Financial Characteristics Identification of Earnings Management Behavior of Listed Companies Based on Cluster Analysis

Yanping Liu[1,a]; Chao Wang [2,b]; Xiu Cheng* [1,c]

[a]sylyp008@dlut.edu.cn, [b]1368716937@qq.com, [c]1172898283@qq.com

[1]School of Economics and Management, Dalian University of Technology, China

[2]School of Economics and Management, Tongji University, China

Abstract. In this paper, cluster analysis method of multivariate statistical method is used to identify the financial characteristics of earnings management behavior of listed companies, proving that the financial characteristics of the company can reflect the earnings management level of the company to some extent. The research shows that there are 17 categories of companies have significant earnings management behavior, and these companies have weak growth ability, but strong profitability and operation ability, with higher incentive of management. After secondary classification of these categories, further analysis shows that there is a significant positive correlation between profitability, operation ability and growth ability and earnings management behavior. There is a significant negative correlation with management incentive and earnings management behavior.

Keywords： Clustering Analysis; Earnings Management; Factor Analysis; Financial Character

1. Introduction

The purpose of regular disclosure of financial reports by listed enterprises is to show their actual operating conditions to the securities market. However, the figures shown in financial reports cannot reflect the real profits of enterprises. This kind of legal and non-compliant earnings management behavior of enterprises is bound to affect the outside world to make a correct judgment on the real business situation of enterprises. Therefore, how to obtain the real financial information investors want from these possible distorted financial reports has been an urgent problem for the accounting practice circle to solve.

2. Earnings management model

2.1 Introduction to earnings management

Earnings management is also known as Earnings Management. Ronen (2011) summarized previous definitions of earnings management and divided it into three categories: white, gray, and black[1]. This article believes that most of the definitions of earnings management in the future have evolved from Scott's (1997) definition, and domestic scholars have also borrowed

ICIDC 2023, June 02-04, Nanchang, People's Republic of China

DOI 10.4108/eai.2-6-2023.2334617

from Scott's (1997) definition of earnings management [2]. The literature on the motivation of earnings management is mainly divided into financing motivation[3], avoidance motivation[4], selfish motivation[5], job motivation[6], and tax motivation[7].

2.2 Accrued earnings management model variable

The Jones model is the most widely used econometric model for accrued earnings management. Jones (1991) abandoned the characteristics of random walk and mean reversion of traditional discretionary accruals, and believed that discretionary accruals changed with the change of main business income and the company's fixed assets[8].The revised Jones cross-section model is adopted in this paper, and the variables selected are listed companies' net profit, cash flow generated from operating activities, main business income, fixed assets, accounts receivable and total assets[9]. The extent to which operational accruals correspond to earnings management is estimated.

$$\mathrm{TA}_{i,t} = NI_{i,t} - CFO_{i,t} \tag{1}$$

$$\mathrm{NDA}_{i,t} = \alpha_0 + \alpha_1 \times \frac{1}{A_{i,t-1}} + \alpha_2 \times \frac{\Delta REV_{i,t} - \Delta REC_{i,t}}{A_{i,t-1}} + \alpha_3 \times \frac{PPE_{i,t}}{A_{i,t-1}} + \alpha_4 \times ROA_{i,t} \tag{2}$$

$$\frac{\mathrm{TA}_{i,t}}{A_{i,t-1}} = \alpha_0 + \alpha_1 \times \frac{1}{A_{i,t-1}} + \alpha_2 \times \frac{\Delta REV_{i,t} - \Delta REC_{i,t}}{A_{i,t-1}} + \alpha_3 \times \frac{PPE_{i,t}}{A_{i,t-1}} + \alpha_4 \times ROA_{i,t} + \varepsilon_{i,t} \tag{3}$$

Where, TAi,t is the total accrued profit of Company i in year t, ΔREVi,t is the change amount of company i's operating income in year t, ΔRECi,t is the change amount of Company i's accounts receivable in year t, PPEi,t is the total fixed assets of Company i in year t, ROAi,t is the return rate of total assets of Company i in year t, εi,t is the residual term, Ai,t-1 is the total assets of Company i in year t-1.

