SFB 649 Discussion Paper 2005-062

The Impact of Industry Classification Schemes on Financial Research

Christian Weiner*

* School of Business and Economics,
Humboldt-Universität zu Berlin,
Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

The Impact of Industry Classification Schemes on Financial Research*

Christian Weiner[†]

December 20, 2005

Abstract

This paper investigates industry classification systems. During the last 50 years there has been a considerable discussion of problems regarding the classification of economic data by industries. From my perspective, the central point of each classification is to determine a balance between aggregation of similar firms and differentiation between industries. This paper examines the structure and content of industrial classification schemes and how they affect financial research. I use classification systems provided by the Worldscope and the Compustat database. First, this study gives a detailed description of the structure and methodology of industrial classification systems and the relevance in leading finance and accounting journals. Second, I construct a benchmark classification system to measure the performance of different systems and provide evidence that some systems a more homogeneous in terms of value drivers than others. Third, I examine how multiple valuation is influenced by industry classification and show that the results vary significantly for different systems.

*I would like to thank Roumiana Slavova for her support, Ernst Maug, Ingolf Dittmann and Niels Ulbricht for helpful discussions and comments. I gratefully acknowledge support by the Rudolf von Bennigsen-Foerder foundation and the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

[†]Humboldt-Universität zu Berlin, School of Business and Economics, Spandauer Str. 1, 10178 Berlin, Germany; e-mail: weiner@wiwi.hu-berlin.de

1 Introduction

In this paper I investigate industry classification systems and their implications for financial research. The main question is how these classification systems affect research results and why. This problem is especially of relevance because a significant fraction of empirical papers require classification systems. To provide a structured contribution I first examine the use of classification systems to show the importance of this paper. I further investigate six systems that are commonly used in financial research. Finally, I use these systems for a company valuation approach to show that the results are different and depend on the underlying classification system.

This paper provides several results. I find that on average 30% of the papers in the top 3 finance and the top 2 accounting journals use industry classification systems. The main purposes are sample restriction (34%), comparable company selection (31%) and detection of industry effects (12%). This paper also provides evidence that about 45% of the companies change their industry over time based on Worldscope SIC-codes (Standard Industrial Classification) and 25% based on Compustat, respectively. If Compustat GICS-codes (Global Industry Classification Standard) are used, 20% of the observations change their classification. Finally, I show that GICS and SIC are more homogeneous in terms of financial characteristics like profitability, leverage or market-to-book value than the other systems and that there is a high correspondence between SIC and GICS.

The main results of the empirical analysis are that the selection of comparable firms from the same industry is of critical importance. I find that there is a significant difference between the classification systems in terms of valuation accuracy. I further provide evidence that valuation accuracy can be improved by using not only industry membership for comparable selection but also other key financial ratios like profitability or size. I also document that the size of an industry has a high relevance for the selection of peer groups. Valuation errors increase if the industry is too small and also increase if the industry is too large.

This paper has the following structure: In section 2 I discuss the motivation for this analysis. Section 3 describes the research design. The next section summarizes the previous research in the field of industry classification systems. Section 5 presents an analysis how industry classifications are used in financial research. To do so, I cover the leading finance and accounting journals. This section also shows detailed descriptive statistics and the correspondence between different systems. In section 6 I present two financial applications. I show how different classification scopes affect the accuracy of multiple valuation. In the second application I construct a benchmark classification system based on value drivers. Based on this benchmark this study compares different systems in their ability to value firms with the multiple valuation approach. I also discuss the relevance of industry size for valuation accuracy. Section 7 concludes.

2 Motivation

Industry classification systems are used for several reasons in academic research and for a variety of practical purposes. Common applications are the selection of comparable companies, performance measurement or segment valuation. The main implication of classification schemes for companies is to combine entities that are as homogeneous as possible in terms of different financial, organizational or further characteristics depending on the underlying methodology of the classification system. This is the main reason why I am able to detect significant differences between several system but also between different data vendors. For instance, if one system uses segment sales figures to allocate firms to industries and another system uses segment earnings, I would expect significant differences in the classification structure because earnings are - compared to sales - influenced by profitability, accounting standards and additional determinants. It is also possible that firms are assigned to different industries in the same classification system depending on the underlying data vendor and its methodology. In the Worldscope database the firm Texaco Inc. is classified in the SIC-system as Crude Petroleum and Natural Gas (code: 1311), while it is classified as Petroleum Refining (code: 2911) in Compustat. Obviously, the user gets completely different peer groups for his applications. One can also observe different distributions of firms across industry categories according to the level of detail. While some systems have a more uniform distribution over all categories, other systems are extremely skewed with few large and many small industries.

The selection and understanding of an appropriate classification scheme is of critical importance to reduce the possibility of systematical measurements errors or selectivity biases. For instance, the diversification discount[1] will typically be shown by the comparison between the value of all segments of a diversified firm and corresponding single-segment firms. This calculation is based on average multiples of comparable firms from the same industry. Depending on the chosen classification system and the level of detail the discount ranges between 18% and 0% and sometimes turns into a premium.

While there is a significant number of papers in the top finance and accounting journals that uses industry classification systems as an instrument, the number of publications that investigates classification systems is very low. One reason could be the fact that the Standard Industrial Classification (SIC) system from the U.S. Office of Management and Budget is the most popular classification structure with the longest history beginning in 1939 and the highest availability. If there exists only one underlying system, then research results are comparable despite potential measurement problems. I extend the analysis to a broader universe of classification systems, which gives the possibility to detect problems of one specific system. I also develop an independent benchmark classification system based on value drivers to measure and compare the performance of the other systems.

3 Research Design

In the first part of this study I use summary and univariate statistics to examine the characteristics of industry classification systems and to show the difference and the correspondence between these systems. This analysis is based on the Worldscope

[1]The empirical literature that explains the diversification discount includes papers by Berger and Ofek (1995), Lang and Stulz (1994), Villalonga (2004a, 2004b) and many other publications.

database from Thomson Financial with the Standard Industry Classification system (SIC), the Worldscope Industry Group system (WSIG), the Dow Jones Global Classification Standard system (DJGCS), and the Compustat database from Standard and Poor's with the Standard Industry Classification system (SIC), the North American Industrial Classification system (NAICS) as well as the General Industry Classification system (GICS). I will also cover the Fama and French classification system, which is a reclassified SIC system. I will not only present differences between several classification systems but also between different data sources and data vendor methodologies. To do so, I compare the SIC systems from Worldscope and Compustat. I will also document that Worldscope and Compustat classification systems are used by a significant fraction of empirical research studies.

The objective of this study is to point out that a critical appreciation of industry classification is of relevance for financial research. I present evidence indicating that the available classification systems differ greatly and systematically in terms of methodology, structure and content. In this context I especially raise three questions:

1. How are industry classification schemes used in financial and accounting research? I investigate the most important journals in finance and accounting and search for the main purposes of use. I also address the question how datasets are modified and selection algorithms are used by researchers in terms of industry membership. For instance, Alford (1992) and many other accounting and finance papers require a minimum sample size of at least 5 comparable firms per industry for their analysis.

2. How do firms change their industry membership over time? This part is of high relevance for time series investigations. One can see that a lot of studies require time series data to support their results. This can be critical if industry membership changes over time. Therefore I document the movement of firms from one industry to another.

