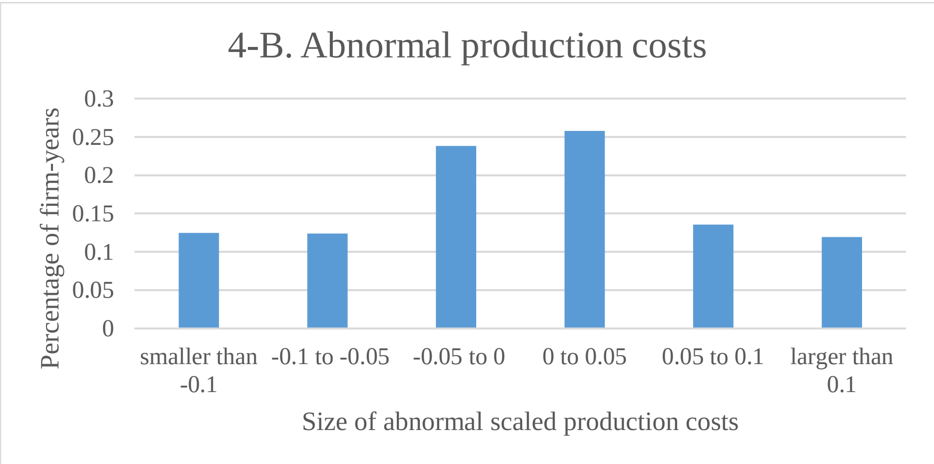
Table 6. Size of EM in the final sample with alternative models

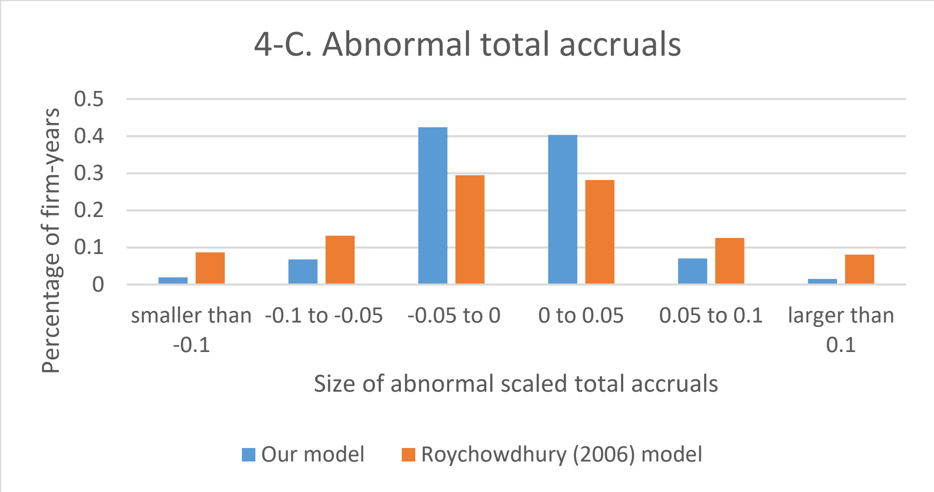
	Panel A. Our models			Panel B. Roychowdhury (2006) models		
	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$
Mean	-0.00050	0.00212	-0.00059	-0.00073	0.00212	0.00020
Std Dev	0.032	0.162	0.047	0.058	0.162	0.102
Absolute mean	0.018	0.080	0.030	0.036	0.080	0.060

The distributions of manipulations generated by our models and Roychowdhury's (2006) models are presented in Figure 4. These figures can be used in two ways – to compare the implications of using different models and to discuss the direction of EM. With respect to discretionary expenses and total accruals, the application of Roychowdhury (2006) model results in fatter tails, with abnormalities over 10 percent being much more frequent. Small manipulations are far less frequent, compared to the results with our estimation models. These differences are obviously a consequence of the difference in the prediction power. Where the prediction power is lower, the actual and predicted values are likely to be more different, generating larger residuals. These figures highlight just how much model specification affects the final result.

Figure 4. Distributions of abnormal EM measures with suspect firm-years due to loss avoidance







Focusing on results from our models, we turn to discussing the direction of EM. Fifty-five percent of discretionary expense residuals are negative. As discussed in Section 3.7, this implies overstatement of earnings. Focusing only on large manipulations of over 5 percent of lagged assets, understatements of earnings are slightly more frequent than overstatements.