The above calculation can obtain total accrued profit and non-manipulative profit, then manipulative profit is equal to total acrued profit minus non-manipulative profit. As shown in Equation (4).

$$\mathrm{DA}_{i,t} = TA_{i,t} - NDA_{i,t} \tag{4}$$

Following the Jones method, we use controllable profit to estimate the degree of accrual earnings management, as shown in Equation (5).

$$AEM = DA_{i,t} \tag{5}$$

The estimated AEM of accrued earnings management is divided into two parts, namely positive earnings management AEM+ and negative earnings management AEM-, and the two methods are identified respectively.

2.3 Real earnings management model variables

Using the Roychowdhury real earnings management measurement model, real earnings management is divided into three proxy variables of abnormal activity cash flow, abnormal discretionary expense and abnormal production cost, which correspond to three real earnings management behaviors of sales manipulation, production manipulation and expense manipulation respectively[10].

$$\frac{NCFO_{i,t}}{A_{i,t-1}} = \beta_0 + \beta_1 \times \frac{1}{A_{i,t-1}} + \beta_2 \times \frac{SALES_{i,t}}{A_{i,t-1}} + \beta_3 \times \frac{\Delta SALES_{i,t}}{A_{i,t-1}} \tag{6}$$

$$\frac{CFO_{i,t}}{A_{i,t-1}} = \beta_0 + \beta_1 \times \frac{1}{A_{i,t-1}} + \beta_2 \times \frac{SALES_{i,t}}{A_{i,t-1}} + \beta_3 \times \frac{\Delta SALES_{i,t}}{A_{i,t-1}} + \xi_{i,t} \tag{7}$$

$$\frac{NDISX_{i,t}}{A_{i,t-1}} = \gamma_0 + \gamma_1 \times \frac{1}{A_{i,t-1}} + \gamma_2 \times \frac{SALES_{i,t-1}}{A_{i,t-1}} \tag{8}$$

$$\frac{DISX_{i,t}}{A_{i,t-1}} = \gamma_0 + \gamma_1 \times \frac{1}{A_{i,t-1}} + \gamma_2 \times \frac{SALES_{i,t-1}}{A_{i,t-1}} + \tau_{i,t} \tag{9}$$

$$\frac{NPROD_{i,t}}{A_{i,t-1}} = \varphi_0 + \varphi_1 \times \frac{1}{A_{i,t-1}} + \varphi_2 \times \frac{SALES_{i,t}}{A_{i,t-1}} + \varphi_3 \times \frac{\Delta SALES_{i,t}}{A_{i,t-1}} + \varphi_4 \times \frac{\Delta SALES_{i,t-1}}{A_{i,t-1}} \tag{10}$$

$$\frac{PROD_{i,t}}{A_{i,t-1}} = \varphi_0 + \varphi_1 \times \frac{1}{A_{i,t-1}} + \varphi_2 \times \frac{SALES_{i,t}}{A_{i,t-1}} + \varphi_3 \times \frac{\Delta SALES_{i,t}}{A_{i,t-1}} + \varphi_4 \times \frac{\Delta SALES_{i,t-1}}{A_{i,t-1}} + \upsilon_{i,t} \tag{11}$$

By referring to Jones' method, the estimated values of the three proxy variables of abnormal activity cash flow, abnormal discretionary expense and abnormal production cost are obtained, as shown in Equations (12), (13) and (14).

$$\mathrm{R}EM_{CFO} = \frac{CFO_{i,t}}{A_{i,t-1}} - \frac{NCFO_{i,t}}{A_{i,t-1}} \tag{12}$$

$$\mathrm{R}EM_{\mathrm{DISX}} = \frac{\mathrm{DISX}_{i,t}}{A_{i,t-1}} - \frac{N\mathrm{DISX}_{i,t}}{A_{i,t-1}} \tag{13}$$

$$\mathrm{R}EM_{PROD} = \frac{PROD_{i,t}}{A_{i,t-1}} - \frac{NPROD_{i,t}}{A_{i,t-1}} \tag{14}$$

Finally, REM, the estimated value of the real degree of earnings management, is obtained, and the calculation formula is shown in Equation (15).

$$\mathrm{R}EM = EM_{PROD} - EM_{DISX} - EM_{CFO} \tag{15}$$

Similarly, REM refers to the direction differentiation method of accrued earnings management. This paper divides REM of real earnings management into two parts, namely REM+ of positive real earnings management and REM- of negative real earnings management.

