Front. Bus. Res. China 2010, 4(3): 423–450
DOI 10.1007/s11782-010-0105-0

RESEARCH ARTICLE

Liangding Jia, Guoguang Wan, Jieyu Zhou, Qian Lu

Do Continuous and Categorical Measures Capture the Same Construct of Diversity? Evidence from China

Abstract Chinese researchers have long preferred substantive research (i.e., examination of relationships among variables) to construct validation research, while failing to realize that construct validation research is the foundation of substantive research. This paper endeavors to answer a question: Do Rumelt's (1974) categorical and Standard Industrial Classification (SIC) and skill-based continuous measures capture the same construct of diversity in the context of China? We take two steps to investigate the question. First, are continuous measures (SIC & skill-based entropy and Montgomery product count) appropriately differentiated among Rumelt's categories? Second, are the categories converted from SIC & Skill-based measures convergent to Rumelt's categorical measures? Results show that the two measures are consistent with each other.

Keywords diversification, continuous measures, categorical measures, Standard Industrial Classification

Received November 9, 2009

Liangding Jia
School of Business, Nanjing University, Nanjing 210093, China
E-mail: jldyxlzs@nju.edu.cn

Guoguang Wan
Business School, Hong Kong University of Science and Technology, Hong Kong, China
E-mail: zjgt2000@gmail.com

Jieyu Zhou (✉)
School of Business, Nanjing University, Nanjing 210093, China
E-mail: juliezhou416@gmail.com

Qian Lu
Business School, National University of Singapore, Singapore 119077, Singapore
E-mail: lqian119@gmail.com

1 Introduction

In empirical studies on diversification strategy, no matter we regard diversification as a dependent, independent or intermediate variable, we would inevitably encounter the measurement problem of diversification. Then, which indicator or measurement reflects the extent and type of enterprise diversification? At present, there are two types of diversification measurement: continuous or categorical. In diversification research, especially in empirical study on the relationship between diversification and corporate performance, both methods have been widely applied (e.g., Christensen and Montgomery, 1981; Rumelt, 1974, 1982; Montogomery, 1982; Markides, 1995; Mayer and Whittington, 2003; Gort, 1962; Markham, 1973; Palepu, 1985; Amit and Livnat, 1988; Baysinger and Hoskisson, 1989; Markides, 1995; Boyd, Gove and Hitt, 2005). However, findings about the relationship between diversification and corporate performance are not consistent (Ramanujam and Varadarajan, 1989). Similarly, different diversification measurement methods have also been employed in the Chinese empirical research. For example, Li and Zhao (1998), and Li (2002) adopted specialization ratio and strategy categories; Yin (1998), Jiang, Liu and Lu (2006), Zhu (1999), Zhang, Yuan and Chen (2002), Yao, Lü and Lan (2004), Ni and Zhang (2005), Shao and Liu (2006), Hong and Xiong (2006), Qiu and Hou (2006) and Huang (2007) used continuous measurements such as business count method, Herfindahl index, and entropy index. In addition, these studies arrived at different conclusions of the relationship between diversification and performance. Why are the findings inconsistent?

Between the 1980s and 1990s, a group of scholars examined the reliability and validity of different measurement methods of diversification. For example, the SIC-based measurement method was considered suitable for study of industry-level differences between diversified and non-diversified enterprises, while the categorical measurement is appropriate for the study of strategy-level differences between related and unrelated diversified companies (Pitts and Hopkins, 1982). Palepu (1985) proposed that different findings from two schools in the research of the relationship between diversification and corporate performance mostly came from the different measurement methods, as the industrial organization school preferred continuous measurement while the strategic management school preferred categorical measurement. A survey of 128 large listed companies in the U.S. demonstrated the consistency between Rumelt's categorical measurement and product count measurement (Montgomery, 1982). Another large scale survey showed that Rumelt's categorical measurement was highly consistent with entropy index approach, while business-count approach was lowly convergent with the two methods above (Hoskisson, Hitt,

Johnson and Moesel, 1993). However, inconsistency was reported among the three measurement methods, i.e., Rumelt's categorical measurement, entropy index approach and business-count approach (Hall and John, 1994). Obviously, the findings of the latter two studies did not support Montgomery's (1982) view. Later, Farjoun (1998) developed the skill-based continuous measurement (similar to the SIC-based entropy measurement), and he confirmed the discriminant validity between them, but did not provide evidence to support the convergent validity between the two.

At present, with more access to data of Chinese listed companies, many scholars can use different methods to measure the type and degree of corporate diversification and to study the relationships between diversification and other variables. As Hoskisson et al. (1993) pointed out that, the above research can be divided into two groups: substantive research that examines the relationship research among variables versus construct validation research that focuses on measurement properties of the construct, such as content validity, reliability, convergent validity, discriminant validity, and criterion validity. Most scholars placed more emphasis on substantive research than construct validation research, regardless of the fact that the latter is the foundation of the former. However, if the constructs have measurement problems, the relationships among these constructs make no sense (Schwab, 1980; Venkatraman and Grant, 1986). Likewise, scholars in China in the area of corporate diversification preferred substantive research to construct measurement research. Under such a circumstance, a basic question remains unanswered: Do different methods that measure corporate diversification capture the same construct? Or are different measurement methods consistent with each other or convergent to the same construct? This article attempts to answer the basic question in the Chinese context by two steps. In step one, we tested whether the SIC-based and skill-based continuous measures were appropriately differentiated among Rumelt's categorical measure. In step two, we transform the SIC-based and skill-based continuous measurements into categories by cluster analysis method, and then test whether the transformed categorical measures are convergent to Rumelt's categorical measure.

2 Theoretical Background

2.1 Categorical Measures

The categorical measurement of corporate diversification derives from the basic theory of strategic management. The theory of strategic management conceives of two main features of diversification. One is the extent to which enterprises

participate in each business; the other is the degree to which a new business relates with the original one. Therefore, the categorical measurement method focuses on two dimensions: specialization ratio and relatedness ratio. An enterprise is categorized into different types of diversification by the two dimensions. Specialization ratio reflects the degree to which an enterprise is engaged in each business, while relatedness ratio captures the degree to which the businesses that an enterprise enters relate with each other. Categorical measurement was firstly proposed by Wrigley (1970), and he divided diversification into four types according to the two dimensions of specialization ratio (SR) and relatedness ratio (RR). The four types include single, dominant, related and unrelated business. Later, Rumelt (1974) improved this method by involving two new dimensions. One is constrained diversification, which is different from linked diversification. The former refers to each business unit in an enterprise is related with its core business, while the latter means that each business is correlated with at least one of the other businesses rather than all. The other is vertical dominant diversification, which refers to vertical corporations. In this way, Wrigley's four types of diversification are further split into eight subtypes, namely single, vertical-dominant, constrained-dominant, related-dominant, unrelated-dominant, constrained-linked, related-linked and unrelated business.

The remarkable feature of categorical measurement is its embodiment of a variety of sharing between businesses in an enterprise, such as production, marketing, channeling and technical sharing. The measurement was widely adopted by researchers (Bettis, 1981; Christensen and Montgomery, 1981; Montgomery, 1982, 1985; Chartterjee and Blocher, 1992; Hoskisson et al., 1993). Montgomery (1982) found that categorical measurement was relevant to continuous measurement. Chartterjee and Blocher (1992) and Hoskisson et al. (1993) also found that Rumelt's categorical measurement has high content validity in certain areas. However, shortcomings of the measurement were also revealed. For example, it may have low reliability due to its high subjectivity because it relies heavily on the researcher's personal understanding and judgments to identify the relationship among businesses an enterprise enters. It is also much time-consuming even impossible to collect sufficient relevant information to identify the relationship among businesses. Therefore, some scholars are skeptical of this measurement. Pitts and Hopkins (1982) argued that the subjective judgment process leads to different outcomes even when studying the same enterprise due to low research reliability.