3. How homogeneous are the covered classification systems in terms of several financial characteristics and what is the grade of correspondence between different systems? Because homogeneity is always related to one or more variables,

I examine the variation for different measures like profitability or size and different scopes of industry groups. I also document how different systems can be substituted. This part is especially relevant for the selection of comparable companies.

Industry classification systems are used by academics and practitioners for different purposes. Therefore, the second part of this paper will provide several empirical and practical contributions that support the high relevance of classification systems. These contributions are linked to the valuation of firms and can be summarized by three points:

1. The first contribution is an application that estimates firm values by using a multiple valuation approach. Here, one essential component is a classification system, where comparable firms can be selected. I present the valuation performance for each available classification system and compare the results between the systems.
2. I document that there is a tradeoff between the size of an industry group and the homogeneity on the one hand as well as the valuation accuracy on the other hand. This is of critical importance for the determination of an appropriate classification level.
3. I develop an alternative classification and benchmark scheme. For this purpose I use cluster analysis methods, which gives the possibility to develop firm portfolios that are homogeneous in terms of several financial measures like profitability, size or leverage. I can calibrate the model for specific group sizes and homogeneity.

4 Related Literature

There exists few literature in finance and accounting that examines industrial classification systems. Most studies that cover classification systems focus on SIC codes reported by Compustat or CRSP because they have the widest distribution and availability in financial research.

The purpose of the study of Clarke (1989) is to answer the question whether industries based on SIC codes are able to separate firms into homogeneous economic groups. He covers a period from 1975 to 1983 and uses the Compustat North America database. The investigation examines different classification scopes. He uses a regression approach and concludes that profit ratios, sales changes and stock price changes of companies cannot be well explained by the SIC industry structure. He further shows that SIC is more effective at dividing firms into coarser 1- or 2-digit groups than into finer 3- or 4-digit groups.

Guenther and Rosman (1994) compare Compustat and CRSP SIC codes and test for homogeneity within industry groups. They use stock returns to measure homogeneity and calculate the Pearson correlation coefficient between monthly stock returns for companies within the same 4-digit industry. The results indicate that Compustat SIC industries are more homogeneous than CRSP industries. They further document that industry classifications based on the 2-digit SIC code between the two databases agree in 62% of the cases. Moreover, they show that the intra-industry variance of 14 key financial ratios is lower and the correlation in stock returns is higher for Compustat industries than for CRSP industries.

Kahle and Walking (1996) show that 21% of the firms linked to Compustat and CRSP SIC codes have the same 4-digit code and 79% have the same 1-digit code. One major reason of the inaccuracy is that the primary SIC code data item is based on the current primary SIC code of a given firm although a large number of firms changes their primary SIC code over time. Thus, changes over time are only covered by CRSP historical codes but not by Compustat codes. About 24% of the firms change their 4-digit SIC code at least once over a period from 1974 to 1993. Since 1987 Compustat reports also the historical primary SIC code. They further point out that the Compustat SIC system is superior in explaining abnormal performance compared to CRSP and a comparison based on the 4-digit level is superior to the 2-digit level.

Fama and French (1997) develop their own classification system, which links the existing SIC classification codes to 48 industry groups. The companies within these groups are expected to have more similar risk characteristics. The primary focus of this paper is not the development of a new classification system. Fama and French show that the estimation errors of cost of equity are quite high due to uncertainty about true factor risk premiums and the best asset pricing model.

Fan and Lang (2000) measure relatedness and complementarity of firms by considering commodity flows from input-output data. They construct an alternative measure based on these flows and show that this measure better explains firm relatedness than common SIC-codes.

Bhojraj, Lee and Oler (2003) compare the 4 different classifications systems SIC, GICS, NAICS and Fama and French. They use descriptive statistics to compare each system and perform several regressions in their investigation. They show that the Global Industry Classification System provided by S&P and MSCI exceeds other systems by explaining stock returns and cross-sectional variations of different multiples. They provide two explanations for the superior performance of GICS. First, GICS is financial orientated and therefore better in explaining financial ratios. Second, this system is based on the individual assignment of companies to industry classes by financial specialists.

Summarized, the most papers investigate Compustat industry classifications and especially the SIC system. Therefore it might be possible that all results are biased due to the methodology of a single data provider. The most comprehensive study of classification systems comes from Bhojraj, Lee and Oler (2003) who cover four broadly available systems. The other papers only compare two systems or examine a single system.

5 Industrial Classification in Financial Research

Industrial classification schemes will be used for several reasons in theory and practice. In this section I document the relevance of company classification systems in finance and accounting research. I also document the correspondence and difference between the systems and show time series changes of company classifications. Finally, I present evidence that some classification systems are more homogeneous in terms of financial characteristics than others.

5.1 Description of Classification Systems

In this section I describe the structure, methodology and content of each classification system that is used in financial and accounting research. While SIC and NAICS only provide an industry structure developed by the U.S. Government that has to be linked to companies by data vendors, other systems like GICS, DJGCS, WSIG, Value Line and FF provide a structure, methodology and company individual classification information[2].

Standard Industrial Classification (SIC)

The Standard Industrial Classification system has been developed by the U.S. Office of Management and Budget. It is the widely-used classification system for researchers and practitioners. The system is based on a four digit code. The first digit covers 10 divisions like Mining or Manufacturing, the first two digits 81 major groups like Oil and Gas Extraction or Paper and Allied Products, the first three digits industry groups like Converted Paper, all four codes industries like Envelopes. Each company is linked to one specific code typically based on sales information from the largest segment.

North American Industry Classifications System (NAICS)

The North American Industry Classifications System is planned to be the successor of the SIC system. It covers 20 sectors like Information defined by the first two digits, 96 sub-sectors like Broadcasting and Telecommunications defined by the first three digits, 311 industry groups like Telecommunications defined by the first four digits, 721 industries like Wireless Telecommunications Carriers defined by the first five digits and 1170 country specific sub-industries like Paging defined by all six digits together.

[2]SIC: Standard Industrial Classification; NAICS: North American Industry Classifications System; GICS: Global Industry Classification Standard; DJGCS: Dow Jones Global Classification Standard; WSIG: Worldscope Industry Groups; FF: Fama and French

Global Industry Classification Standard (GICS)

The Global Industry Classification Standard is a development by Morgan Stanley Capital International (MSCI) and Standard & Poor's (S&P). The GICS system consists of 10 sectors like Energy or Financials, 23 industry groups like Oil and Gas or Insurance with 59 industries like Oil and Gas Drilling or Insurance Brokers and 122 sub-industries. The system links an eight-digit code to each company. Data are available from December 1994 for S&P 1500 companies, and from June 1999 for non-S&P companies. The classification of companies is primarily based on revenues but also on earnings and market perception. Diversified companies are members of separate industry groups or industries.

Dow Jones Global Classification Standard (DJGCS)

The Dow Jones Global Classification Standard provided by Dow Jones covers approximately 45,000 securities worldwide. Companies are classified into 10 general economic sectors like Financial or Consumer - Cyclical, 18 market sectors like Banks or Automobiles, 51 Industry Groups like Auto Parts and finally 89 sub-groups like Tires. The classification of individual companies is based on revenues from dominant lines of business. Worldscope provides current DJGCS data based on the sub-group level. This means that every company is linked to one sub-group that consists of three characters.