3. Research design

3.1 Variable selection

Clustering variable.

Table 1. Clustering variable table

Category	Influencing factor	Financial index
Earnings re-	Profitability	Return on equity, return on total assets, return on total assets

lates to financial indicators	Operational capacity	Total assets turnover, current assets turnover, accounts receivable turnover
	Solvency	Current ratio, quick ratio, cash from sales ratio, asset-liability ratio
	Growth ability	Growth rate of total assets, growth rate of operating profit, growth rate of operating revenue, return on equity
Earnings management related to financial indicators		The ratio of cash flow from operating activities to last year's total assets
		The ratio of accrued profit to last year's total assets, the ratio of change in accounts receivable to last year's total assets
		Change in revenue compared to last year's total assets
		Ratio of fixed assets to last year's total assets
Non-financial indicators of earnings management	Ownership concentration	Shareholding ratio of the largest shareholder, the top five shareholders and the top ten shareholders
	Internal control of the company	DIB Internal control index
	Management incentive situation	Equity incentive plans, executive compensation
	Institutional investor shareholding	Proportion of shares held by institutional investors
	Earnings management preference of listed companies	Earnings management preference is measured by the financial restatement of the previous year. 0 means no financial restatement of the previous year and 1 means financial restatement of the previous year

3.2 Data processing

In this paper, A-share companies listed on the main board of Shanghai and Shenzhen Stock exchanges from 2010 to 2018 are taken as clustering objects, and companies that issue B shares or H shares at the same time are excluded. At the same time, the following steps are processed: samples of financial listed companies, data of IPO listed companies in the same year, data of PT, ST, and *ST "hat" companies, and samples with missing values and wrong values are excluded.

The sample data mainly came from RESSET database, and some data came from CSMAR database and DIB internal control and risk management database. EXCEL was used for processing, and finally 9286 samples were obtained. SPSS 22.0 software was used for analysis.

3.3 Cluster analysis model

Factor analysis.

Factor analysis is divided into exploratory factor analysis and confirmatory factor analysis. In this paper, exploratory factor analysis is used to reduce the dimension of 26 cluster variables.

Each variable in exploratory factor analysis can be expressed as a linear combination of common factors and the sum of special factors, namely:

$$X_i = a_{i1}F_1 + a_{i2}F_2 + a_{i3}F_3 + \cdots + a_{in}F_n, i = 1,2,\ldots,m \tag{16}$$

Where, m≤26, X_i is the cluster variable No. 1-26, $F_1, F_2, F_3, \cdots, F_n$ is the common factor of X_i, and ε_i is the special factor of X_i. The matrix of this model is expressed as:

$$\mathrm{X} = AF + \varepsilon \tag{17}$$

The following conditions should be met between these indicators:

$$m \geq n \text{ ;} \tag{18}$$

$$Cov(F, \varepsilon) = 0 \text{ ;} \tag{19}$$

$$D(F) = I_n \tag{20}$$

$$D(\varepsilon) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m^2 \end{bmatrix} \tag{21}$$

Cluster analysis.

After dimensionality reduction by factor analysis, common factor $F_1 - F_n$ is obtained. Cluster analysis is carried out on samples by calculating common factor value $f_{11} - f_{pn}$ corresponding to each sample object. Sample space is represented by Y.

The main idea of cluster analysis is to cluster according to distance. After obtaining the sample space matrix Y, the distance matrix R(r_{ij}) of the sample space can be obtained through calculation. Euclidean distance is used to calculate the distance, namely:

$$r_{ij} = \sqrt{\sum_{k=1}^{n} \left(f_{ik} - f_{kj} \right)^2} \text{ 。} \tag{22}$$