2.2 Continuous Measures

Continuous measurement of corporate diversification is theoretically based on

the industrial organization theory, which defines corporate diversification from the angle of industry in which the business activities locate. Therefore, continuous measurement measures diversification according to the industrial distribution of an enterprise. The measurement employs the SIC code to define various indices to reflect the degree of diversification. SIC code was created by the U.S. National Bureau of Statistics in accordance with the "product categories." SIC code uses a 7-digit figure to distinguish goods and services, each figure representing different levels of classification. The 4-digit SIC code has been widely used in academic research. The first two digits represent an industry sector; the latter two specific business category. The continuous measurement consists of several approaches, which are listed below.

2.2.1 Business-Count Approach

Gort (1962) defined diversification as an increased number of heterogeneous markets an enterprise enters. He is the first scholar to measure corporate diversification using SIC code. Later, several similar business count measurements emerged. Wood (1971) identified broad spectrum diversity (BSD) and narrow spectrum diversity (NSD), with BSD referring to the number of businesses that an enterprise is engaged in 2-digit SIC industries, and NSD referring to the number of businesses in 4-digit SIC industries. Based on the two concepts, Varadarajan and Ramanujam (1987) invented mean narrow spectrum diversity (MNSD), namely the ratio of NSD divided by BSD.

2.2.2 Entropy Index Approach

In order to overcome the subjectivity of categorical measurement and the shortcomings of business-count approach (i.e., cannot distinguish between related and unrelated diversification) (Montgomery, 1982), some scholars attempted to create a new continuous measurement to integrate the merits of both categorical and continuous measurements. The new measurement can reflect the diversification degree as well as can be transformed into categorical measure to describe the diversification type as Rumelt's measurement does. Entropy index is created as one of the integrated approaches. Entropy index approach distinguishes the relatedness among businesses in different 2-digit SIC industries from the relatedness among businesses within the same 2-digit SIC industry. Jacquemin and Berry (1979) constructed the equation of entropy index approach as follows:

$$EDI = \sum_{i=1}^{n} p_i \ln(1/p_i), \tag{1}$$

where p_i refers to the sales proportion of the businesses in 2-digit or 4-digit SIC industries in the total sales of an enterprise; n refers to the number of 2-digit or 4-digit SIC industries in which an enterprise is engaged in. Total diversification (DT) and unrelated diversification (DU) can be calculated through the entropy index approach according to different meanings p_i and n represent. When p_i is the sales proportion of the businesses in 4-digit SIC industries in the total sales, and n is the number of 4-digit SIC industries, entropy index refers to DT; when p_i is the sales proportion of the businesses in 2-digit SIC industries in the total sales, and n is the number of 2-digit SIC industries, the entropy index refers to DU. Related diversification (DR) is the difference between DU and DT.

2.2.3 Product Count Measures (Herfindahl Index)

Berry (1971) and Mcvey (1972) independently used Herfindahl index to measure corporate diversification. The formula is as follows:

$$HDI_1 = \sum_{i=1}^{n} p_i^2, \tag{2}$$

where p_i refers to the sales or asset proportion of the businesses in 2-digit, 3-digit or 4-digit SIC industries in the total sales of an enterprise, and n refers to the number of 2-digit, 3-digit or 4-digit SIC industries in which an enterprise is engaged in. The product count measures were firstly proposed by Montgomery (1982). The measurement is similar to Herfindahl index. Its formula is as follows:

$$Diversification_{2,3,4} = 1 - \frac{\sum_j m_{ij}^2}{(\sum_j m_{ij})^2}, \tag{3}$$

where m_{ij} refers to the percentage of firm i's total sales in market j.

2.3 Skill-Based Measurement

Although the relatedness among businesses cannot be directly captured, it can be measured indirectly through inputs (such as skilled employees and specialized facilitates). If the inter-industry skill relatedness can be measured, so can corporate diversification using the skill relatedness between industries in which an enterprise is engaged in. Farjoun (1998) focused on two important kinds of relatedness: physical and skill relatedness. Physical relatedness concerns with relations among the physical characteristics of products, whereas skill relatedness highlights the resources such as research and development teams, experienced salespersons, and managerial and other skills common to two or more products. Similarity or complementarity of the physical attributes of products is the

overarching criterion for relatedness used by the SIC code. Therefore, SIC-based measures of corporate diversification such as entropy and Herfindahl indices reflect the physical relatedness between businesses within an enterprise rather than the skill relatedness. Therefore, SIC-based measurement and inter-industry skill relatedness measurement can be integrated to reflect both physical and skill relatedness of an enterprise. Using the skill-based approach, each industry is characterized by its underlying profile of specialties defined as the different types and extent of human skills required in the industry, as indicated by occupational distributions. Using cluster analysis with the Occupational Employment Survey data from United States Department of Labor and Statistics, Farjoun (1998) regrouped the 90 3-digit SIC industries within 20 2-digit SIC sectors with similar skill profiles into 8 skill-related industry groups. After the classification, we can calculate the skill-based Herfindahl and entropy indices to measure corporate diversification.

3 Sample and Measurements

3.1 Sample Selection

We focused on the companies in manufacturing sector listed in Shanghai Stock Exchange and Shenzhen Stock Exchange to control industrial effects. The *Industrial Classification of National Economy (GB/T4754-2002)* issued by the National Council of China in 2002 was followed in coding different industries. Data of the sampled companies in 2006 were collected from the Wind database. After deleting the missing data, the final sample consisted of 222 listed manufacturing companies.

3.2 Diversification Measurement

3.2.1 Rumelt's Categorical Measurement

In step 1, we calculated each company's specialized ratio (SR), vertical ratio (VR) and related ratio (RR) according to the methods suggested by Rumelt (1974), Palepu (1985) and Hall and John (1994). To guarantee the reliability of our calculation, we took the following steps. Firstly, three coders coded the sample double-blindly and classified Rumelt's categories independently. Then, the research team leader combined the encoding results and discussed them with the three coders. The three-coder-agreement rate for the coding of companies was more than 70% and the two-coder-agreement rate exceeded 90%. For those inconsistent results, we gathered more information about the companies from

their official website, related websites, and other materials. After discussion, the three coders agreed on the final results.

In step 2, three of the authors classified all sampled companies into eight Rumelt's categories on the basis of the above VR, RR and SR in a double-blind fashion and using the criterion suggested by Rumelt's (1974), Palepu (1985), and Hall and John (1994). Then the research team leader summarized the classification results. The three-researcher-agreement rate exceeded 90%. The disagreement mainly lied in the classification between constrained business and linked business. We continued to discuss the disagreed classification, using more company information and consulting relevant experts till a fully-agreed classification result was achieved. Following Rumelt's (1974) suggestion, the final sample categories were reduced from the original eight to five.

3.2.2 SIC-Based Continuous Measurement

We used two types of SIC-based continuous measurements, including the Herfindahl index proposed by Montgomery (1982), and the entropy index suggested by Palepu (1985). The listed companies reported the data according to *The Guidelines of Industrial Classification of Listed Companies* issued by China Securities Regulatory Commission (CSRC). We employed 3-digit SIC industry data to calculate the total diversity (DT) because the 4-digit SIC codes in CSRC guidelines are correspondent with the 3-digit SIC codes in the guidelines of National Council of China. We took the same coding method and procedure as above. The three-researcher-agreement rate was more than 70% and the two-researcher-agreement rate exceeded 90%. We finally arrived at a consensus after long discussion between research team leader and the three coders.

3.2.3 Skill-Based Continuous Measurement

We used the data on "industry" and "profession" from *The Fifth Nationwide Census* in 2000. There are 64 types of professions in *The Guidelines of Professions Classification of the People's Republic of China*. Some professions were deleted or merged due to the following reasons: First, there were two few people in the profession. For example, we deleted the professions of "sports worker" and "religion workers" for there were two and zero person in the category, respectively; second, we combined some similar profession types and the professions having a very small proportion in the manufacturing sector. For example, we combined three similar and peripheral types of professions, i.e., "the head of Central Committee of the Communist Party of China and of various Communist Party of China organizations," "the head of various China's state organs and their affiliations," and "the head of the democratic parties and various

social groups, and their affiliations." In so doing, we reduced the original 64 types of professions into 50 types. Using the data on "industry" and "profession" from *The Fifth Nationwide Census*, we calculated the distribution of the 50 professions across the 168 three-digit SIC manufacturing industries, which belonged to thirty-three two-digit SIC industries. Following Farjoun's (1998) procedures, we used hierarchical and nonhierarchical clustering algorithms and Ward's method to sort the 168 three-digit SIC manufacturing industries into thirty-three skill-based groups.