Worldscope Industry Groups (WSIG)

Thomson Financial provides a four digit numeric code system where each company is linked to one code based on the net sales or revenues figures. The first two digits represent one of 27 major industry groups. Major groups are for instance Aerospace, Automotive or Chemicals. The next two digits represent sub-groups that cover a more detailed industry classification within the major groups. The major group Financial contains the most sub-groups (12). The major group Beverages on the other hand has only three sub-groups. Diversified companies with no clear primary segment but several similar important segments have own sub-groups. Companies that cannot be linked to a major group are classified in the group Miscellaneous.

Value Line (VL)

Value Line is a comprehensive source of information and covers approximately 100 industries. The Value Line database contains fundamental data (both current and historical) on more than 7,500 publicly traded North American, European, and Asian firms. It includes hundreds of items on each firm, with balance sheet and income statement data. Companies are assigned to industries by sales information. Industries are for instance Tobacco or Medical Services. Some industries are separated into specialty and diversified classes. There exist no sub-categories.

Fama and French (FF)

Fama and French (1997) develop a classification system that links the existing SIC groups based on 4-digits to 48 industries. Their intention is not to develop a new classification structure. They are only interested in a manageable number of industries. Anyway, this classification system is used by many papers[3]. They further point out that differences in cost of capital across industry groupings are high and economically significant. Therefore industry membership should be an important determinant in cost of capital estimations.

As mentioned, companies are not classified by a public institution but by several data vendors. Thus, one can distinguish between organizations that provide the structure and methodology of classification systems and commercial organizations that link firms to these systems. Table 1 shows how developers of the methodology and data provider are related. Here, I cover only these databases that are used by most papers.

In the case of VL, GICS and WSIG the developer of the methodology and the data provider are the same company. All other classification systems provide only the structure that has to be filled by another party. The main databases that collect financial information are Compustat from Standard and Poor's, Worldscope from Thomson Financial, Value Line from ValueLine Inc. and CRSP from the Center for Research in Security Prices. Each data vendor can use his own company information and system interpretation to link companies to a classification structure. I cover all

[3]See for instance Lee, Myers and Swaminathan (1999), Gebhardt, Lee and Swaminathan (2001), Purnanandam and Swaminathan (2003), Loughran and Marietta-Westberg (2005) and many others.

Table 1: Provider of classification systems

This table shows how organizations that develop the methodology and structure and the data provider that link firms to the classification structure are related. SIC is Standard Industrial Classification, NAICS is North American Industry Classification System, WSIG is Worldscope Industry Groups, GICS is Global Industry Classification System, DJGCS is Dow Jones Global Classification Standard, VL is Value Line, FF is Fama and French. The X shows in which database which classification system is available. I present Compustat from Standard & Poor's, Worldscope from Thomson Financial, CRSP from the Center for Research in Security Prices and Value Line from Value Line Inc..

developer	system	database			
		Compustat	CRSP	Worldscope	Value Line
Government	SIC	X	X	X	
Government	NAICS	X			
Thomson Financial	WSIG			X	
S&P	GICS	X			
Dow Jones	DJGCS			X	
Value Line	VL				X
Fama, French	FF	X	X	X	

systems that are available in Compustat and Worldscope. Additionally, I use the Fama and French system which is derived from SIC codes.

5.2 The Use of Industry Classification Systems

There is a large number of different purposes, where a separation of companies into homogeneous entities is required. In this section I present a detailed descriptive analysis how industrial classification systems are used in the financial and accounting literature. I find a significant number of papers that use classification systems and conclude that most of these papers do not cover the problems and critical effects that are related to industry classification systems. For instance, it is an open question, what SIC level should be used. Some papers use the 4-digit level with few companies per group, other papers select larger classes and use 3- or 2-digit classifications and there are also papers that use mixed levels with a minimum class size. In this paper I argue that a better understanding of the subject can help to reduce biases and increase the accuracy of the results.

This approach investigates every article that has been published in four representative financial and accounting journals between 1995 and 2003. The main research questions of this section can be summarized by three points:

1. First, I ask whether industry classification system are required and used in each of the papers. If I find an indication that an industry classification systems is used, I identify the exact system and the underlying classification provider as well as the database used. I analyse and summarize this coverage over a period from 1995 to 2003 to show which systems are commonly used and which are not.

2. If I am able to detect that there is an explicit statement about the use of industry classification systems, I determine the main purpose of use. Similar purposes will be classified into several groups. I am able to identify 7 different purposes for the use of industry classification systems: comparable selection, sample restriction, industry dummies, industry effects, industry distribution diversification and international use. In some papers more than one purpose can be identified.

3. Finally, I will discuss the scope of classification (for instance, 4-digit, 3-digit, ...) and the motivation for this restriction.

In this analysis I cover all available articles from 1995 to 2003 that have been published in the following journals to examine the use of industrial classification systems in financial and accounting literature:

1. American Economic Review (AER)

2. Journal of Accounting and Economics (JAE)
3. Journal of Accounting Research (JAR)
4. Journal of Finance (JoF)
5. Journal of Financial Economics (JFE)

The total number of papers that have been published during the time period from 1995 to 2003 in the American Economic Review is 226, 249 in the Journal of Accounting and Economics, 235 in the Journal of Accounting Research, 493 in the Journal of Finance and 716 in the Journal of Financial Economics.

Table 2 summarizes how industrial classification systems are used in financial and accounting research. Panel A displays for each year the number of papers that use industry classification systems. The percentage of papers that contain industry classification systems related to all papers that have been published within this period are presented in parentheses. In Panel B I show the classification system and the database where it is drawn from. I present figures for each journal and pool results over all years. Values in parentheses display the percentage use of classification systems.

I document that the classification of companies into homogeneous groups has a strong influence in financial and accounting publications. In the Journal of Accounting and Economics 35% of the published papers require industry classifications on average. In 1996 and 2003 45% of the papers require industry classification. The Journal of Accounting Research shows that more than 50% of the papers require classification systems. In the finance journals 19% of the papers use classification systems on average. In the American Economic Review there are only few papers that deal with an empirical topic and use industry classification. In the Journal of Finance and Journal of Financial Economics the number of papers is much higher. Accounting papers provide much more empirical investigations. Therefore, the number of papers that require classification systems is higher than in finance, where I detect more theoretical papers.

Most papers that require industry classification for sample restriction use the highest classification level. Only about 15% select another level. For comparable

Table 2: The use of industry classification

This table provides an analysis how industry classification is used in the leading economics, finance and accounting journals. Panel A displays the total number of papers that use industry classification schemes. Percentage values of journals that use industry classification are in parentheses. 100% refers to the total number of papers published in this year and journal. AER is American Economic Review, JAE is Journal of Accounting and Economics, JAR is Journal of Accounting Research, JoF is Journal of Finance, JFE is Journal of Financial Economics. Panel B shows which systems and data vendors are used.