Suppose s classes are generated in the clustering result, then the original large sample space is divided into s small sample space, namely: $Y_1, Y_2, Y_3, \cdots, Y_s$, where the sample capacity of each class is $q_1, q_2, \cdots, q_s$, and $q_1 + q_2 + \cdots + q_s = p$. It also produces s small distance matrices $R_1, R_2, R_3, \cdots, R_s$, In addition, there is also a distance between two samples of each two classes, which is expressed by O(O_{ij}), wherein the sum of internal distances of each class is K(K_i), and the distance between samples of each class is expressed by L(l_{ij}). The idea of cluster analysis is to make the distance K within a class small enough and the distance L between two classes large enough through the clustering analysis algorithm, that is, an optimization problem as shown below:

$$\begin{aligned} &Min. \sum_{i=1}^{s} K_i \\ &Max. \sum_{i=1}^{s} \sum_{j=1}^{s} l_{ij} \\ &st. \begin{cases} K_i = \dfrac{1}{q_i^2} \sum_{s=1}^{q_i} \sum_{t=1}^{q_i} r_{st} \\ l_{ij} = \dfrac{1}{q_i q_j} \sum_{s=1}^{q_i} \sum_{t=1}^{q_j} r_{st} \end{cases} \end{aligned} \tag{23}$$

This optimization problem is optimized to find its optimal solution, that is to say, an optimal classification, so as to conduct a classification of listed company samples.

4. Empirical analysis

4.1 Descriptive Statistical Analysis

Through descriptive statistical analysis of research variables, it was found that among 9286 samples, the mean and variance of financial indicators were within a reasonable range, and singular values were removed during Winsorize processing. After comparing the standard deviation of earnings management indicators with Jones' results, it was found that the values calculated by domestic data were smaller than the results of Jones' model, indicating that the control over earnings management in China is not as strong as that of American companies, and domestic companies are more cautious in earnings management. The average value of the company's internal control is 648, which indicates that the internal control level of domestic enterprises is constantly improving. The average shareholding ratio of the largest shareholder in the shareholding ratio is as high as 35.4%, and the shareholding ratio of the top five shareholders is as high as 51.5%, which indicates that the phenomenon of one share dominance in China's listed enterprises is relatively serious, and the equity concentration ratio is relatively high. The highest shareholding ratio of institutional investors does not exceed 2%, indicating that the proportion of institutional holdings in domestic enterprises is relatively small and institutional investors are still in a state of development. The average financial restatement for the previous year was only 0.18, indicating that the quality of financial reporting for domestic enterprises is gradually increasing. The average of the relevant measurement indicators for earnings management is relatively small, but the maximum value is relatively large, indicating that the overall earnings management is relatively small, but there are still a few companies with serious earnings management phenomena.

4.2 Factor analysis

Principal component analysis was used and Varimax orthogonal rotation was performed on the factor load matrix. The results of factor analysis are as follows.

Table 2 shows KMO and Bartlett sphericity test, where KMO value is 0.641, greater than 0.6, and significance is 0.00 and less than 0.05, indicating strong correlation between these variables, which is suitable for factor analysis.

As shown in the total variance interpretation table in Table 3 and the macrubble diagram in Figure 1, there are 9 characteristic roots larger than 1 in the initial factor solution, while the characteristic roots of the remaining factors are relatively small. The cumulative contribution rate of the difference of the first 9 common factors is 69.62%. Therefore, the first 9 common factors are selected in this paper to establish the factor loading matrix.

Table 2. KMO and Bartlett tests

KMO sampling appropriateness measure		.641
Bartlett's sphericity test	Last chi-square read	143730.961
	Degree of freedom	325
	Significance	.000

Table 3. Total variance interpretation

Component	Initial eigenvalue			Extract the sum of loads squared			Sum of the squares of rotating loads		
	Total	Variance (%)	Accumula-tion (%)	Total	Variance(%)	Accumu-lation (%)	Total	Variance (%)	Accumula-tion(%)
1	4.075	15.671	15.671	4.075	15.671	15.671	2.708	10.415	10.415
2	2.931	11.275	26.946	2.931	11.275	26.946	2.705	10.406	20.821
3	2.584	9.937	36.883	2.584	9.937	36.883	2.598	9.993	30.814
4	2.049	7.880	44.763	2.049	7.880	44.763	2.178	8.379	39.193
5	1.669	6.421	51.184	1.669	6.421	51.184	1.987	7.644	46.837
6	1.367	5.257	56.441	1.367	5.257	56.441	1.801	6.928	53.765
7	1.324	5.090	61.532	1.324	5.090	61.532	1.761	6.772	60.537
8	1.091	4.198	65.729	1.091	4.198	65.729	1.271	4.890	65.427
9	1.012	3.891	69.621	1.012	3.891	69.621	1.090	4.194	69.621