The figure of abnormal production costs does not show a comparison between models because for that measure we do not change the Roychowdhury (2006) model, so actually there is no difference. The figure is presented to discuss the direction of EM. We consistently find the frequently large abnormal production costs. This means that production costs are used the most as the vehicle for EM. As mentioned previously, this might be because they are very variable and also under the discretion of management. It might be difficult for auditors to rule manipulations as fraud. Indeed as Roychowdhury's (2006) definition of real-activities management implies, manipulating production costs does not result in fraud. But it does result in over- or underproduction, which are done to mislead users of financial statements about the true underlying production costs. Large positive residuals (over 5 percent of lagged assets) are slightly more frequent than large negative residuals i.e. earnings overstatements are slightly more frequent.

With respect to total accruals, negative residuals are slightly more frequent than positive residuals. This trend also applies to large residuals over 5 percent of lagged assets. It implies that earnings understatements are slightly more frequent.

Three main conclusions can be drawn so far. Firstly, all three types of EM are used approximately equally often to understate earnings as to overstate earnings. Secondly, large manipulations are by far most frequent in production costs. Thirdly, we do not know whether overstatements or understatements are more frequent in the Chinese market overall - that depends on how these three measures are jointly used.

4.5Real activities manipulation

The final puzzle of the size of EM in the Chinese market is regarding the understanding that different EM techniques are not used in isolation. To be able to say definitively how often and how much Chinese reported earnings are manipulated on aggregate, we must consider the above EM measures together. However, due to the construction of the measures, we cannot easily aggregate abnormal production costs and abnormal total accruals. It is likely that there is some overlap in what these measures capture. Aggregating abnormal production costs and abnormal discretionary expenses gives the size and frequency of real activities manipulation (RAM). This measure is constructed by using the residuals from applying our models with the industry-specific estimation approach and with designation of suspicious firm years which report a small loss. That is, using the discretionary expenses and production costs residuals presented in Figure 4. The formula of RAM is:

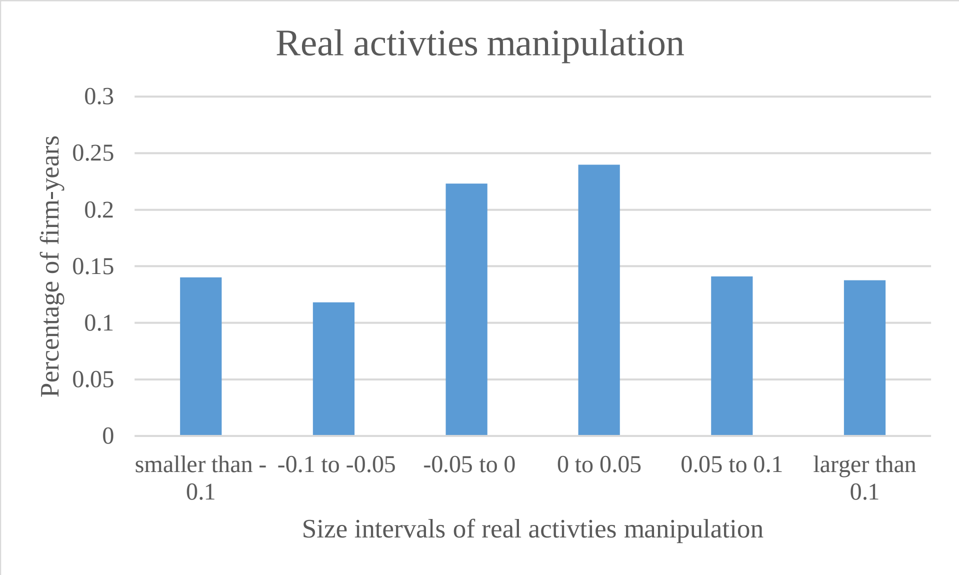
$$\text{RAM} = \text{Abnormal } PRODC_t/A_{t-1} - \text{Abnormal } DISEXP_t/A_{t-1}$$

The aggregated measure RAM is largely similar to abnormal production costs, which is not surprising since those residuals are by far the largest of the three. The summary statistics are shown in Table 7. RAM has an even higher absolute mean. This is also reasonable, since both types of real activities EM show more frequent overestimations of earnings rather than underestimations. On average, real activities manipulation of earnings is equivalent in size to 8.7 percent of lagged total assets. The distribution of RAM is presented in Figure 5. Manipulations larger than 10 percent of lagged total assets occur 27.9 percent of the time.