The 3-digit SIC industries originally within the same 2-digit SIC industry might be classified into different skill-based groups. For example, in a skill-based group, there are fourteen 3-digit SIC industries (i.e., C368, C391, C392, C393, C395, C397, C409, C411, C412, C413, C414, C415, C419, C275), which are from five different two-digit SIC industries (i.e., C27, C36, C39, C40, C41). In this group, the largest proportion of profession is "mechanical and electrical product assembling," amounting to about 23.24%. Although the 14 3-digit-SIC industries were classified into different 2-digit SIC industries in accordance with the SIC code system, they shared the highest proportion of professional workers. Therefore, we regarded them as skill-based related industries, i.e., skill relatedness.

Three researchers of this paper calculated the skill-based entropy indices (i.e., DU_Skill and DR_Skill) and the skill-based Herfindahl index (i.e., H_Skill) according to the skill-based classification separately and independently. The three-researcher-agreement rate was more than 80% and the two-researcher-agreement rate exceeded 90%. The researchers discussed the inconsistent results until a consensus was reached.

4 Data Analysis

4.1 Comparison of Continuous Measures and Rumelt's Categorical Measures

The first objective of this study is to determine whether the continuous measures of diversification can be used to differentiate between Rumelt's categorical measures. For each of eight and five Rumelt's different categories, we estimated the mean scores of SIC-based entropy indices, mean scores of SIC-based Herfindahl indices, and mean scores of skill-based entropy and Herfindahl indices. Using ANOVA method and the *post hoc* Scheffe's tests, we determined differences between a category mean and the overall mean for the other categories. The larger the differences are, the more possible that the continuous measures can be differentiated among Rumelt's categorical measures.

The results of this analysis are shown in Table 1.

Table 1 Continuous Diversification Measures vs. Categorical Diversification Measures[a]

Rumelt's category		Mean score of SIC-based entropy indices[b]			Mean score of SIC-based Herfindahl indices[c]		Mean score of skill-based entropy indices[d]		Mean score of skill-based Herfindahl indices[d]
	N	DT_SIC	DR_SIC	DU_SIC	H_SIC3	H_SIC2	DR_Skill	DU_Skill	H_Skill
1. Single business	45	0.102 1 (–)***	0.024 8 (–)***	0.077 3 (–)***	0.054 5 (–)***	0.040 3 (–)***	0.023 3 (–)***	0.078 7 (–)***	0.040 8 (–)***
2. Dominant-vertical	70	0.480 3	0.179 1 (+)**	0.301 2 (–)*	0.271 1	0.170 4 (–)**	0.176 8	0.303 4	0.172 6 (–)*
3. Dominant-constrained	11	0.212 8 (–)**	0.108 1	0.104 7 (–)**	0.112 9 (–)**	0.049 8 (–)**	0.122 0	0.090 8 (–)**	0.043 6 (–)**
4. Dominant-linked	7	0.327 5	0.105 3	0.222 2	0.186 8	0.125 5	0.113 3	0.214 1	0.124 5
5. Dominant-unrelated	21	0.488 6	0.049 1	0.439 6	0.279 4	0.252 9	0.044 7 (–)*	0.443 9	0.255 2
6. Related-constrained	13	0.552 2	0.077 7	0.474 5	0.335 0	0.291 2	0.127 0	0.425 2	0.254 3
7 Related-linked	13	0.725 5 (+)**	0.307 6 (+)***	0.417 9	0.449 6 (+)**	0.259 0	0.323 6 (+)**	0.401 9	0.254 2
8. Unrelated business	42	0.979 2 (+)***	0.151 6	0.827 6 (+)***	0.567 7 (+)***	0.494 2 (+)***	0.220 7 (+)**	0.758 5 (+)***	0.454 5 (+)***
Mean		0.499 3	0.126 1	0.373 2	0.287 8	0.217 9	0.142 5	0.356 8	0.208 7
F-Statistic[e]		26.802***	3.327***	23.870***	30.423***	25.588***	3.507***	19.315***	20.818***
R^2		0.364***	0.011	0.363***	0.399***	0.385***	0.035***	0.313***	0.335***
1. Single business	45	0.102 1 (–)***	0.024 8 (–)***	0.077 3 (–)***	0.054 5 (–)***	0.040 3 (–)***	0.023 3 (–)***	0.078 7 (–)***	0.040 8 (–)***
2. Dominant-vertical	70	0.480 3	0.179 1 (+)**	0.301 2 (–)*	0.271 1	0.170 4 (–)**	0.176 8	0.303 4	0.172 6 (–)*

(*To be continued*)

(*Continued*)

Rumelt's category		Mean score of SIC-based entropy indices [b]			Mean score of SIC-based Herfindahl indices [c]		Mean score of skill-based entropy indices [d]		Mean score of skill-based Herfindahl index [d]
	N	DT_SIC	DR_SIC	DU_SIC	H_SIC3	H_SIC2	DR_Skill	DU_Skill	H_Skill
3. Dominant business	39	0.381 9 (–)*	0.075 8	0.306 1	0.215 8 (–)**	0.172 8	0.078 8	0.303 1	0.172 0
4. Related business	26	0.638 9 (+)*	0.192 7	0.446 2	0.392 3 (+)**	0.275 1	0.225 3 (+)*	0.413 5	0.254 2
5. Unrelated business	42	0.979 2 (+)***	0.151 6	0.827 6 (+)***	0.567 7 (+)***	0.494 2 (+)***	0.220 7 (+)**	0.758 5 (+)***	0.454 5 (+)***
Mean	222	0.499 3	0.126 1	0.373 2	0.287 8	0.217 9	0.142 5	0.356 8	0.208 7
F-Statistic [f]		43.998***	4.033***	37.843***	49.369***	40.533***	4.947***	29.837***	32.299***
R^2		0.381***	0.015*	0.368***	0.413***	0.387***	0.043***	0.315***	0.336***

Note: a. A plus or minus sign in brackets indicates category means that are significantly greater than (+) or less than (–) the mean for all other categories (*t*-tests);

b. Mean score of SIC-based entropy indices were calculated according to the methods proposed by Palepu (1985);

c. Mean score of SIC-based Herfindahl indices were calculated according to the methods proposed by Montgomery (1982), H_SIC3 is the Herfindahl index for 3-digit SIC industries, H_SIC2 the Herfindahl index for 2-digit SIC industries;

d. Following the method of Farjoun (1998), we reclustered the 168 three-digit SIC industries into 33 skill-based groups using the distribution of 50 professions across the 168 three-digit SIC manufacturing industries. Then we computed skill-based entropy indices according to the method proposed by Palepu (1985) and Herfindahl index according to the method suggested by Montgomery (1982). H_Skill is the Herfindahl index fro 2-digit SIC industries;

e. The degree of freedom of *F* test is (7, 214);

f. The freedom degree of *F* test is (4, 217);

g. The total number of observation is 222. * indicates significant at level 0.10, ** indicates significant at level 0.05, *** indicates significant at level 0.01.