Extraction method: principal component analysis

Table 4. Factor load matrix after rotation

	Component								
	1	2	3	4	5	6	7	8	9
Share of the top five shareholders	.968	.092	.052	.023	.016	-.010	.015	.045	-.007
The proportion of shares held by the top 10 shareholders	.938	.107	.075	.078	.001	-.003	.024	.091	.005
Shareholding ratio of the largest shareholder	.851	.067	-.053	-.091	.067	-.015	-.030	-.102	-.054
Return on assets	.091	.867	.143	.164	.134	.157	.054	.093	-.021
Return on equity	.101	.828	.002	.167	.149	.271	-.063	.163	-.047
Operating margin	.113	.790	.264	.063	-.221	.150	.148	.099	.028
Current ratio	.029	.042	.948	-.028	-.074	.003	-.110	.027	-.010
Quick ratio	.035	.047	.947	-.007	-.061	-.001	-.062	.025	-.013
Asset-liability ratio	.002	-.198	-.738	.015	.098	.019	-.111	.113	.000
Change in accounts receivable	.009	.047	.010	.686	.360	.083	-.022	.058	-.091
Growth rate of total assets	.006	.202	-.052	.664	-.128	-.010	.030	.102	.063
Revenue growth rate	.002	.029	.025	.662	.031	.293	.094	.126	-.003
Change in operating revenue	.002	.069	-.033	.656	-.056	-.069	-.270	-.022	-.023
Turnover of total assets	.037	.011	-.060	.108	.892	.034	-.212	.026	-.089
Turnover of current assets	.036	.010	-.204	.046	.848	-.010	.265	-.085	-.039
Accounts receivable turnover rate	.024	.123	-.014	-.191	.390	-.056	.030	.145	.254
Operating profit growth rate	-.014	.160	-.007	.068	-.009	.881	.021	.010	.008
Net profit growth rate	-.008	.256	-.011	.094	-.001	.858	.003	-.016	.000
Fixed assets	-.032	.026	-.212	.250	-.014	-.092	.730	-.164	.068
Cash on sales ratio	.067	.331	.175	-.180	-.103	.055	.701	.076	-.036
Accrued profit	.006	.215	.017	.340	-.173	-.098	-.668	-.072	.029
Executive compensation	-.009	.244	-.160	-.027	.050	-.048	-.087	.724	-.064
Proportion of shares held by institutional investors	.323	-.098	-.003	.066	-.031	.192	.125	.612	.043
Equity incentive plan	-.132	.088	.076	.177	.014	-.062	-.029	.435	-.023
Financial restatement of previous year	-.023	.018	-.028	.053	.028	.020	-.034	-.079	.867

	Component								
	1	2	3	4	5	6	7	8	9
Internal control situation	.051	.432	-.025	.102	.116	.007	-.075	-.022	-.482

Note: Extraction method: principal component analysis. Rotation method: Kaiser standardized maximum Variance method

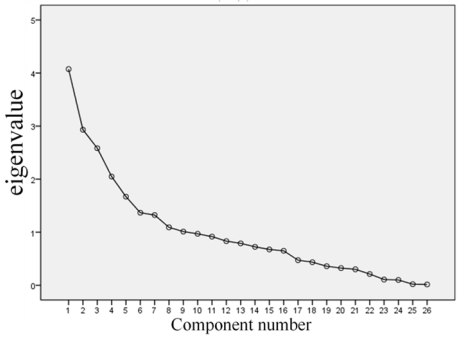

Fig. 1. Lithotripsy

According to the factor loading matrix after rotation in Table 4, each common factor is constituted, as shown in Table 5

Table 5. Names and compositions of common factors

Common factor	Factor name	Factor composition
The first common factor	Ownership concentration	Shareholding ratio of the largest shareholder, top five shareholders and top ten shareholders
The second common factor	profitability	Return on assets, return on equity and operating profit rate
The third common factor	Solvency	Current ratio, quick ratio and asset-liability ratio
The fourth common factor	Earnings management and growth capacity integrated factors	Change in accounts receivable, total assets growth rate, operating revenue growth rate and operating income
The fifth common factor	Operational capacity	Total assets turnover, current assets turnover and accounts receivable turnover
The sixth common factor	Growth ability	Operating profit growth rate, net profit growth rate
The seventh common factor	Earnings management and solvency factors	Fixed assets, cash on sales ratio and accrued profit
The eighth common factor	Governance structure	Executive compensation, institutional investor ownership and equity incentive plans
The ninth common factor	Internal control	Previous year's financial restatement and internal control

4.3 Cluster analysis

This paper uses K-Means clustering analysis algorithm for data clustering analysis. Considering the sample number of each category, we decide to use 50 categories for initial clustering, so that

the number of each category is roughly between 100 and 500, which is more in line with the actual situation.