Table 7. Size of real activities EM

	RAM
Mean	0.002619
Std Dev	0.169
Absolute mean	0.087

Figure 5. Real activities manipulation



4.6EM size and frequency with alternative assumption about suspect firm-years

This section represents a check to the results from Sections 4.3 and 4.4, which are based on the assumption that firm-years with small losses manage earnings more than other firm-years. Because this assumption cannot be tested, we test the reasonableness of the result by applying a different assumption and observing whether this changes the underlying conclusions. This check is not intended to test/confirm any assumption about suspicious firm-years – it is intended to check whether the implications about size and frequency of EM still hold. In other words, are our results based on an untested assumption sensitive to a change in that assumption?

The change in assumption is about which firm-years are assumed to have more EM than the rest of the sample. In this section, we assume those are firm-years will small increases in earnings, as shown in Figure 3. We redo the analysis from the previous two sections and compare the results. The residuals generated under the new assumption are quite similar to those reported in Section 4.3. Table 8 is analogous to Panel A of Table 5, but also includes summary statistics of RAM2. RAM2 is analogous to RAM with the difference being that residuals of production costs and discretionary expenses are those summarized in Table 8 instead of Table 5. The absolute means of all measures are identical.

Looking at the degree of correspondence between residuals generated with alternative suspicious-years assumption, we see that they are highly correlated. The correlation between the abnormal discretionary expenses generated under each assumption is 69.6 percent. The correlation between abnormal production costs is 76.8 percent and between abnormal total accruals 77.9 percent. The distributions of the residuals generated under the second assumption about suspect firm-years lead to the same conclusions about the frequency of understatements and overstatements (shown in Appendix D).

Regarding the aggregated measure RAM, creating an analogous RAM2 leads to similar conclusions. RAM2 has a smaller mean but larger standard deviation than RAM. The absolute mean is the same. The correlation between RAM and RAM2 is 77.2 percent. The distribution of RAM2 is practically the same as that of RAM (shown in Appendix D).

Table 8. EM with alternative suspect years, industry-specific, our model

	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$	RAM2
Mean	4.89E-05	0.000467	-0.00029	0.000419
St dev	0.044	0.152	0.047	0.174
Absolute mean	0.018	0.080	0.030	0.087

The above results suggest that there is no difference in results under two different assumptions about suspicious firm-years. There are three possible reasons. Firstly, the identified thresholds do separate successfully suspicious firm-years and in addition to that, the two thresholds are equally important to Chinese firms. Accordingly, the two thresholds inspire EM in the same way. Alternatively, it could be a coincidence that we obtain the same result. If it is a coincidence, we could expect that designating a sample of random firm-years as suspicious firm years would generate a different result.

To rule out coincidence, STATA is used for selecting 1,000 random firm-years to be considered suspicious. This is approximately the same number of suspicious firm-years under both assumptions discussed above. We then run the estimation and prediction models in the same way as before. The resulting residuals are actually very similar. The absolute mean for all three measures is again identical. This result is shown Table 9.

Table 9. Size of EM with 1,000 random suspicious firm-years

	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$
Mean	-0.00017	0.000639	6.43E-05
Std Dev	0.032	0.140	0.046
Absolute mean	0.018	0.080	0.030

The fact that the absolute mean and distributions do not change significantly under the three assumptions of suspicious years can hardly be a coincidence. At this point we have to consider a third reason for getting similar results - the three uses of suspicious years have something in common. All three designations of suspicious years generate 900-1200 suspicious firm-years (approximately 6 percent of the final sample). Changing the residuals for this percentage of the final sample probably does not have a very significant effect on the sample as a whole simply because the proportion of changed residuals is too small.