Panel A

year	journal				
	AER	JAE	JAR	JoF	JFE
1995	0(0%)	10(42%)	13(61%)	8(17%)	43(40%)
1996	3(2%)	13(45%)	15(65%)	8(15%)	42(50%)
1997	3(2%)	8(29%)	12(52%)	9(17%)	56(27%)
1998	2(1%)	6(43%)	18(49%)	9(18%)	54(28%)
1999	4(3%)	11(26%)	22(62%)	11(23%)	50(40%)
2000	1(1%)	12(40%)	19(51%)	3(5%)	51(31%)
2001	1(2%)	1(6%)	18(55%)	5(9%)	61(25%)
2002	2(2%)	7(44%)	12(38%)	11(15%)	58(26%)
2003	2(2%)	21(45%)	19(57%)	6(11%)	52(42%)

Panel B

code	database	journal				
		AER	JAE	JAR	JoF	JFE
SIC	Compst.	92%	96%	91%	85%	79%
	Worldsc.	0%	2%	3%	2%	6%
	CRSP	5%	1%	4%	3%	5%
NAICS	Compst.	0%	0%	0%	1%	1%
DJGCS	Worldsc.	0%	0%	0%	0%	0%
WSIG	Worldsc.	0%	0%	0%	0%	0%
GICS	Compst.	0%	0%	2%	2%	1%
VL	VL	1%	0%	1%	2%	4%
FF	FF	0%	0%	0%	2%	2%
others	others	2%	1%	1%	2%	2%

selection purposes most papers use a wide level. Regarding SIC, about 40% cover 3-digits, while 30% cover 2-digits. Overall, about 40% of the papers use a 3-digit level for SIC or a corresponding level for the other systems. 30% require the narrowest level, while 20% use the 2-digit or a corresponding level. Only 10% require the broadest 1-digit level.

Table 3 provides the purpose of company classification in financial and accounting research. Some papers use classification for more than one reason. Therefore values do not sum to 100% in every case. The first purpose for the use of industry classification systems is the selection of peer groups. These peer groups from the same industry are expected to have similar financial characteristics. The second purpose is sample restriction. A lot of papers exclude industries like banks or utilities because these companies have unusual financial characteristics. The third group of papers uses classification systems for the development of industry dummies in regression analysis. The fourth purpose is the analysis of industry effects and industry specific characteristics without a statistical methodology. The next purpose is based on the classification of segments. Several papers use segment industry classification to measure and value diversification. The last purpose covers all papers that use industry classification for international investigations. This is not a direct purpose but an interesting information. Some of the purposes does not fit the categories or the description in the paper is not clear. Therefore I have an additional category that collects all unusual or unknown purposes.

The results in table 3 show that the selection of comparable firms and the restriction of datasets are most often used purposes that require industry classification. Industries that are typically excluded are banking and financial institutions (more than 50% of the cases) and utilities (15%). Some papers also exclude service companies (10%) or consider only the manufacturing sector (15%). The second important task of industry classifications is the selection of comparable companies from the same industry. The main reason is that companies in the same industry are expected to have similar characteristics like size, profitability and so on. Most papers require at least 5 to 6 comparable companies and therefore use the 3- or 2-digit SIC level.

Table 3: The purpose of industry classification

This table reports the purposes of industry classification systems. The results are based on all papers in the covered journals from 1995 to 2003 that use industry classification schemes. One paper can have one or more than one purpose of use. Therefore the percentage values do not have to sum to 100%. AER is American Economic Review, JAE is Journal of Accounting and Economics, JAR is Journal of Accounting Research, JoF is Journal of Finance, JFE is Journal of Financial Economics. Avg shows the average.

Purpose	AER	JAE	JAR	JoF	JFE	avg
Comparable selection	34%	25%	16%	23%	42%	27%
Sample restriction	22%	39%	44%	46%	29%	40%
Industry dummies	11%	12%	10%	2%	9%	8%
Industry effects	15%	4%	16%	15%	14%	12%
Industry distribution	7%	6%	3%	5%	17%	8%
Diversification	2%	3%	1%	8%	2%	4%
International use	8%	10%	5%	5%	4%	6%
Other	9%	4%	5%	3%	3%	4%

Industry dummies are used in regression analysis to estimate industry effects. About 8% of the papers perform regressions with industry dummies to identify industry effects. A similar purpose is the coverage of industry effects. About 12% of the papers use classifications for this reason. Most of these papers show descriptive statistics of industry members. Another fraction of papers examine the distribution of companies over industries. Several papers investigate diversification of companies and especially whether there is a premium or discount. For this reason the value of each segment of a company will be estimated by single-segment companies from the same industry. The international use of industry classifications is not an original purpose. Nevertheless, I have a separate category because a significant fraction of these papers use Worldscope classification systems instead of Compustat SIC codes.

5.3 Data Selection

This section discusses the data and variables that are required for the following investigations. I cover only companies from the United States.

Compustat provides SIC, GICS and NAICS. The product line accounting for the largest percent of sales provided in the 10-K statements will determine the primary SIC code. For SIC I use the current SIC code (SIC) and the historic SIC code (SICH). The current SIC code is based on fiscal year end data from 2004. Historical SIC codes are available from 1985 to 2004. I can identify 136,325 company-year historic SIC codes and 21,000 companies with a current SIC code. Companies are classified to GICS segments by revenue and earnings based on annual reports and financial statements. For GICS I use the annual GICS code (SPGIC) which is available from 1994 to 2004 and can identify 63,338 company-years. Companies are classified to NAICS groups by sales information from annual reports. To identify NAICS classifications I use the current NAICS variable (NAICS) and the historical data item (NAICSH). I identify 19,217 company classification information and 68,484 company-year information, respectively.

From Worldscope I can extract SIC, Worlscope Industry Group and Dow Jones Industry Group. Companies are classified to SIC groups by the data item net sales or revenues (wc01001). SIC contains timeseries information (wc19506) and current data information (wc07021) from 2004. To get a comparable dataset I use a history that begins in 1985. It is possible to identify 101,016 company-year historic SIC codes and 15,117 companies with a current SIC code. The Worldscope IG classifies companies into groups by using the data item net sales or revenues (wc01001). For Worldscope IG I use current information (wc06011) from 2004. The database contains 15,399 companies with a valid value. DJGCS links companies to groups by using the data item net sales or revenues (wc01001). For Dow Jones codes (wc07040) the database provides current values for 9,779 companies.

Fama and French provide no original classification system. Therefore, these industry groups are derived from Compustat and Worldscope SIC-codes. To do so, I use the definition provided by Fama and French (1997), Appendix A and the Compustat and Worldscope timeseries SIC codes, respectively.

5.4 Classification Structure

The main purpose of this section is to examine the structure of different classification systems. The term "structure" refers to the distribution of companies over a classification system. Table 4 presents the general structure of the industrial classification systems that are available in the Worldscope and Compustat universe. For each system I show the horizontal statistics, which refers to the number of companies per industry group and the vertical information, which covers every available classification level. Besides the distribution of each system and level I also display the number of official categories (N-OF) published by the classification developer and the number of categories used and filled by the database vendor (N-DB). I present data for 2002 because structural changes of the classification systems over time could lead to biased results. The only requirement for the dataset is that the specific classification code for each company has to be available. I select all companies from the United States. Worldscope SIC consists of 4 levels and covers 8,711 companies in 2002. About 75% of the possible categories are filled with data on the two top levels, while almost 100% are used on the first and second level. The SIC system in Compustat has the same functional structure and covers 7,515 companies. The number of categories that is used by Compustat are below the Worldscope figures for all levels. The Industry Group system from Worldscope covers 11,075 companies. The number of possible categories and categories used is equal. The Dow Jones Global Classification System has data for 9,337 companies available in Worldscope. The number of official categories and categories used is equal. The Global Industry Classification System from Compustat contains 9,337 companies. Because the developer and the database vendor are the same, the number of official and categories used is equal. The North American Classification System available in Compustat covers with 19,217 the largest number of companies. The number of categories provided and categories used is similar. The Fama and French system based on Worldscope and Compustat is derived from the SIC system. It links the common SIC codes to

48 industries. Both databases fill all 48 industries with companies.