Table 6 is the variance analysis table. It can be seen from the table that these factors make significant contributions to the classification results, that is, the classification results are significant.

Table 7 shows the number of samples in each category. It can be seen from the table that the distribution of the number of different samples is relatively uniform, which conforms to the hypothesis of cluster analysis. However, since the number of samples in categories 3, 10, 14, 15, 17, 27, 28, 30, 42 and 48 are not more than 100, they will not be considered in subsequent studies, and only those categories with more than 100 samples will be considered.

Table 8 is a comparison table of the mean values of various samples of earnings management indicators (due to the length, this paper only shows the mean values of categories with higher earnings management degree). It can be seen from the table that the earnings management indicators of categories 10, 50, 28, 18, 27, 9, 46, 21, 36, 31, 17, 30, 41, 12, 13, 14 and 20 are higher. That is, the positive accrual earnings management indicators of 10, 50, 28, 18 and 27 are high; the negative accrual earnings management indicators of 9, 46, 21, 36 and 31 are high; the positive real earnings management indicators of 17, 30, 10, 50 and 28 are high; the negative real earnings management indicators of 18, 41, 12, 13, 14 and 20 are high. However, there are listed companies that use both positive accrual earnings management and positive real earnings management to control profits, such as the 10th, 50 and 28 categories. There are also listed companies that use accrued earnings management and real earnings management to control earnings from different directions.

Table 6. ANOVA analysis table

	Clustering		Error		F	Significance
	Mean square	df	Mean square	df		
REGR factor score 1 for analysis 1	103.627	49	.456	9236	227.489	.000
REGR factor score 2 for analysis 1	110.201	49	.421	9236	261.974	.000
REGR factor score 3 for analysis 1	127.266	49	.330	9236	385.518	.000
REGR factor score 4 for analysis 1	106.313	49	.441	9236	240.921	.000
REGR factor score 5 for analysis 1	119.228	49	.373	9236	319.849	.000
REGR factor score 6 for analysis 1	134.095	49	.294	9236	456.278	.000
REGR factor score 7 for analysis 1	102.946	49	.459	9236	224.213	.000
REGR factor score 8 for analysis 1	115.713	49	.391	9236	295.630	.000
REGR factor score 9 for analysis 1	158.190	49	.166	9236	952.629	.000

Table 7. Number of samples in each category

1	210.000	11	130.000	21	160.000	31	47.000	41	273.000
2	272.000	12	275.000	22	458.000	32	312.000	42	92.000
3	65.000	13	274.000	23	330.000	33	143.000	43	124.000
4	423.000	14	11.000	24	255.000	34	34.000	44	122.000
5	297.000	15	8.000	25	283.000	35	222.000	45	117.000
6	166.000	16	217.000	26	352.000	36	121.000	46	264.000

7	474.000	17	32.000	27	43.000	37	113.000	47	164.000
8	194.000	18	106.000	28	1.000	38	54.000	48	88.000
9	231.000	19	117.000	29	162.000	39	3.000	49	365.000
10	48.000	20	171.000	30	79.000	40	325.000	50	459.000

In this paper, samples of the category with high earnings management degree and ordinary samples were separated for average comparison. The difference between the two types of samples in terms of profitability is mainly reflected in the two indicators of return on equity and operating cost rate of return. The two indicators of the company with more serious earnings management degree are higher, indicating that the company with more serious earnings management degree has stronger profitability. The growth ability is mainly reflected in the growth rate of operating profit, operating income and net profit. Companies with more serious earnings management degree have lower growth indicators, which indicates that their growth ability is weak. In terms of operating capacity, it is mainly reflected in the turnover of total assets and accounts receivable. Companies with more serious earnings management have higher indicators of these indicators, indicating that their operating capacity is higher. In terms of management incentive, both equity incentive plan and executive compensation are relatively high, indicating that these companies have strong management incentive. To sum up, companies in these categories with a high degree of earnings management have strong profitability and operation ability, weak growth ability and strong incentive of management.