To rule out the third reason as the cause of stable results, STATA is used again but this time to generate 2,000 random firm-years as suspicious (approximately 11 percent of the final sample). The summary statistics are shown in Appendix E. The absolute means of all three measures are still very similar but increase very slightly. This can be expected because we have un-specific coefficients being applied to a larger proportion of the sample. Here we encounter a similar relationship as when analysing the firm-specific approach. The more specific coefficients are, the smaller the residuals and vice versa. And how specific coefficients are depends on the degree of overlap between the estimation period and the event period. The bigger the number of suspicious firm years, the larger the residuals and so the estimate of EM will be.

In terms of solving the bias introduced by overlapping estimation and event period in the common application of the industry-specific method, it seems that thresholds are not an effective way to address this bias. This is because the result with threshold suspicious years and without is the same – distributions in Figures 1 and 4 relating to our model and industry-specific estimation are practically the same. Similarly, the absolute mean and standard deviations in Panel A of Tables 5 and 6 and in Table 8 and Table 9 are highly similar. On the other hand, we should not increase the interval width in threshold analysis just because the proportion of suspicious firm-years is small, as it is now. The interval width should always be based on what can be defined as a small loss or small decrease.

If the stable estimates of abnormal EM (in absolute terms) do not change because of few suspicious years or coincidence, then we could be capturing a result which is not related to thresholds but rather a general characteristic of the market. Even with the presence of the bias due to overlapping estimation and event period, we believe that the estimation of size and frequency of EM is reliable in this case. We consistently get a stable result, regardless of whether the percentage of suspicious years is zero, 6 percent or 11 percent (approximately).

5. Conclusion

The first conclusion from our analysis is that the choice of estimation method significantly affects the estimation result. It is not certain whether or not that is due to the specificity of coefficients, which is an inherent characteristic of the firm-specific approach. It is likely that the primary cause is the lack of degrees of freedom. This problem cannot be solved easily. Therefore, the industry-specific approach is preferable, despite the shortcoming that estimation period and event period normally coincide.

However, if this shortcoming is addressed by using threshold analysis to separate estimation firm-years and event-firm-years, then there is no significant effect on the final result. Both the summary statistics and distributions of residuals are relatively similar irrespective of whether the industry approach is applied uniformly on the whole sample or there is some division of the estimation and event period. This is a very surprising result. Ideally, we expect that the estimation models should be estimated on observations where the dependent variable is surely not manipulated; then the coefficients should be applied to event years where it is manipulated, in order to measure the manipulation. We do this by dividing firm-years in a group which is likely to have manipulated the EM variable versus a group which is less likely to do so. The likelihood is not estimated statistically – it is based on consideration of incentives. However, the division based on this assumption results in a small percentage of the total sample to be designated as suspicious. Thus, even though the abnormal levels for these firm-years changes significantly, the impact on the conclusions for the whole sample is minimal.

We find large earnings manipulation in the Chinese market. The average size of the manipulation *on an individual EM measure basis* is equivalent in size to1.8-7.9 percent of lagged total assets. Large manipulations occur frequently – between 6.8 and 50.4 percent of the time, depending on the type of manipulation! The largest manipulation is via production costs, followed by accruals. Does this mean that that reported earnings are on average misstated by 1.8-7.9 percent? Not necessarily, because the overall manipulation per firm-year depends on how the combination of EM types is used. All three EM measures considered here are not easily aggregated – there is probably some overlap between total accruals and production costs. The size of real activities earnings management is 8.7 percent on average. Large manipulations equivalent to over 5 percent of lagged total assets occur 54 percent of the time. Even though we cannot easily combine total accruals and the aggregated measure, a consideration of the numbers involved suggests that EM is probably considerable. Abnormal total accruals would have to be very large to offset real activities management, and we can see that they are not.

Another conclusion of our study is that the explanatory power of the estimation models of EM measures significantly affects the result. As can be seen from the comparison in Table 6, the absolute mean of abnormal EM measures is very different when different models are applied to the same dataset and in the same way (industry-specific estimation). When the R-squared of the estimation doubles, the size of the absolute mean halves. This is a statistical consequence, so this conclusion is relevant for all studies of EM which rely on regression analysis of EM measures.