I will now observe the distribution of companies and have several implications for the use of the classification codes in financial research. The SIC code from Compustat, which is a widely-used classification system, shows especially on the narrowest level that the first quartile contains only groups with one or two companies. The median number of 3 companies indicates that there is a large fraction of small groups. For most applications the third SIC level can be recommended because the mean and median number of firms is large enough and the standard deviation within the group is relatively low. For Worldscope SIC the median is 8 and the mean is 26. This is a reasonable tradeoff between an appropriate size of each group and enough heterogeneity between different groups. The distribution of the Compustat SIC system shows different results. Here the fourth level has a median number of companies per industry of 8 and a mean of 17 which is similar to the third level in Worldscope. The number of groups (424) on the fourth level is about 30% higher than the Worldscope number of groups (327) on the third level. Worldscope's Industry Groups have an average number of companies of 63 on the first level. This number increase sharply to 426 on the second level. These two level are comparable with the 1- and 2-digit SIC level in terms of industry size and standard deviation. The Dow Jones system has only one level which is similar to the second SIC-level. Furthermore, the GICS system has a coarser structure than SIC. The number of categories on the first level is 123 which leads to an average number of 70 companies per industry. This figure increase to 933 on the last level, which is almost comparable to the SIC system. NAICS provides a structure that is similar to SIC on the first three levels. The Fama and French system has 48 categories with an average number of about 300 companies per group.

5.5 Correspondence between Classification Systems

I could show that the Compustat SIC system is commonly used for classification purposes but there are also other systems that could be chosen. The intention of this section is to identify classification systems that have a large concordance to the

Table 4: Distribution of companies

This table shows the distribution of companies for different classification systems. I cover all companies in the United States for the year 2002. System WSSIC is Worldscope SIC, CSSIC is Compustat SIC, WSIG is Worldscope Industry Group, DJGCS is Dow Jones Classification, GICS is Compustat Global Industry, NAICS is Compustat NAICS, WSFF is Worldscope Fama and French and CSFF is Compustat Fama and French. Lev is the observed level which always goes from the broadest scope to the narrowest, N-OF are the numbers of official categories published by the developer of the system, while N-DB are the numbers of categories with companies available in the database. The company columns report the univariate statistics for the number of companies per category.

system	lev	n	N-OF	N-DB	companies per category						
					min	P25	P75	max	med	mean	std
WSSIC	4	8711	1004	752	1	2	8	452	3	11	33
	3	8711	416	327	1	3	21	1050	8	26	80
	2	8711	72	72	1	18	119	1319	45	121	214
	1	8711	11	10	1	388	1579	1862	800	871	687
CSSIC	4	7515	1004	424	1	4	17	435	9	17	34
	3	7515	416	270	1	5	25	850	11.5	27	67
	2	7515	81	68	1	24	107	1099	48.5	110	174
	1	7515	11	9	99	369	1112	1919	807	835	561
WSIG	4	11075	170	170	1	9	46	1202	18	63	143
	2	11075	26	26	2	88	452	2199	169.5	426	634
DJGCS	3	9779	134	134	1	13	89	634	31.5	73	104
GICS	8	9337	123	123	1	19	84	834	41	70	99
	6	9337	59	59	1	39	226	844	96	148	150
	4	9337	23	23	4	216	497	990	308	389	261
	2	9337	10	10	298	357	1540	1920	819	933	633
NAICS	6	19217	1170	1010	1	2	13	1008	5	18	65
	5	19217	721	664	1	3	20	1008	7	28	88
	4	19217	311	311	1	7	50	1736	21	59	152
	3	19217	96	96	1	35	188	2099	88	198	349
	2	19217	20	9	48	232	1975	8334	572	2135	3030
WSFF	2	8711	48	48	1	22	143	1632	52	344	300
CSFF	2	7515	48	48	1	26	154	1544	50	301	284

Compustat SIC system which is the reference and anchor system. In this case concordance means that the structure of the reference and the corresponding system as well as the distribution of the firms over industries are similar or equal. Three reasons influence the concordance between two systems. First, the number of companies between the two systems due to several data sources, could be different. Second, the structure of the systems could be different. Third, the assignment to industries could deviate. A high concordance allows to give the recommendation to substitute the Compustat SIC system with a corresponding system. This is of relevance if the Compustat system is not available or one of the other systems has some features that are required by researchers.

Bhojraj, Lee and Oler (2003) show the concordance between four commonly used classification systems (Compustat SIC, Worldscope FamaFrench, Compustat GICS and Compustat NAICS). They display for each Compustat SIC industry the corresponding industry of the other systems, where the number of equal firms is maximal. For instance, the 2-digit industry 20 contains 38 firms while the corresponding industry of the NAICS system covers 30 of these firms. The correspondence is 79%. The lack of this approach is that it does not consider the overall distribution of firms but only one category with the highest number of firms. In some cases this process covers only a small fraction of firms because the distribution of firms over industries is almost equal. The second problem is that this approach compares classification systems with different numbers of firms. If one system contains 100 firms and the other system only 50 then the expected concordance is only 50%. This problem has not been reflected in the analysis.

I use another approach that covers all aspects of concordance. I take the Compustat SIC classification system as the reference system because it is most commonly used. Each industry within this systems will be called a reference industry. The system I want to compare is referred to as the corresponding system. For each 4-digit industry I select all firms within that industry and calculate a concordance measure for all other corresponding systems based on these firms. I measure concordance C between the Compustat SIC industry i and one corresponding system as

$$C_i = \sum_{ind=1}^{IND} \left(\frac{n_{ind}}{f_i} \right)^2, \tag{1}$$

where IND consists of all industries in the corresponding system from that I detect firms from my reference industry, n is the number of corresponding firms within an industry and f is the number of firms in the reference industry. I cover the intersection of companies between two systems to control for different numbers of companies. For instance, the 4-digit reference industry 1 contains 100 firms f. The corresponding system, I want to examine, has 5 industries IND, where these 100 firms are distributed. Firms that are not available in one of the systems will not be considered. The corresponding industry 1 contains 40 firms (n=40), industry 2 20, industry 3 20, industry 4 10 and industry 5 10, which adds up to 100. The concordance measure is now 0.26 $\left[\left(\frac{40}{100}^2 \right) + \left(\frac{20}{100}^2 \right) + \left(\frac{20}{100}^2 \right) + \left(\frac{10}{100}^2 \right) + \left(\frac{10}{100}^2 \right) = 0.26 \right]$. This is relatively low due to the high dispersion over industries. It is obvious, that the concordance measure ranges between 0 and 1. The reference industry always has a concordance measure of one.