Table 8. Comparison of average earnings management indicators of various samples

Category	Sample number	Positive accrued earnings management	Negative accrued earnings management	Positive real earnings management	Negative real earnings management
10	166	.082	.004	.089	.066
50	496	.067	.004	.070	.071
28	118	.047	.019	.172	.145
18	177	.060	.013	.008	.363
27	208	.050	.005	.049	.083
9	166	.014	.036	.058	.069
46	174	.009	.037	.042	.130
21	193	.009	.047	.058	.046
36	188	.009	.047	.049	.103
31	109	.005	.046	.016	.104
17	187	.024	.028	.146	.108
30	181	.019	.033	.108	.181
41	152	.039	.020	.019	.198
12	265	.031	.017	.017	.211
13	281	.020	.021	.030	.192
14	171	.017	.023	.001	.308
20	195	.015	.028	.004	.205

4.4 Discriminant analysis

In order to test the relationship between categories and earnings management related indicators, discriminant analysis was carried out on each cluster variable. The observation coefficients

showed that the discriminant function coefficients of each category had a big difference in earnings management indicators. Table 9 showed that positive accrual earnings management was stronger in categories 10, 50, 28, 18 and 27. In the discriminant function, the coefficients of accrual profit, change of accounts receivable, change of operating income and fixed assets differ greatly, indicating that the main discriminant basis of the discriminant function is the relevant indicators of earnings management.

Table 9 shows the validity test of discriminant equations. It can be seen from the table that the validity of discriminant functions of all categories of criterion is significant.

Table 9. shows the validity test of discriminant equation

Category	Λ Statistic	Chi-square	Degree of freedom	Significance
Class 10	.863	1362.149	26	.000
Class 50	.812	1934.145	26	0.000
Class 28	.882	1168.296	26	.000
Class 18	.804	2025.713	26	0.000
Class 27	.903	950.797	26	.000
Class 9	.875	1240.040	26	.000
Class 46	.930	674.045	26	.000
Class 21	.875	1240.040	26	.000
Class 36	.780	2300.284	26	0.000
Class 31	.896	1013.547	26	.000
Class 17	.837	1645.944	26	0.000
Class 30	.868	1315.904	26	.000
Class 41	.856	1441.299	26	.000
Class 12	.870	1293.603	26	.000
Class 13	.831	1715.610	26	0.000
Class 14	.854	1459.221	26	.000
Class 20	.873	1262.806	26	.000

To sum up, there are large differences in the coefficients of accrual profit, change in accounts receivable, change in operating income and fixed assets in the discriminant function for each cluster variable, indicating that the main discriminant basis of the discriminant function is the relevant indicators of earnings management. Companies in these categories with higher earnings management degree have stronger profitability and operating capacity, but weaker growth capacity. Have strong management incentive.

4.5 Further analysis

Table 10. Regression analysis table

Model	Unstandardized coefficient		Standardization coefficient	t	Significance
	B	Standard error	Beta		
(Constant)	-.087	.010		-9.074	.000
Equity concentration factor	.026	.010	.025	2.743	.006
Profitability factor	-.221	.010	-.207	-23.028	.000
Solvency factor	-.099	.010	-.093	-10.312	.000
Earnings management and growth capacity factors	.221	.010	.207	23.012	.000

Operational capacity factor	.160	.010	.150	16.677	.000
Growth capacity factor	.018	.010	.017	1.891	.059
Earnings management and solvency factors	-.279	.010	-.262	-29.044	.000
Governance structure factor	-.262	.010	-.247	-27.370	.000
Internal control factor	-.009	.010	-.008	-.931	.352

In order to further prove that different categories with higher earnings management degree obtained through cluster analysis have strong profitability and operating ability, weak growth ability, strong management incentive and other financial characteristics, this paper carries on a further analysis. The results are shown in Table 10. It can be seen from Table 10 that there is a significant positive correlation between profitability factor and earnings management category variable CLA. The relationship between operating capacity and growth capacity and earnings management is similar to that between profitability and earnings management, while there is a significant negative correlation between management incentive and earnings management category variables.