Under the eight categories, the results show that: (1) Both DT and DU can distinguish single and unrelated business, and DR can distinguish single business, and all are significant at the 0.01 level. In addition, DT can distinguish dominant-constrained and related-linked significantly at the level of 0.05; DR can distinguish single business significantly ($p<0.01$) and distinguish dominant-vertical business significantly ($p<0.05$). DU can distinguish dominant- onstrained significantly ($p<0.05$) and distinguish dominant-vertical marginally ($p<0.1$). (2) SIC-based Herfindahl indices (including H_SIC3 and H_SIC2) can distinguish single business and unrelated business significantly ($p<0.01$). In addition, H_SIC3 can distinguish dominant-constrained and related-linked both significantly at the level of 0.05, and H_SIC2 can distinguish dominant-vertical and dominant- constrained business significantly ($p<0.05$). (3) The skill-based entropy indices (including DR_Skill and DU_Skill) can distinguish single business significantly ($p<0.01$). In addition, DR_Skill can not only distinguish related-linked and unrelated business significantly at the level of 0.05, but also distinguish dominant-unrelated marginally ($p<0.1$); DU_Skill can distinguish unrelated business significantly ($p<0.01$), and distinguish dominant-constrained business at the significant level of 0.05. (4) The skill-based Herfindahl index can separate single business and unrelated business significantly ($p<0.01$), and can distinguish dominant-constrained business at the significance level of 0.05 and distinguish the dominant-vertical business marginally ($p<0.1$). The F test of ANOVA shows that the overall models of SIC-based entropy indices, SIC-based Herfindahl indices, skill-based entropy indices and skill-based Herfindahl index are significant. These results indicate that these four types of continuous measurement can be effectively distinguished between Rumelt's 8 categories.

Under the five categories, the results show that: (1) DT and DU can both distinguish single business and unrelated business significantly, and DR can distinguish single business significantly, and all are significant at the level of 0.01. In addition, DT can distinguish dominant business and related-linked marginally ($p<0.1$); DR can distinguish dominant-vertical business significantly ($p<0.05$); DU can distinguish dominant-vertical business marginally ($p<0.1$). (2) SIC-based Herfindahl indices (including H_SIC3 and H_SIC2) can distinguish single business and unrelated business significantly ($p<0.01$). In addition, H_SIC3 can distinguish dominant business and related-linked business and H_SIC2 can distinguish dominant-vertical both at the significant level of 0.05. (3) The skill-based entropy indices (DR_Skill and DU_Skill) can distinguish single business significantly ($p<0.01$), and DR_Skill can distinguish related business marginally ($p<0.1$); (4) Skill-based Herfindahl index can separate single business and unrelated business significantly ($p<0.01$), and can distinguish dominant-vertical business marginally ($p<0.1$). The F tests of

ANOVA show that the models of SIC-based entropy indices, SIC-based Herfindahl indices, skill-based entropy indices and skill-based Herfindahl index are all significant. These results demonstrate that these four types of continuous measurement can be effectively differentiated between Rumelt's 5 categories.

In order to analyze the ability of the continuous measures more directly to distinguish between related and unrelated diversification strategy types, we conducted Scheffe's test to test for each pairwise difference in category means. The results are shown in Table 2.

Table 2 Results of *post hoc* Analysis Using Scheffe's Difference of Means Test for Continuous vs. Categorical Diversification Measures[a]

Rumelt category	DT_SIC							DR_SIC							DU_SIC							H_SIC3						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1. Single business																												
2. Dominant-vertical	***							–							**							***						
3. Dominant-constrained	–	–						–	–						–	–						–	–					
4. Dominant-linked	–	–	–					–	–	–					–	–	–					–	–	–				
5. Dominant-unrelated	***	–	–	–				–	–	–	–				***	–	–	–				***	–	–	–			
6. Related-constrained	***	–	–	–	–			–	–	–	–	–			**	–	–	–	–			***	–	–	–	–		
7. Related-linked	***	–	**	–	–	–		**	–	–	–	–	–		*	–	–	–	–	–		***	–	***	–	–	–	
8. Unrelated business	***	***	***	***	***	***	–	–	–	–	–	–	–	–	***	***	***	***	***	**	***	***	***	***	***	***	**	–

Rumelt category	H_SIC2							DR_Skill							DU_Skill							H_Skill						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1. Single business																												
2. Dominant-vertical	**							–							**							**						
3. Dominant-constrained	–	–						–	–						–	–						–	–					
4. Dominant-linked	–	–	–					–	–	–					–	–	–					–	–	–				
5. Dominant-unrelated	***	–	–	–				–	–	–	–				***	–	–	–				***	–	–	–			
6. Related-constrained	***	–	–	–	–			–	–	–	–	–			*	–	–	–	–			**	–	–	–	–		
7. Related-linked	**	–	–	–	–	–		*	–	–	–	–	–		*	–	–	–	–	–		**	–	–	–	–	–	
8. Unrelated business	***	***	***	***	***	*	**	*	–	–	–	–	–	–	***	***	***	***	**	*	**	***	***	***	***	**	*	*

(*To be continued*)

(*Continued*)

Rumelt category	DT_SIC				DR_SIC				DU_SIC				H_SIC3			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Single business																
2. Dominant-vertical	***				**				***				***			
3. Dominant business	***	–			–	–			**	–			***	–		
4. Related business	***	–	**		*	–	–		***	–	–		***	*	***	
5. Unrelated business	***	***	***	***	–	–	–	–	***	***	***	***	***	***	***	***

Rumelt category	H_SIC2				DR_Skill				DU_Skill				H_Skill			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Single business																
2. Dominant-vertical	***				**				***				***			
3. Dominant business	**	–			–	–			**	–			**	–		
4. Related business	***	–	–		**	–	–		***	–	–		***	–	–	
5. Unrelated business	***	***	***	***	**	–	–	–	***	***	***	***	***	***	***	***

Note: $N = 222$. * denotes $p < 0.10$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

Under the eight categories, we found that the single business can be significantly distinguished from dominant-vertical, dominant-unrelated, related-constrained and related-linked business, and the unrelated business can be significantly differentiated from other six categories except related-linked business, by the measures of DT_SIC, DU_SIC, H_SIC3, H_SIC2, DU_Skill and H_Skill. However, few of the eight categories can effectively be distinguished from each other by the measures of DR_SIC and DR_Skill. The results are consistent with what Pitts and Hopkins (1982) proposed that SIC-based measurement might be appropriate for studying the industry differences between diversified and non-diversified companies. However, it can hardly distinguish related diversification from unrelated one.

Under the five categories, results show that the single business can effectively be distinguished from dominant-vertical, dominant business, related business and unrelated business by the measures of DT_SIC, DU_SIC, H_SIC3, H_SIC2, DU_Skill, and H_Skill. Unlike the results from eight categories, the single business can be significantly distinguished from dominant-vertical and related business by the measures of DR_SIC and DR_Skill, and the result under

DR_Skill is better than others because single business can be significantly distinguished from unrelated business by the skill-based entropy indices.

Further, we performed discriminant analysis using Rumelt's five strategy categories as class variables and DR and DU as predictor variables to determine whether SIC-based entropy measures (DR_SIC and DU_SIC) and skill-based entropy measures (DR_skill and DU_skill) could be used to predict a company's membership in one of the Rumelt's strategy categories. We added the analysis of skill-based entropy measures into this research in addition to the SIC-based entropy measures, which were used in Hall and John (1994). To allow for cross-validation of the sample results, we divided the sample into two groups. Using the first group with 122 companies as calibration sample, we developed a discriminant function using DR and DU scores to assign companies to Rumelt's strategy categories. The hit matrix is shown in Table 3. The second group with 100 companies served as the holdout cross-validation sample. We used the discriminant function established by the first group to determine whether the SIC-based and skill-based entropy measures (DR and DU) of 100 companies in this group could effectively predict the companies into Rumelt's strategy categories.