This approach has one key feature. The measure takes into account the distribution of firms in an industry and gives more weight to industries with more matches. It approaches one when an industry consists of a large number of reference firms, while it is close to zero if firms are distributed over several industries and show a high dispersion.

Table 5 provides the results of the analysis. I display the average concordance measure for each year for the corresponding system. Worlscope SIC, IG and FF classification is based on 4 digits, Compustat FF is based on 4 digits, GICS is based on 8 digits, NAICS is based on 6 digits. I cover all firms that are available in both databases and have a valid Compustat SIC classification code. Data for the corresponding system are also required because missing values would also affect the concordance measure. The matching between Compustat and Worldscope firms is based on Cusip information.

The results of table 5 show that the concordance decreases over time. This result is valid for all corresponding systems. Worldscope SIC almost halves from 1990

Table 5: Concordance between systems

This table displays the average concordance measure (CON) over all industries between the Compustat SIC system and 6 corresponding systems (Worldscope SIC, Fama and French, DJGCS, WSIG and Compustat GICS, NAICS). For one specific 4-digit Compustat SIC industry this measure covers the distribution of underlying firms over the reference industry. The best concordance is one, while the worst is zero. The Compustat system itself has a concordance measure of one. The Fama and French classification system based on Compustat data has also the measure one because Fama and French (1997) convey the SIC system into their own system. Worldscope SIC and Fama and French Worldscope SIC (Worldscope FF) cover timeseries data from 1990. Compustat GICS contains timeseries data and begins in 1994. Compustat NAICS begins in 1997 and has static data, Worldscope DJGCS and Worldscope IG begin in 1990 and have static data. Ind refers to the number of Compustat SIC industries, where corresponding data are available. The matching between Compustat and Worldscope data is based on the Cusip identifies.

year	WS SIC		GICS		WS FF		NAICS		DJGCS		WSIG	
	ind	CON	ind	CON	ind	CON	ind	CON	ind	CON	ind	CON
1990	225	0.79			225	0.92			344	0.71	352	0.69
1991	327	0.67			327	0.85			359	0.70	361	0.69
1992	348	0.65			347	0.84			370	0.70	371	0.68
1992	356	0.61			355	0.82			368	0.67	369	0.65
1994	385	0.57	338	0.83	384	0.80			392	0.64	396	0.62
1995	389	0.54	373	0.83	389	0.78			390	0.63	395	0.61
1996	400	0.53	377	0.84	400	0.78			401	0.61	406	0.61
1997	403	0.50	390	0.82	403	0.76	396	0.79	404	0.59	408	0.59
1998	413	0.48	397	0.83	412	0.75	407	0.78	413	0.55	416	0.56
1999	419	0.49	421	0.64	419	0.75	414	0.78	418	0.54	422	0.56
2000	421	0.48	422	0.63	421	0.74	417	0.77	420	0.53	422	0.56
2001	420	0.46	422	0.63	420	0.73	416	0.77	421	0.53	423	0.55
2002	417	0.46	419	0.64	417	0.72	412	0.76	418	0.54	419	0.55

to 2002 with a strong decrease in 1991. On the other hand the number of valid industries doubles. GICS has a constant development from 1994 to 1998. Then there is a sharp decrease in 1999 from 0.83 to 0.64 and again a constant movement from 1999 to 2002. Worldscope FF has the best concordance but also decreases from 0.92 to 0.72 over all years. NAICS provides similar correspondence figures for all available years from 1997 to 2002. DJGCS shows a strong reduction from 0.71 in 1990 to 0.54 in 2002. WSIG provides a similar result. All systems lead to an increase of valid industries over time, which is caused by a general existence of new firms. The figures between the systems show that Worldscope FF, NAICS and GICS provide the best correspondence followed by Worldscope SIC, DJGSC and WSIG. The recommendation is that research based on Worldscope should use the Fama and French classification system, which is derived from the SIC system. In the Compustat universe researches should prefer NAICS to GICS.

5.6 Homogeneity

A further informative investigation is how homogenous firms within one industry group are and how this changes for different industry group levels. Homogeneity always depends on one or more measures that define the distance between two different objects. For this analysis I take six financial measures into account. These are return on assets (roa) defined by EBIT divided by total assets, leverage (lev) defined by total debt divided by market value of equity and book value of total debt, yearly stock return (sr) defined by the change between the year-end price in year t and year t-1, natural logarithm of total assets (ta), natural logarithm of sales (sa) and market to book value (mb) defined by market value of equity divided by book value of equity. For each industry group I calculate the standard deviation between all group members based on the selected variable and the median value of the standard deviation of all industry groups available on this level. The median is used to reduce the influence of outliers.

This investigation has two dimensions that are of relevance for researchers and practitioners. The first dimension shows the differences between industry levels within one classification system. It is an obvious fact that a higher level with a more inhomogeneous structure is associated with a higher industry variation but the increase from one level to the other might be quite different. The other dimension is the comparison of different classification systems. Some systems have a different industry variation by definition[4], while other systems have a different variation due to firm classification and data availability.

Table 6 presents the results of this analysis. I display the median variation measured as standard deviation for each classification system and level. The underlying dataset covers all companies in the United States for the year 2002. The requirement for each classification system is that the firm identifier, the year and the specific classification value are available. I further require that all six variables are available.

Obviously, the variance increases if one moves from a detailed industry level to a broader level. Worldscope SIC shows a small increase from the 4-digit to the 2-digit level for all variables, while the change from 2-digits to 1-digit is extraordinary high for return on assets, stock return and market to book value. If I compare this with the Compustat results it can be seen that the variation within industry groups is lower for almost all cases. For instance, return on assets have a standard deviation based on the 2-digit level that is close to 0.30, while it is 0.45 for Worldscope data. The 4-digit Industry Groups figures are comparable to the 3-digit SIC level for every variable, while the Dow Jones system can be related to the 2-digit SIC level. GICS has also a four-level structure but the variation within industries is different to the SIC system. The 8-digit GICS level is similar to the 2-digit Compustat SIC level and the 4-digit GICS level is similar to the 1-digit level but there is no correspondence for the 6- and 2-digit level. NAICS has a six-level structure. The most detailed NAICS level has the lowest within industry variation from all systems. The 4- to 2-digit levels are similar to the corresponding SIC levels. The two Fama and French systems are derived from SIC. Both system show significant differences for return on assets, stock returns and market to book value. The closest relation to the SIC system is based on the 2-digit level.

[4]For instance, Compustat SIC based on 4 digits has a much more detailed structure than Dow Jones industry groups.