5. Conclusion

This paper first condenses 26 indicators into 9 factors, including equity concentration ratio factor, profitability factor, solvency factor, comprehensive factor of earnings management and growth capability, operational capability factor, growth capability factor, comprehensive factor of earnings management and solvency, governance structure factor and internal control factor, which contain most information of the original variables. Use these 9 factors as clustering objects, clustering analysis was conducted on listed companies, dividing the original sample into different categories and comparing the earnings management levels of different categories. The conclusion was drawn that there are significant differences in financial characteristics between different categories. After secondary classification of these categories, the conclusion was verified that there are significant differences in overall profitability, operational ability, growth ability, and management incentives among different categories. The final regression analysis shows that there is a significant positive correlation between profitability, operational ability, and growth ability and earnings management behavior. High or low levels of these three indicators may lead to the company's accrual of earnings management behavior, while high or low levels of these indicators may lead to the company's actual earnings management behavior.

Through cluster analysis, this paper finds that there is a significant negative correlation between management incentive and earnings management behavior. Appropriate management incentive can suppress earnings management level, but low or high management incentive will make companies generate accrued earnings management behavior, while high or low management incentive can make companies generate real earnings management behavior. To sum up, the earnings management behavior of the company will be reflected in the financial characteristics of the company to some extent. By analyzing the financial characteristics of the company, the earnings management behavior of the company can be identified to a certain extent.

Acknowledgment: This paper is supported by "Research on Modern Financial Supervision System Based on Institutional Investors and Earnings Management Perspective" of National Social Science Fund（18BJY230）

References

[1]. Joshua Ronen ,VardaYaari. Earnings Management: Emerging Insights in Theory, Practice, and Research[M].USA: Springer Science + Business Media, LLC,2011.

[2]. Chen Jufen (2012). Empirical Study on Earnings Management of Small and Medium Board IPOs Based on Governance Structure (Master's Thesis, Zhejiang University) https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2012&filename=1012302659.nh

[3]. Qi Huaijin&Huang Youwei (2016). The Choice of Earnings Management Behavior of IPO Company and the Differences between Different Markets Accounting Research (08), 34-41+96.https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KJYJ201608005&DbName=CJFQ2016.

[4]. Cai Chun, Zhu Rong, He Hui&Xie Liufang (2012). Selection of Earnings Management Methods, Implicit Behaviors, and Improvement of the Situation of Endangered Enterprises: Empirical Evidence from A-share Special Treatment Companies Accounting Research (09), 31-39+96.https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KJYJ201209007&DbName=CJFQ2012.

[5]. Yuan Zhizhu, Hao Wenhan&Wang Zeshen (2014). Empirical Study on the Impact of Management Incentives on Enterprise Accrual and Real Earnings Management Behavior Management Review (10), 181-196. doi: 10.14120/j.cnki.cn11-5057/f.2014.10.017

[6]. He Weifeng, Xiong Hui&Xuan Wenqi (2013). Research on Promotion Incentive and Earnings Management Behavior China soft science (10), 111-123. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=ZGRK201310011&DbName=CJFQ2013.

[7]. Gaidi&Hu Guoqiang (2012). A Trade off Study on Tax Evasion and Financial Reporting Costs - Evidence from China's 2008 Income Tax Reform Accounting Research (03), 20-25+94.https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KJYJ201203005&DbName=CJFQ2012

[8]. Yang Zhiqiang&Wang Hua (2014). Compensation gap within the company, equity concentration ratio and earnings management behavior - based on comparative analysis of compensation within the senior management team and between senior managers and employees Accounting Research (06), 57-65+97.
https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KJYJ201406008&DbName=CJFQ2014.

[9]. Gong Yifei and Xie Yuanfang (2018). Research on the Impact of Internal Control Deficiencies and Rectifications on Earnings Sustainability: Empirical Evidence from A-share Listed Companies Accounting Research (05), 75-82. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=KJYJ201805010&DbName=CJFQ2018.

[10]. Sun Guangguo, Liu Shuang&Zhao Jianyu (2015). Major shareholder control, institutional investor shareholding, and earnings management Nankai Management Review (05), 75-84. https://kns.cnki.net/kcms/detail/detail.aspx?FileName=LKGP201505008&DbName=CJFQ2015.