Appendix

A. Jones (1991) model

Estimation model: $NA_\tau = \alpha_1(1/A_{\tau-1}) + \alpha_2(\Delta REV_\tau) + \alpha_3(PPE_\tau)$

Where

$NA_\tau = normal\ level\ accruals\ in\ year\ \tau$;

$\Delta REV_\tau = revenues\ in\ year\ \tau\ less\ revenues\ in\ year\ \tau - 1\ scaled\ by\ total\ assets\ at\ \tau - 1;$

$PPE_\tau = ross\ property\ plant\ and\ equipment\ in\ year\ \tau\ scaled\ by\ total\ assets\ at\ \tau - 1;$

$A_{\tau-1} = total\ assets\ at\ \tau - 1;\ and$

$\alpha_1, \alpha_2, \alpha_3 = firm - specific\ coefficients$

Prediction model: $TA_t = a_1(1/A_{t-1}) + a_2(\Delta REV_t) + a_3(PPE_t) + v_t$

Where

$a_1, a_2\ and\ a_3\ denotes\ the\ OLS\ estimations\ of\ \alpha_1, \alpha_2\ and\ \alpha_3$, and TA is the total accruals scaled by lagged total assets.

B. Dechow et al (1995) – Modified Jones model

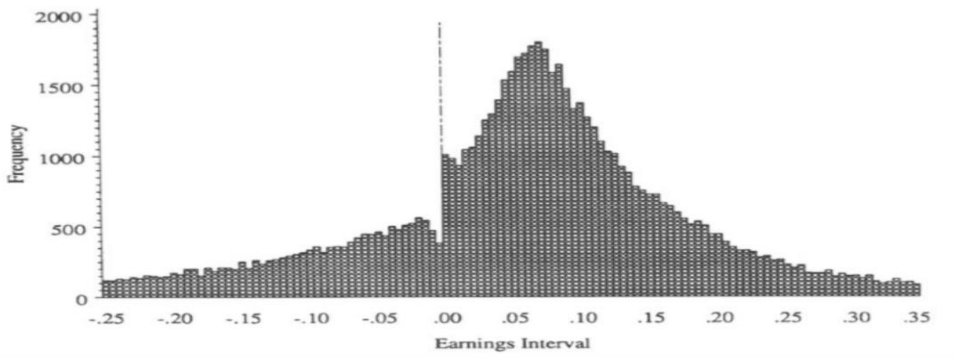
Prediction model: $TA_t = a_1(1/A_{t-1}) + a_2(\Delta REV_t - \Delta REC_t) + a_3(PPE_t) + v_t$

Where

$\Delta REC_t = net\ receivables\ in\ year\ t\ less\ net\ receivables\ in\ year\ t - 1\ scaled\ by\ total\ assets\ at\ t - 1$

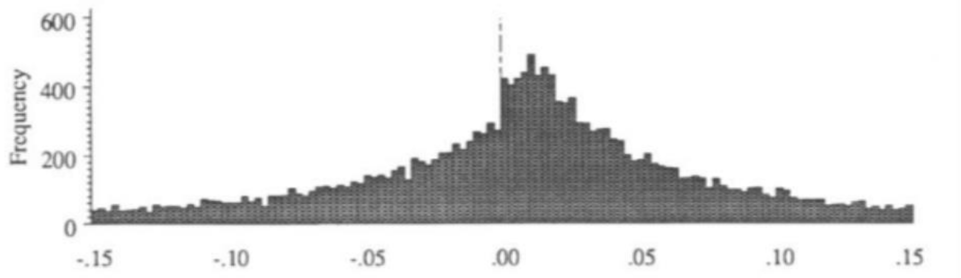
C. Burgstahler and Dichev (1997): Distribution around thresholds