Table 6: Variance analysis

This table displays the homogeneity of different industrial classification systems and different classification levels. Roa is return on assets, sr is the yearly return of the stock, lev is leverage, ta is the natural logarithm of total assets, sa is the natural logarithm of net sales and mb is market value of equity to book value of equity. The values represent the median of the standard deviation within each industry group for the year 2002.

code	level	roa	sr	lev	ta	sa	mb
Worldscope SIC	4	0.150	0.821	0.249	2.254	2.231	0.631
	3	0.217	1.144	0.273	2.322	2.338	0.850
	2	0.457	3.087	0.287	2.499	2.570	1.936
	1	9.526	5.952	0.290	2.890	2.873	8.838
Compustat SIC	4	0.143	0.844	0.247	2.181	2.151	0.648
	3	0.143	0.896	0.255	2.201	2.212	0.762
	2	0.306	1.546	0.279	2.365	2.355	1.271
	1	2.949	5.152	0.280	2.725	2.645	11.050
Industry Groups	4	0.234	1.306	0.271	2.445	2.458	0.832
	2	2.403	6.612	0.280	2.770	2.822	4.827
Dow Jones	3	0.500	2.313	0.266	2.535	2.587	1.420
GICS	8	0.446	1.521	0.270	2.396	2.416	1.545
	6	0.660	2.318	0.271	2.475	2.495	2.771
	4	2.987	3.641	0.273	2.555	2.723	5.594
	2	3.478	6.830	0.270	2.575	2.818	18.171
NAICS	6	0.117	0.689	0.229	2.089	2.041	0.569
	5	0.122	0.753	0.244	2.152	2.121	0.646
	4	0.155	0.941	0.260	2.176	2.209	0.816
	3	0.200	1.534	0.270	2.209	2.328	1.156
	2	0.486	3.833	0.270	2.373	2.489	1.843
Worldscope FF	2	0.865	5.647	0.276	2.627	2.736	3.729
	1	7.828	10.286	0.293	2.631	2.724	5.652
Compustat FF	2	0.539	2.213	0.277	2.499	2.526	2.096
	1	2.550	3.321	0.278	2.594	2.595	4.314

The six variables show different behaviors. Leverage has a low median variation within industries independently of the classification level. The range is between 0.229 and 0.293. Return on assets and stock return have a significantly larger dispersion over all systems.

Altogether, I find that several systems provide an analogy to the SIC system. This is of great relevance for substitutional reasons. For most systems one can observe a large increase of variation from the second-lowest to the lowest level.

5.7 Time Series Analysis

A large number of papers uses time series data for their investigations. The underlying industry classification system might also be time series or possibly static. In general, changes of the classification of a firm or the system structure will not be considered. This section examines the changes within industrial classification systems over time and shows that about 50% of the companies change their industry class. I cover the SIC-systems available in Worldscope and Compustat from 1994 to 2002 and the GICS-system available in Compustat from 1994 to 2002. I do not consider NAICS because the time period is too short to get reliable results.

If I observe some changes of the classification of companies then I have two possible causes. It is obvious that a company can move from one primary segment to another due to an increase or decrease of segment sales and assets, respectively. The second reason for a change of the class occurs if the data vendor or the classification developer changes the methodology. The second reason is of minor relevance. The last fundamental change of the SIC structure was in 1993 which is not in my research period. The GICS-system starts in 1994 and shows no significant change in the research period.

Table 7 presents the changes of firms between different industry groups over time. I cover a time period from 1994 to 2002 for SIC-codes and for GICS-codes. The first change will be covered from 1994 to 1995 the last from 2001 to 2002. I will have one requirement for this dataset. For every company industry membership information have to be available for each year from 1994 to 2002. I also include companies that have a shorter history. The final dataset for Worldscope SIC classification for the United States contains 34,353 company-years, for Compustat SIC classification 40,704 company-years and for Compustat GICS classification 37,053 company-years. I display the number of companies that never change their industry group and the distribution of companies that moved once or more between different industry groups. The results are presented for each SIC-level (4,3,2,1) and for each GICS-level (8,6,4,2).

Obviously, the number of firm movements decreases from the 4-digit SIC level to the 1-digit SIC level and from the 8-digit GICS level to the 2-digit GICS level, respectively. This shows the column "no change". For Worldscope SIC 54.12% of the companies do not change their industry based on the 4-digit level while 83.10% do not change on the 1-digit level. The distribution of changes within one system and between different systems is more interesting. Considering the 4-digit SIC level from Worldscope, I see that about 46% of the companies change their industry once or more. The Compustat data present quite different results. Here, the number of changes is only 25%. Based on the 3-digit level Worldscope shows that 66% of the observations remain constant. Compustat has a fraction of unchanged observations of 77%. While the first dataset shows an increase from the fourth to the third level of about 12%, Compustat has a lower increase of about 2%. The second SIC-level in Worldscope has about 74% of unchanged observations, while there are 80% in Compustat. The changes from one level to the next are 12% and 3%, respectively. On the fourth level both databases present similar figures of about 83% to 85%. One can conclude that the number of changes in Worldscope is higher and decreases by 10% per SIC-level. In Compustat the changes are much less and also the decrease from 75% to 85%. GICS from Cumpustat is different because I cover only 5 years. The number of classification changes based on the 8-digit level is 85%, while it is 92.4% on the 2-digit level.

Based on the results of this part I have several implications. First, Worldscope SIC shows a high fraction of changes on the 4-digit level. This fraction decrease by about 10% from one level to the next. To reduce the influence of changes it is recommended to use the 3- or 4- digit level. SIC and GICS from Compustat are more

Table 7: Firm classification movements

This table presents how firms change between industry groups within the SIC and GICS system over time. I use a time period from 1994 to 2002 for Worldscope and Compustat SIC-codes and also for Compustat GICS-codes. Worldscope SIC refers to code "19506", Compustat SIC to code "sich", Compstat GICS to code "spgic". The requirements for the datasets are that every firm has SIC or GICS codes for each year of the time period. All results are percentage values. I consider firms that change from one code to another and then change back to the previous as double changing firms. The number of observations counts firm-years.

system	level	#obs.	no change	number of changes			
				1	2	3	>4
Worldscope SIC	4	34353	54.12	25.12	13.32	5.23	2.21
	3	34353	65.62	18.83	10.63	3.62	1.30
	2	34353	74.28	14.19	8.21	2.42	0.90
	1	34353	83.10	9.64	5.64	1.21	0.41
Compustat SIC	4	40704	74.65	15.79	6.45	2.20	0.91
	3	40704	77.01	14.24	6.03	2.01	0.71
	2	40704	80.38	12.36	5.03	1.61	0.62
	1	40704	85.01	9.53	3.95	1.11	0.40
Compustat GICS	8	27053	81.51	15.26	2.81	0.30	0.10
	6	27053	86.62	10.91	2.03	0.31	0.13
	4	27053	90.71	7.71	1.42	0.16	0.00
	2	27053	92.40	6.28	1.21	0.11	0.00

robust in terms of industry changes. The difference of changes from the narrowest to the widest level is about 10%. Thus, the influence of changes is similar on each level.

6 Industrial Classification: Empirical Investigations

This section consists of empirical investigations related to industry classification systems. I will document the performance of different classification systems for some typical research questions in finance and accounting. In this section I perform two empirical investigations where industry classification systems are key essential components. I document - based on multiple valuation analysis - that the selection and use of a classification system is an important determinant. For instance, several multiple valuation papers like Alford (1992) require a minimum number of peer firms and use a flexible selection procedure from a narrow to a broad classification definition to determine comparable firms. He shows that the 3-digit SIC level on average provides the best multiple valuation accuracy. Further representative papers are from Villalonga (2004a) who uses the same method but at least 5 peer firms and Berger and Ofek who require at least 5 firms. Some papers use a fixed industry level or a fixed number of firms. Beatty, Riffe and Thompson (1999) require at least 20 firms per industry and cover the 3-digit SIC level. Lins and Servaes (1999) classify firms at the 2-digit SIC level and also require a minimum number of firms.