The results of calibration sample and cross-validation sample are shown in the first two panels of Table 3. The overall percents correctly classified were 45.90% (SIC-based entropy measures) and 43.44% (skill-based entropy measures), which was lower than the result (56%) obtained by Hall and John (1994), but still, they are significantly higher than the largest chance hit ratio of 37.91% at alpha = 0.01. The results of the validation group are different from the result of Hall and John (1994). The result of Hall and John (1994) showed no significant difference between the hit ratio and the largest hit ratio. However, the hit ratios of our validation groups are 48% (SIC-based) and 43% (skill-based), significantly higher than the largest chance hit ratio of 41.25% at alpha = 0.01 level. These results show that, in the Chinese context, we can classify a company into Rumelt's strategy categories more effectively by using its SIC-based and Skill-based entropy measures. In Table 3, we also found that the hit ratios of both SIC-based and skill-based entropy indices predict single business, dominant- vertical and unrelated business much more effectively than they predict dominant and related business. It is also in consistent with the result of Hall and John (1994). In addition, the hit ratio of SIC-based entropy index is slightly higher than that of Skill-based entropy index (48% and 43%, respectively). This shows that the SIC-based entropy measures may have higher predictability for Rumelt's strategy categories than the skill-based entropy measures do.

Table 3 Results of Discriminant Analysis, Cross Validation and Random Replication [a]
Part A

Actual group	Calibration sample result										
	Designated group of SIC-based entropy indices (using DR_SIC and DU_SIC)					Designated group of skill-based entropy indices (using DR_Skill and DU_Skill)					
	1	2	3	4	5	1	2	3	4	5	Total
1. Single business	20 (90.9%)	0 (0%)	0 (0%)	0 (0%)	2 (9.1%)	20 (90.9%)	1 (4.5%)	0 (0%)	0 (0%)	1 (4.5%)	22
2. Dominant-vertical	11 (29.7%)	18 (48.6%)	0 (0%)	0 (0%)	8 (21.6%)	11 (29.7%)	15 (40.5%)	0 (0%)	0 (0%)	11 (29.7%)	37
3. Dominant business	4 (17.4%)	19 (82.6%)	0 (0%)	0 (0%)	0 (0%)	4 (17.4%)	18 (78.3%)	0 (0%)	0 (0%)	1 (4.3%)	23
4. Related business	3 (20%)	9 (60%)	0 (0%)	0 (0%)	3 (20%)	3 (20%)	8 (53.3%)	0 (0%)	0 (0%)	4 (26.7%)	15
5. Unrelated business	1 (4%)	6 (24%)	0 (0%)	0 (0%)	18 (72%)	1 (4%)	6 (24%)	0 (0%)	0 (0%)	18 (72%)	25
	The hit ratio is 45.90%, the largest chance hit ratio is 37.91%, Press's $Q = 51.16$ ($p < 0.001$)					The hit ratio is 43.44%, the largest chance hit ratio is 37.91%, Press's $Q = 41.90$ ($p < 0.001$)					

Actual group	Cross-validation sample result										
	Designated group of SIC-based entropy indices (using DR_SIC and DU_SIC)					Designated group of skill-based entropy indices (using DR_Skill and DU_Skill)					
	1	2	3	4	5	1	2	3	4	5	Total
1. Single business	17 (73.9%)	6 (26.1%)	0 (0%)	0 (0%)	0 (0%)	16 (69.6%)	7 (30.4%)	0 (0%)	0 (0%)	0 (0%)	23
2. Dominant-vertical	9 (27.3%)	17 (51.5%)	0 (0%)	0 (0%)	7 (21.2%)	9 (27.3%)	16 (48.5%)	0 (0%)	0 (0%)	8 (24.2%)	33
3. Dominant business	5 (31.3%)	8 (50%)	0 (0%)	0 (0%)	3 (18.8%)	5 (31.3%)	9 (56.3%)	0 (0%)	0 (0%)	2 (12.5%)	16
4. Related business	0 (0%)	8 (72.7%)	0 (0%)	0 (0%)	3 (27.3%)	0 (0%)	8 (72.7%)	0 (0%)	0 (0%)	3 (27.3%)	11
5. Unrelated business	0 (0%)	3 (17.6%)	0 (0%)	0 (0%)	14 (82.4%)	0 (0%)	6 (35.3%)	0 (0%)	0 (0%)	11 (64.7%)	17
	The hit ratio is 48%, the largest chance hit ratio is 41.25%, Press's $Q = 49.00$ ($p < 0.001$)					The hit ratio is 43%, the largest chance hit ratio is 41.25%, Press's $Q = 33.06$ ($p < 0.001$)					

Part B

	Result of random replications using DR_SIC and DU_SIC)						Result of random replications using DR_Skill and DU_Skill)					
	Calibration sample			Cross-validation sample			Calibration sample			Cross-validation sample		
	Hit ratio (%)	Press's *Q*	Significance [a]	Hit ratio (%)	Press's *Q*	Significance [a]	Hit ratio (%)	Press's *Q*	Significance [a]	Hit ratio (%)	Press's *Q*	Significance [a]
1	45.08	47.97	***	46.00	42.25	***	42.62	39.02	***	44.00	36.00	***
2	45.90	51.16	***	48.00	49.00	***	43.44	41.90	***	43.00	33.06	***
3	44.26	44.89	***	47.00	45.56	***	41.80	36.25	***	47.00	45.56	***
4	45.08	47.97	***	47.00	45.56	***	44.26	44.89	***	45.00	39.06	***
5	48.36	61.33	***	43.00	33.06	***	46.72	54.44	***	43.00	33.06	***
6	44.26	44.89	***	44.00	36.00	***	45.08	47.97	***	45.00	39.06	***
Mean	45.49			45.83			43.99			44.50		

Note: a. Compared with the largest chance hit ratio. *** indicates significant at level 0.01.

To guarantee the credibility of the above conclusion about the effects of the cross-validation sample on the interpretation of the original holdout sample, we conducted several replications (Frank, Massy and Morrison, 1965). In each replication, the sample was split into two groups randomly: one serving as a calibration sample for development of a discriminant function and another for cross-validation of the results. The results are shown in the final panel of Table 3. In all six replications, the hit ratios are all significantly higher than the largest chance hit ratio. These results confirmed our conclusion: Under the Chinese context, the SIC-based entropy measures including DR_SIC and DU_SIC scores, and the skill-based entropy measures including DR_skill and DU_dkill scores, could be effectively used to predict Rumelt's strategy category membership.

4.2 Comparison of Continuous-to-Categorical Measures and Rumelt's Categorical Measures

The second objective of this study is to compare Rumelt's categorical measures with the categorical measures, which are transformed from the four types of continuous measurement including SIC-based entropy measures, SIC-based Herfindahl measures, skill-based entropy measures, and skill-based Herfindahl measures by using cluster analysis, to see whether the two categories are consistent with each other. Following the method proposed by Ketchen and Shook (1996), both hierarchical and non-hierarchical clustering methods were adopted in turns. According to the recommendation of Hall and John (1994), we did not include the 70 dominant-vertical samples when using the 4-group Rumelt's topology to cluster Herfindahl measures (Table 4). Firstly, we conducted hierarchical cluster analysis with the method of the least distance between groups to classify 222 samples into 8 categories and to classify 152 samples into 4 categories. We used the mean of each variable in the eight and four categories as the initial center of the non-hierarchical cluster analysis to refine the results of the hierarchical cluster analysis. The results of cluster analysis are shown in Table 4.