Earnings scaled by market value

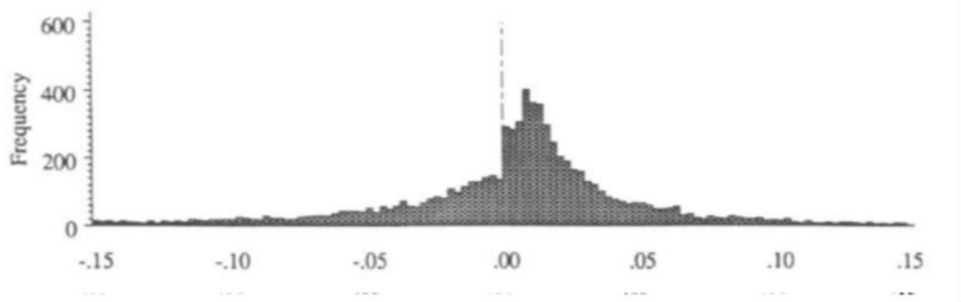


Changes of earnings scaled by market value

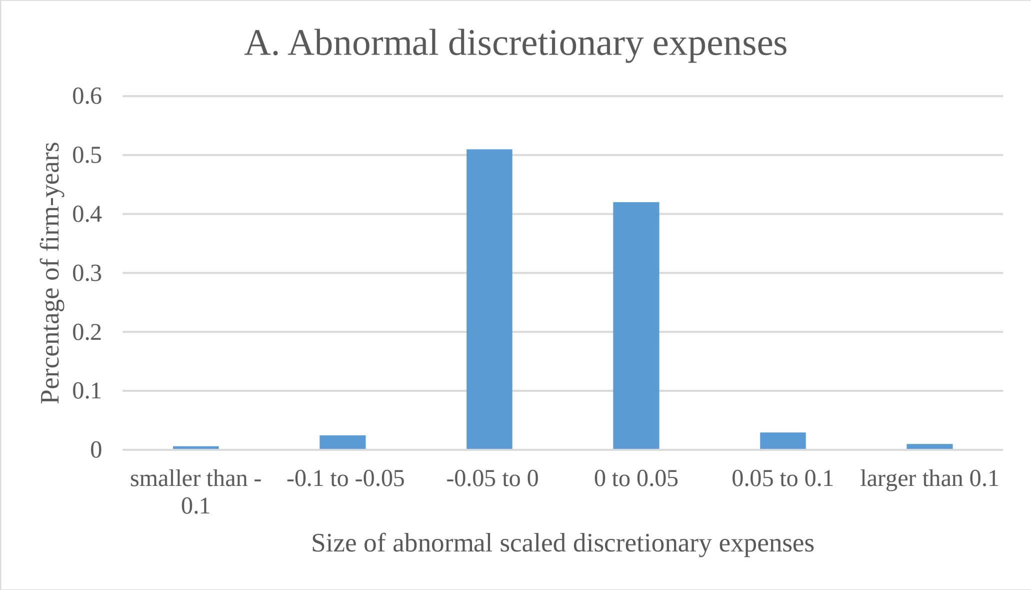
Panel B: Year Subsequent to 1 or 2 Years of Earnings Increases

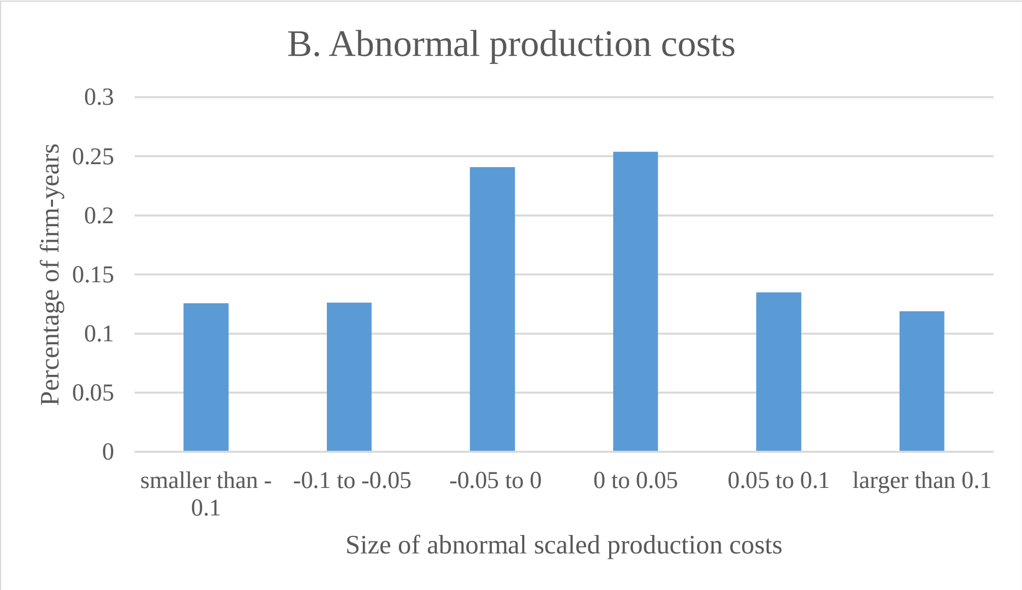


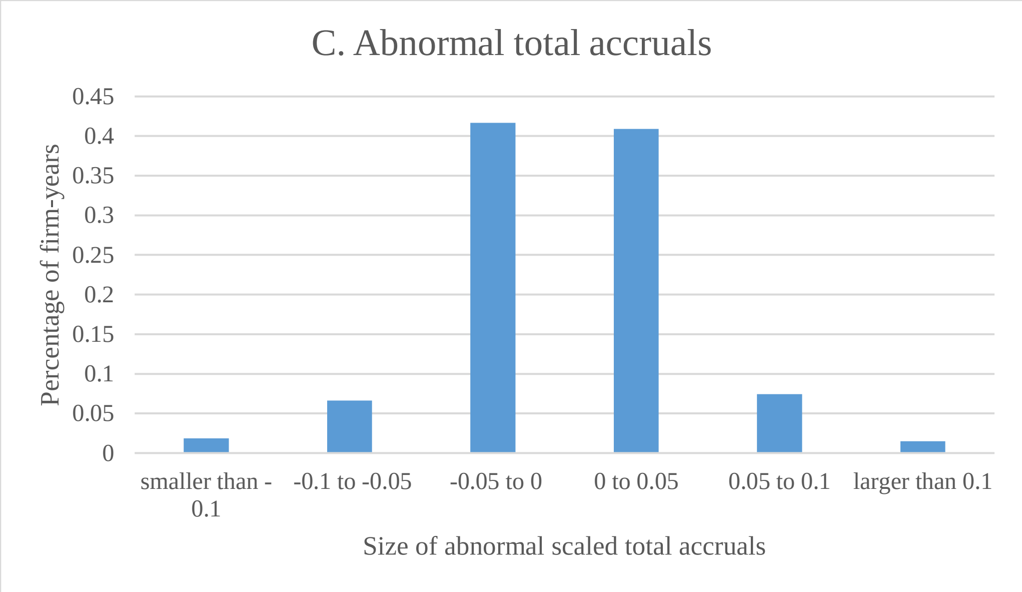
Panel C: Year Subsequent to 3 or More Years of Earnings Increases



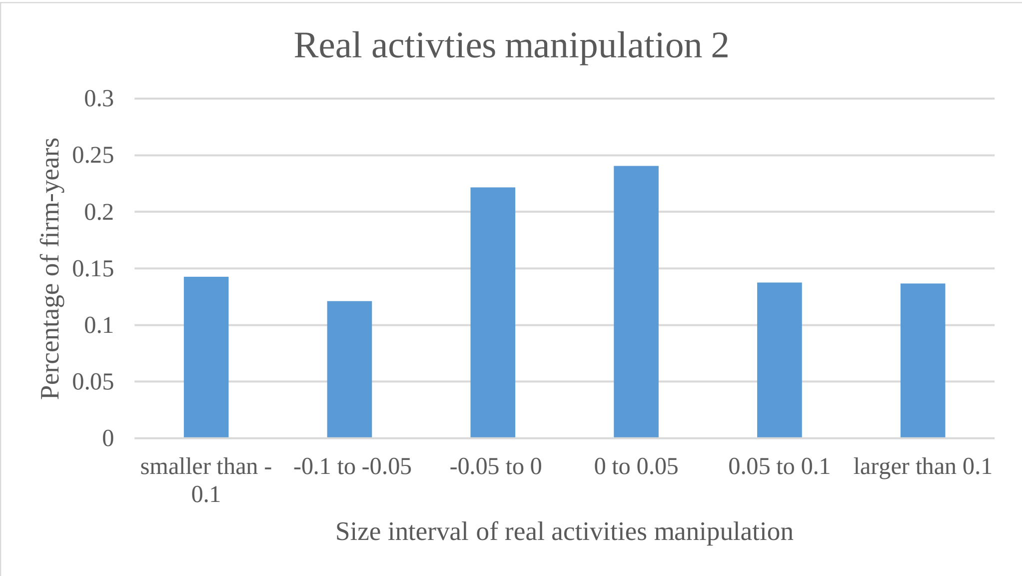
D. Distributions of abnormal EM measures with alternative suspect firm-years