For both SIC-based and skill-based entropy measures, we clustered the 222 samples into 8 categories and the 152 samples into 4 categories. Following the procedures suggested by Baysinger and Hoskisson (1989), Palepu (1989), and Hall and John (1994), we labeled the clustering results corresponding to Rumelt's eight types and four types respectively. For the results of the eight categories, the procedures were as follows: first, we named the group with the lowest mean of both DR and DU as "single business," and the one with the highest DU mean and the lowest DR mean as "unrelated business"; second, we named the group with

Table 4 Strategy Categories Resulting from Cluster Analysis of Continuous Measures for Diversification

	Cluster analysis of SIC-based entropy measures [a]			Cluster analysis of SIC-based Herfindahl measures [b]				Cluster analysis of skill-based entropy measures [c]			Cluster analysis of skill-based Herfindahl measures [c]
Strategy category	*N*	DR_SIC	DU_SIC	*N*	H_SIC3	*N*	H_SIC2	*N*	DR_Skill	DU_Skill	*N*	H_Skill
1. Single business [d]	89	0.005 7 (0.028 74)	0.062 3 (0.101 26)					94	0.008 5 (0.037 28)	0.075 9 (0.116 83)		
2. Dominant-vertical	56	0.013 9 (0.038 86)	0.565 2 (0.126 28)					61	0.017 0 (0.044 52)	0.634 8 (0.138 20)		
3. Dominant-constrained	27	0.562 4 (0.122 62)	0.110 6 (0.123 31)					20	0.525 1 (0.138 71)	0.034 5 (0.069 78)		
4. Dominant-linked	6	0.376 8 (0.087 02)	0.913 0 (0.098 16)					11	0.440 5 (0.102 48)	0.526 2 (0.098 52)		
5. Dominant-unrelated	23	0.027 8 (0.062 68)	0.954 1 (0.115 45)					5	0.368 4 (0.090 60)	0.895 6 (0.114 44)		
6. Related-constrained	6	1.041 9 (0.141 97)	0.104 5 (0.085 74)					6	1.117 9 (0.096 55)	0.106 9 (0.081 02)		
7. Related-linked	8	0.407 2 (0.090 04)	0.515 3 (0.108 85)					10	0.679 8 (0.114 76)	0.287 0 (0.067 03)		
8. Unrelated business	7	0.000 0 (0.000 00)	1.373 1 (0.129 65)					15	0.028 9 (0.064 23)	1.170 7 (0.199 31)		
1. Single business [e]	65	0.007 8 (0.040 29)	0.067 4 (0.106 48)	59	0.021 8 (0.038 11)	72	0.023 1 (0.037 44)	67	0.006 7 (0.039 09)	0.076 8 (0.115 29)	78	0.027 9 (0.042 94)
2. Dominant business	45	0.035 1 (0.099 40)	0.605 0 (0.119 96)	27	0.265 3 (0.068 32)	28	0.229 8 (0.070 02)	47	0.049 1 (0.117 87)	0.638 2 (0.123 88)	29	0.258 1 (0.070 35)
3. Related business	18	0.670 1 (0.241 12)	0.150 6 (0.120 62)	44	0.488 1 (0.054 66)	38	0.479 0 (0.053 55)	23	0.719 4 (0.254 40)	0.171 2 (0.157 71)	34	0.483 6 (0.047 76)
4. Unrelated business	24	0.091 3 (0.163 87)	1.106 7 (0.202 02)	22	0.681 4 (0.061 94)	14	0.678 3 (0.066 85)	15	0.086 7 (0.164 54)	1.169 3 (0.200 16)	11	0.683 4 (0.068 03)

Note: a. Means are shown with standard deviations in parentheses. We calculated them by using the method suggested by Palepu (1985);
b. Calculated by using the method suggested by Montgomery (1982);
c. Following the method recommended by Farjoun (1998), we reclustered the 168 three-digit SIC industries into 33 skill-based groups using the distribution of 50 professions across the 168 three-digit SIC manufacturing industries. Then we computed skill-based entropy indices according to the method proposed by Palepu (1985) and Herfindahl index according to the method suggested by Montgomery (1982). H_Skill is the Herfindahl index fro 2-digit SIC industries;
d. *N*=222;
e. *N*=152. We did not include the 70 dominant-vertical samples in the cluster analysis since the existing studies have found no evidence to support the relationship between dominant-vertical group and Herfindahl measures.

the highest DU mean and low DR mean as "related-constrained," and the one with slightly higher DU mean and moderately high DR mean as "related-linked"; third, in the remaining groups, we named the group with the largest DU mean as "dominant-unrelated"; fourth, we named the group with higher DU mean in the remaining categories as "dominant-linked," and the one with high DR and lower DU mean as "dominant-constrained"; finally, the remaining group with low DR and moderately low DU mean as "dominant-vertical." For the results of four categories, the procedures were as follows: First, we named the group with the lowest DR and DU mean as "single business," and the one with the highest DU mean and low DR mean as "unrelated business"; second, we named the group with the highest DR mean and low DU mean as "related business," and the one with moderately low DU mean and low DR mean as "dominant business."

For the Herfindahl measures including H_SIC3, H_SIC2 and H_Skill, we only conducted cluster analysis on the 152 samples into eight categories following the suggestion of Hall and John (1994). The rule of naming the results corresponding to Rumelt's eight types is as follows: the group with the largest mean was named as "unrelated business"; the one with second largest mean as "related business," the one with the smallest mean as "single business," and the remaining category as "dominant business."

The results of cross tabulation of continuous-to-categorical measures and Rumelt's categorical measures are shown in Table 5, in which the chi-square values and Cramer's V of all the models are significant at the level near 0.000. These results are quite different from the conclusions drawn by Hall and John (1994). Results show that the continuous measurement of corporate diversification can well reflect Rumelt's categorical measures in the Chinese context. This is consistent with the conclusions obtained in the above discriminant analysis. In addition, for the 8 categories, SIC-based entropy measures were more convergent to Rumelt's categorical measures than skill-based entropy measures (the chi-square value and Cramer's V of the former are larger than those of the latter). This result is also consistent with the conclusion of the above discriminant analysis. For the 4 categories, three-digit SIC Herfindahl measure is the best indicator to predict Rumelt's four categories among other measures because it has the highest chi-square value and the highest Cramer's V. These results demonstrate that when converting the continuous measures into categorical measures using cluster analysis, both SIC-based entropy and Herfindahl measures could better reflect Rumelt's categorical measures than skill-based entropy and Herfindahl measures.

Table 5 Results of Cross Tabulation of Cluster Analysis from Table 4
Part A

Rumelt category	Cluster using SIC-based entropy indices (DR_SIC, DU_SIC)									Cluster using SIC-based Herfindahl indices	
	1	2	3	4	5	6	7	8	Total	3-digit SIC Herfindahl index	2-digit SIC Herfindahl index
1. Single business	40	2	2	0	1	0	0	0	45		
	88.9%	4.4%	4.4%	0%	2.2%	0%	0%	0%			
2. Dominant-vertical	27	15	12	1	8	3	4	0	70		
	38.6%	21.4%	17.1%	1.4%	11.4%	4.3%	5.7%	0%			
3. Dominant-constrained	7	1	3	0	0	0	0	0	11		
	63.6%	9.1%	27.3%	0%	0%	0%	0%	0%			
4. Dominant-linked	4	2	1	0	0	0	0	0	7		
	57.1%	28.6%	14.3%	0%	0%	0%	0%	0%			
5. Dominant-unrelated	5	13	1	0	1	0	1	0	21		
	23.8%	61.9%	4.8%	0%	4.8%	0%	4.8%	0%			
6. Related-constrained	3	7	1	0	1	1	0	0	13		
	23.1%	53.8%	7.7%	0%	7.7%	7.7%	0%	0%			
7. Related-linked	1	5	5	1	0	0	1	0	13		
	7.7%	38.5%	38.5%	7.7%	0%	0%	7.7%	0%			
8. Unrelated business	2	11	2	4	12	2	2	7	42		
	4.8%	26.2%	4.8%	9.5%	28.6%	4.8%	4.8%	16.7%			
Total	89	56	27	6	23	6	8	7	222		
	40.1%	25.2%	12.2%	2.7%	10.4%	2.7%	3.6%	3.2%			
Significance test	Chi-square = 160.755, df = 49, $p < 0.000$; Cramer's $V = 0.322$, $p < 0.000$										

(*To be continued*)

(*Continued*)

Rumelt category	Cluster using SIC-based entropy indices (DR_SIC, DU_SIC)					3-digit SIC Herfindahl index					2-digit SIC Herfindahl index				
	1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total
1. Single business	40	2	2	1	45	40	2	3	0	45	42	1	2	0	45
	88.9%	4.4%	4.4%	2.2%		88.9%	4.4%	6.7%	0%		93.3%	2.2%	4.4%	0%	
2. Dominant business	18	16	4	1	39	14	18	6	1	39	17	16	6	0	39
	46.2%	41.0%	10.3%	2.6%		35.9%	46.2%	15.4%	2.6%		43.6%	41.0%	15.4%	0%	
3. Related business	5	11	8	2	26	4	6	14	2	26	9	8	9	0	26
	19.2%	42.3%	30.8%	7.7%		15.4%	23.1%	53.8%	7.7%		34.6%	30.8%	34.6%	0%	
4. Unrelated business	2	16	4	20	42	1	1	21	19	42	4	3	21	14	42
	4.8%	38.1%	9.5%	47.6%		2.4%	2.4%	50.0%	45.2%		9.5%	7.1%	50.0%	33.3%	
Total	65	45	18	24	152	59	27	44	22	152	72	28	38	14	152
	42.8%	29.6%	11.8%	15.8%		38.8%	17.8%	28.9%	14.5%		47.4%	18.4%	25.0%	9.2%	
Significance test	Chi-square = 101.637, df = 9, $p < 0.000$; Cramer's $V = 0.472$, $p < 0.000$					Chi-square = 136.213, df = 9, $p < 0.000$; Cramer's $V = 0.547$, $p < 0.000$					Chi-square = 113.237, df = 9, $p < 0.000$; Cramer's $V = 0.498$, $p < 0.000$				

Part B

Rumelt category	Cluster using skill-based entropy indices (DR_Skill, DU_Skill)									Cluster using skill-based Herfindahl index
	1	2	3	4	5	6	7	8	Total	
1. Single business	40	3	2	0	0	0	0	0	45	
	88.9%	6.7%	4.4%	0%	0%	0%	0%	0%		
2. Dominant-vertical	28	20	9	5	1	1	4	2	70	
	40.0%	28.6%	12.9%	7.1%	1.4%	1.4%	5.7%	2.9%		
3. Dominant-constrained	7	1	3	0	0	0	0	0	11	
	63.6%	9.1%	27.3%	0%	0%	0%	0%	0%		
4. Dominant-linked	4	2	1	0	0	0	0	0	7	
	57.1%	28.6%	14.3%	0%	0%	0%	0%	0%		

(*To be continued*)

(*Continued*)

Rumelt category	Cluster using skill-based entropy indices (DR_Skill, DU_Skill)									Clusters using skill-based Herfindahl index				
5. Dominant-unrelated	8	10	1	2	0	0	0	0	21					
	38.1%	47.6%	4.8%	9.5%	0%	0%	0%	0%						
6. Related-constrained	4	5	2	0	0	1	0	1	13					
	30.8%	38.5%	15.4%	0%	0%	7.7%	0%	7.7%						
7. Related-linked	1	6	2	1	0	1	2	0	13					
	7.7%	46.2%	15.4%	7.7%	0%	7.7%	15.4%	0%						
8. Unrelated business	2	14	0	3	4	3	4	12	42					
	4.8%	33.3%	0%	7.1%	9.5%	7.1%	9.5%	28.6%						
Total	94	61	20	11	5	6	10	15	222					
	42.3%	27.5%	9.0%	5.0%	2.3%	2.7%	4.5%	6.8%						
Significance test	Chi-square = 143.752, df = 49, $p < 0.000$; Cramer's $V = 0.304$, $p < 0.000$													
	1	2	3	4	Total					1	2	3	4	Total
1. Single business	40	3	2	0	45					42	1	2	0	45
	88.9%	6.7%	4.4%	0%						93.3%	2.2%	4.4%	0%	
2. Dominant business	20	14	5	0	39					19	17	3	0	39
	51.3%	35.9%	12.8%	0%						48.7%	43.6%	7.7%	0%	
3. Related business	5	11	9	1	26					11	7	8	0	26
	19.2%	42.3%	34.6%	3.8%						42.3%	26.9%	30.8%	0%	
4. Unrelated business	2	19	7	14	42					6	4	21	11	42
	4.8%	45.2%	16.7%	33.3%						14.3%	9.5%	50.0%	26.2%	
Total	67	47	23	15	152					78	29	34	11	152
	44.1%	30.9%	15.1%	9.9%						51.3%	19.1%	22.4%	7.2%	
Significance test	Chi-square = 94.878, df = 9, $p < 0.000$; Cramer's $V = 0.456$, $p < 0.000$									Chi-square = 103.183, df = 9, $p < 0.000$; Cramer's $V = 0.476$, $p < 0.000$				

5 Conclusion and Discussion

The results of this research show that Rumelt's categorical measures and continuous measures are capturing the same construct of diversification to a large extent. The results of ANOVA show that the SIC-based entropy and Herfindahl measures and skill-based entropy and Herfindahl measures can be effectively differentiated among Rumelt's strategy categories. Entropy and Herfindahl measures can not only separate single business from unrelated business strategic types significantly, but also distinguish effectively related business from unrelated one.

And the results of Scheffe's test present that, for Rumelt's eight categories, the single business and unrelated business can be significantly distinguished from other categories by DT_SIC, DU_SIC and DU_Skill. However, the majority of categories can not be effectively distinguished from each other by DR_SIC and DR_Skill. The latter result is surprisingly similar to the view of Pitts and Hopkins (1982). SIC-based measures may be indeed appropriate for studying the differences at industry level between diversified and non-diversified companies, rather than the distinction between related and unrelated types of diversification. Therefore, what indices can better measure related diversification? Answers may lie in the following analyses.

For the five categories, the single business, dominant-vertical and related business can be effectively distinguished from each other by DR_SIC and DR_Skill. DR_Skill is better as it can also significantly distinguish single business from unrelated business. This, to some extent, indicates that skill-based diversification measurement may capture the relatedness between businesses better. Comparing with SIC-based measures, skill-based measures may capture more relatedness underlying intangible resources such as experiences, knowledge, and technology than physical resources. This might help to explain why skill-based measures are better in explaining related diversification. As pointed out by Nayyar (1992), researchers usually use potential and external relatedness to capture the relationships among businesses, while ignoring the actual and internal relatedness. In fact, potential relatedness usually deviates from the actual relatedness because of the obstacles from organization and the external environment in practice. Therefore, the actual and internal relatedness more accurately reflects the degree of corporate diversification than potential and external relatedness does. It is suggested that future research could use skill-based measures to capture the corporate diversification.

The results of discriminant analysis further confirmed our above conclusions. Rumelt's categorical measures can be effectively predicted by SIC-based and skill-based entropy measures respectively.

To answer the second question of this paper, we classified the companies into

8 categories or 4 categories using cluster analysis with the four types of continuous measurements. The cross-tabulation analysis shows that the continuous measures are convergent to Rumelt's strategy types very well. This result is inconsistent with Hall and John (1994), who pointed out that SIC-based entropy index and Herfindahl index can not reflect diversification types due to the failure of SIC code to capture the real relatedness among businesses. However, our conclusions show that the SIC code can reflect well the types of corporate diversification in the Chinese context. In addition, we added skill-based entropy and Herfindahl's measures into the analysis. Results show that these measures are convergent to Rumelt's strategy types quite well even though they are slight "inferior" to SIC-based measures.

To check the robustness of the conclusions, we replicated our studies using the data of sample companies in 2004 and 2005 respectively. All the conclusions are in line with each other. Therefore, in general, the continuous measures and the categorical measures do capture the same construct of diversification in the Chinese context.

However, it is worth noticing that: First, as the 4-digit SIC code in *The Guidelines of Industrial Classification of Listed Companies* refers actually to the 3-digit SIC code in *The Industrial Classification of National Economy (GB/T4754-2002)*, listed companies reported indeed the 3-digit SIC industry data. We calculated DT (total diversity) in this paper using the 3-digit SIC data rather than the 4-digit SIC data. It may lead to DT scores as a whole smaller than the would-be DT scores when using the 4-digit SIC data. This may bring problems to the analysis of DR. For example, the cluster analysis using SIC-based entropy measures (DR_SIC and DU_SIC) and skill-based entropy measures (DR_skill and DU_skill) may classify the companies that are truly related-constrained or related-linked into other types.