Real activites manipulation with alternative assumption about suspicious firm-years (RAM2)



E. Choose 2000 “suspicious” years randomly

	Abnormal $DISEXP_t/A_{t-1}$	Abnormal $PRODC_t/A_{t-1}$	Abnormal $TACC_t/A_{t-1}$
Mean	-0.00025	-0.00085	-0.00017
Std Dev	0.048	0.205	0.049
Absolute mean	0.018	0.081	0.031

Bibliography

Barth, M.E., Elliott, J.A. & Finn, M.W. 1999, "Market rewards associated with patterns of increasing earnings", *Journal of Accounting Research*, vol. 37, no. 2, pp. 387-413.

Burgstahler, D. & Dichev, I. 1997, "Earnings management to avoid earnings decreases and losses", *Journal of Accounting and Economics*, vol. 24, no. 1, pp. 99-126.

Cohen, D.A. & Zarowin, P. 2010, "Accrual-based and real earnings management activities around seasoned equity offerings", *Journal of Accounting and Economics*, vol. 50, no. 1, pp. 2-19.

Dechow, P.M., Sloan, R.G. & Sweeney, A.P. 1995, "Detecting Earnings Management", *The Accounting Review*, vol. 70, no. 2, pp. 193-225.

Dechow, P.M., Kothari, S.P. & L. Watts, R. 1998, "The relation between earnings and cash flows", Journal of Accounting and Economics, vol. 25, no. 2, pp. 133-168.

Dechow, P.M. & Dichev, I.D. 2002, "The quality of accruals and earnings: The role of accrual estimation errors", Accounting Review, vol. 77, no. SUPPL., pp. 35-59.

Degeorge, F., Patel, J. & Zeckhauser, R. 1999, "Earnings management to exceed thresholds", *Journal of Business*, vol. 72, no. 1, pp. 1-33.

Eng, L.L. & Lin, Y.C. 2012, "Accounting quality, earnings management and cross-listings: Evidence from China", *Review of Pacific Basin Financial Markets and Policies*, vol. 15, no. 2.

Healy, P.M. & Wahlen, J.M. 1999, "A review of the earnings management literature and its implications for standard setting", *Accounting Horizons*, vol. 13, no. 4, pp. 365-383.

Ho, L.-.J., Liao, Q. & Taylor, M. 2015, "Real and accrual-based earnings management in the pre- and post-IFRS periods: Evidence from China", *Journal of International Financial Management and Accounting*, vol. 26, no. 3, pp. 294-335.

Hribar, P. & Collins, D.W. 2002, "Errors in estimating accruals: Implications for empirical research", *Journal of Accounting Research*, vol. 40, no. 1, pp. 105-134.

Jones, J.J. 1991, "Earnings Management During Import Relief Investigations", *Journal of Accounting Research*, vol. 29, no. 2, pp. 193-228.

Kuo, J.-., Ning, L. & Song, X. 2014, "The Real and Accrual-based Earnings Management Behaviors: Evidence from the Split Share Structure Reform in China", *International Journal of Accounting*, vol. 49, no. 1, pp. 101-136.

Li, T., Luo, M. & Ng, D. 2014, "Earnings management and listing regulations in China", *China Finance Review International*, vol. 4, no. 2, pp. 124-152.

Roychowdhury, S. 2006, "Earnings management through real activities manipulation", *Journal of Accounting and Economics*, vol. 42, no. 3, pp. 335-370.

Xiao, G. 2015, "Trading and earnings management: Evidence from China's non-tradable share reform", *Journal of Corporate Finance*, vol. 31, pp. 67-90.

Zang, A.Y. 2012, "Evidence on the Trade-Off between Real Activities Manipulation and Accrual-Based Earnings Management", *The Accounting Review*, vol. 87, no. 2, pp. 675-703.