Second, we referred Rumelt's strategy types as the benchmark to explore the consistence between the continuous measures and Rumelt's categorical measures. The results show they are quite consistent with each other. The logic underlying our assumption is that Rumelt's categorical measures validly capture the construct of diversification. However, do Rumelt's categorical measures really reflect the concept of diversification, rather than other concepts? This question worth analyzing further in future studies. Future studies can design a nomological network with some relevant constructs, such as size, financial leverage and corporate performance, to explore the convergent validity of these different measures of corporate diversification, and the discriminant validity between these measures of diversification and other relevant constructs, and the criterion validity of these measures of diversification with corporate performance.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 70572049) and the Research Fund of Business School of Nanjing University.

References

Amit R, Livant J (1988a). A concept of conglomerate diversification. Journal of Management, 14(4): 593–604

Amit R, Livant J (1988b). Diversification strategies, business cycles and economic performance. Strategic Management Journal, 9(2): 99–110

Baysinger B, Hoskisson R E (1989). Diversification strategy and R&D intensity in multiproduct firms. Academy of Management Journal, 32(2): 310–332

Berry C H (1971). Corporate growth and diversification. Journal of Law and Economics, 14(2): 371–383

Bettis R A (1981). Performance differences in related and unrelated diversified firms. Strategic Management Journal, 2(4): 379–393

Boyd B K, Gove S, Hitt M A (2005). Consequences of measurement problems in strategic management research: The case of Amihud and Lev. Strategic Management Journal, 26(4): 367–375

Chartterjee S, Blocher J D (1992). Measurement of firm diversification: Is it robust? Academy of Management Journal, 35(4): 874–888

Christensen H K, Montgomery C A (1981). Corporate economic performance: Diversification strategy vs. market structure. Strategic Management Journal, 2(4): 327–343

Farjoun M (1998). The independent and joint effects of the skill and physical bases of relatedness in diversification. Strategic Management Journal, 19(7): 611–630

Frank R E, Massy W F, Morrison D G (1965). Bias in multiple discriminate analysis. Journal of Marketing Research, 2(3): 250–258

Gort M (1962). Diversification and Integration in American Industry. Princeton, NJ: Princeton University Press

Hall E H, John C H (1994). A methodological note on diversity measurement. Strategic Management Journal, 15(2): 153–168

Hoskisson R E, Hitt M A, Johnson R A, Moesel D D (1993). Construct validity of an objective (entropy) categorical measure of diversification strategy. Strategic Management Journal, 14(3): 215–235

Jacquemin A, Berry C (1979). Entropy measure of diversification and corporate growth. Journal of Industrial Economics, 27(4): 359–369

Ketchen D, Shook C (1996). The application of cluster analysis in strategic management research: An analysis and critique. Strategic Management Journal, 17(6): 441–458

Markham J W (1973). Conglomerate Enterprise and Economic Performance. Cambridge, MA: Harvard University Press

Markides C C (1995). Diversification, restructuring and economic performance. Strategic Management Journal, 16(2): 101–118

Mayer M, Whittington R (2003). Diversification in context: A cross-national and cross-temporal extension. Strategic Management Journal, 24(8): 773–781

McVey J S (1972). The industry diversification of multi-establishment manufacturing firms: A developmental study. Canadian Statistical Review, 47(1): 112–117

Montgomery C A (1982). The measurement of firm diversification: Some new empirical evidence. Academy of Management Journal, 25(2): 299–307

Montgomery C A (1985). Product-market diversification and market power. Academy of

Management Journal, 28(12): 789–798
Nayyar P R (1992). On the measurement of corporate diversification strategy: Evidence from large U.S. service firms. Strategic Management Journal, 13(3): 219–235
Palepu K (1985). Diversification strategy, profit performance and the entropy measure. Strategic Management Journal, 6(3): 239–255
Pitts R, Hopkins D (1982). Firm diversity: Conceptualization and measurement. Academy of Management Journal, 7(4): 620–629
Ramanujam V, Varadarajan P (1989). Research on corporate diversification: A synthesis. Strategic Management Journal, 10(6): 523–551
Rumelt R P (1974). Strategy, Structure and Economic Performance. Cambridge, MA: Harvard University Press
Rumelt R P (1982). Diversification strategy and profitability. Strategic Management Journal, 3(4): 359–369
Schwab D (1980). Construct validity in organizational behavior, in: Staw B M, Cummings L L (eds.), Research in Organizational Behavior, Vol.2. Greenwich, CT: JAI Press, 2–43
Varadarajan P, Ramanujam V (1987). Diversification and performance: A reexamination using a new two-dimensional conceptualization of diversity in firms. Academy of Management Journal, 30(2): 380–393
Venkatraman N, Grant J H (1986). Construct measurement in organizational strategy research: A critique and proposal. Academy of Management Review, 11(1): 71–87
Wood A (1971). Diversification Merger and Research Expenditures: A Review of Empirical Studies. Cambridge, MA: Harvard University Press
Wrigley L (1970). Divisional autonomy and diversification. Unpublished doctoral dissertation, Harvard Business School
洪道麟, 熊德华 (Hong Daoling, Xiong Dehua) (2006). 中国上市公司多元化与企业绩效分析 (The analysis of the relationship between diversification of China's listed companies and firm performance). 金融研究, (11): 33–43
黄海波 (Huang Haibo) (2007). 中国房地产类上市公司多元化与企业绩效的实证分析 (The empirical study of the relationship between diversification of China's listed real estate companies and firm performance). 西安交通大学学报(社会科学版), 27(1): 39–43
姜付秀，刘志彪, 陆正飞 (Jiang Fuxiu, Liu Zhibiao, Lu Zhengfei) (2006). 多元化经营、企业价值与收益波动研究——以中国上市公司为例的实证研究 (Diversification, firm value, and the return variance: An empirical study of China's listed companies). 财经问题研究, (11): 27–35
李敬 (Li Jing) (2002). 多元化战略 (Diversification Strategy). 上海：复旦大学出版社
李玲, 赵瑜纲 (Li Ling, Zhao Yugang) (1998). 中国上市公司多元化经营的实证研究 (An empirical study of diversification of China's listed companies). 证券市场导报, (5): 31–42
倪桂平, 张晖 (Ni Guiping, Zhang Hui) (2005). 上市公司的业绩、股权结构与多元化经营研究 (Firm performance, ownership structure, and diversification of China's listed companies). 山东工商学院学报, 19(2): 43–48
邱金辉, 候剑平 (Qiu Jinhui, Hou Jianping) (2006). 多元化对上市公司生产效率影响的实证研究 (An empirical study of the effect of diversification on the productivity of China's listed companies). 系统工程, 24(11): 85–89
邵军，刘志远 (Shao Jun, Liu Zhiyuan) (2006). 多元化战略对管理层激励的影响——来自中国资本市场的经验证据 (The effect of diversification strategy on motivating the manager: The evidence from China capital market). 当代经济科学, 28(2): 1–11
姚俊, 吕源, 蓝海林 (Yao Jun, Lü Yuan, Lan Hailin) (2004). 我国上市公司多元化与经济绩效关系的实证研究 (An empirical study of the relationship between diversification of

China's listed companies and the economic performance). 管理世界, (11): 117–125

尹义省 (Yin Yisheng) (1998). 中国大型企业多角化实证研究——兼与美国大公司比较分析 (An empirical study of diversification of China's large enterprises: Comparing with US). 管理工程学报, 12(3): 1–12

张卫国，袁芳，陈宇 (Zhang Weiguo, Yuan Fang, Chen Yu) (2002). 上市公司多元化战略与经济绩效关系实证分析 (An empirical study of the relationship between diversification strategy and economic performance of China's listed companies). 重庆大学学报, 25(11): 135–139

朱江 (Zhu Jiang) (1999). 我国上市公司的多元化战略和经营业绩 (Diversification strategy of China's listed companies and firm performance). 经济研究, (11): 54–61