The first approach is based on a typical multiple valuation procedure. I test different sets of industry peer groups from a narrow to a wide focus to show the optimal scope for the selection of comparable firms. I use three common classification systems with timeseries data available in Worldscope and Compustat, respectively, to document similarities and differences. The assumption is that a narrow focus with many different industries leads to a high similarity in terms of financial characteristics. On the other hand small groups can be influenced by outliers. A broad definition of industries overcomes the problem of outliers but the firms within an industry become more inhomogeneous. I would expect that a tradeoff between these situations leads to the best results.

In the second approach I develop an artificial classification system used as a benchmark. This system combines similar firms in terms of value drivers like profitability, leverage, growth and firm size. To do so, I use the statistical method of cluster analysis. Based on multiple valuation I test whether this system can forecast firm values in a better way than other classification systems. This part also documents the individual valuation performance and deviation for each system from the benchmark. The idea behind this methodology is that value driver should be superior in explaining cross-sectional differences in firm valuation than pure industry membership. It should also be possible to detect advantages for all systems where financial analysts assign companies to industry groups. The reason is that value drivers should implicitly be covered by financial analysts.

6.1 Multiple Valuation

Multiple valuation is commonly used in corporate finance and accounting. The typical approach is to select comparable companies from the same industry. The underlying assumption is that these firms share the same risk, profitability and accounting methods[5]. While most papers try to improve valuation accuracy by using different multiples or methods, I focus on different classification systems and scopes of industry definition to ask whether there is a tradeoff between the number of firms and homogeneity within one industry.

Data

This study uses three industry classification systems where timeseries data are available. Compustat provides data for SIC and GICS, Worldscope provides data for SIC. Additionally, I require market and accounting data from both databases. The sample covers firms from the United States for all years from 1990 and 2002 for the SIC systems and from 1994 to 2002 for the GICS system. I build one separate dataset for each classification system. The datasets have the following restrictions: The primary identifier and the year have to be available in Worldscope. Cusip and

[5] See Alford (1992).

year have to be available in Compustat. The accounting data total debt, EBIT (earnings before interest and taxes) and total assets have to be available for both databases. Market value of equity has to be available. For the Compustat and Worldscope SIC dataset a timeseries SIC-code has to be available. If this is not available I use the static SIC-code. I exclude datasets where no SIC-code is available, where the code has not 4-digits and where the SIC-code has the number 9999 (non-classifiable establishments) or 6000 to 6999 (financial industry). SIC code 9999 covers for instance firms without any operations, for financial industries I cannot compute EBIT without problems. For the GICS dataset I require that the code is available and it consists of 8-digits. The fiscal year end for all three datasets is the calendar year. Market price will be observed on the last trading day of April in year t+1. The company has only one type of stocks. After elimination through the restrictions 35,807 firm-year observations remain in the Compustat SIC sample, 50,009 in the Worldscope SIC sample and 18,110 in the Compustat GICS sample. If I compare the three samples from 1994 to 2002, I see that the number of observations for the SIC datasets is higher for Worldscope (41,209) than Compustat (29,130), while the number of observations in the GICS dataset (18,110) is lower.

Methodology

The goal is to provide a general indication from what scope of industry definition comparable firms should be selected. To invest whether one level of classification is better than the other, I compare the valuation accuracy based on each level. Each of the three classification systems has 4 classification levels. For every firm in each dataset I estimate 4 enterprise values based on peer groups with equal 4-digit-, 3-digit-, 2-digit- and 1-digit-SIC codes, respectively. For GICS I use 8-digits, 6-digits, 4-digits and 2-digits to determine peer groups. The estimation for firm i's enterprise value $\hat{EV_i}$ is given by

$$\hat{EV_i} = \left[median_{j \in C_i} \left[\frac{EV_j}{EBIT_j} \right] * EBIT_i \right], \tag{2}$$

where EV_j is the enterprise value of firm j defined by the sum of market value and total debt, $EBIT_j$ is EBIT for firm j and C_i is the set of comparable firms based on the underlying classification level used for valuing firm i. I use the median to average comparable firms because I want to control for outliers. I calculate the absolute prediction error APE_i for firm i by

$$APE_i = \left| \frac{\hat{EV_i} - EV_i}{EV_i} \right|, \tag{3}$$

where $\hat{EV_i}$ is the estimated enterprise value for firm i and EV_i is the observed market value for firm i.

Valuation results

Table 8 presents the results of the analysis. I want to find the classification level that determines the optimal number of peer firms for valuation. For practitioners it is useful to know, which classification level provides the lowest forecast errors on average. Therefore, I count all firms with the lowest estimation error for one classification system. E.g. in 2002 for 970 firms the optimal estimation peer group is based on the 4-digit Worldscope SIC-code, while for 1254 firms the best peer group is based on the 3-digit SIC code. 2-digit peer groups show the worst performance with 771 firms. For this example I recommend a selection of peer firms based on the 3-digit SIC-level. Additionally, I present the average valuation errors for each year and level. The table displays all periods from 1990 to 2002 for SIC and 1994 to 2002 for GICS to document time effects.

The figures in table 8 show some clear patterns. Based on Worldscope SIC peer group selection the best prediction can be obtained with the 3-digit SIC code. Only in 1990 and 2000 the 2-digit and 1-digit code, respectively, are better. The 4-digit and 1-digit level shows a similar performance from 1990 to 1994, thereafter I detect more accurate predictions based on the 4-digit level. Only in 2000 and 2001 the 1-digit code exceeds the 4-digit code. I see similar results for Compustat SIC but the results are more significant. In every year except 1994 the 3-digit level shows the best performance. The second best level is based on the 2-digit code, while the 1- and 4-digit level show a weak performance. GICS is the third classification system, where I have timeseries data. From 2000 to 2002 the 4-digit level should be preferred, while from 1995 to 1999 the 6-digit level shows a better accuracy. While there is a

Table 8: Determination of peer groups and valuation

This table provides the results of a valuation procedure. For all firms in the sample the valuation error will be calculated based on peer groups from different classification levels. For each firm the peer group focus with the best prediction accuracy will be counted. For Compustat and Worldscope SIC I use all peer groups from 4 to 1 digit. For GICS I use all peer groups from 8 to 2 digits. The table counts all firms where the valuation accuracy has a minimum. For instance, in 1990 for 348 firms the best valuation result is based on peer groups chosen from the 4-digit SIC level of the Worldscope database. The second row of each year displays the average valuation error.

year	Worldscope SIC				Compustat SIC				GICS			
	4	3	2	1	4	3	2	1	8	6	4	2
1990	348	571	586	434	328	576	329	183				
	0.41	0.40	0.38	0.39	0.43	0.38	0.40	0.46				
1991	494	664	447	472	364	529	443	202				
	0.44	0.40	0.40	0.43	0.44	0.39	0.40	0.45				
1992	599	630	483	556	440	729	360	156				
	0.42	0.39	0.40	0.39	0.44	0.39	0.41	0.44				
1993	596	806	548	566	482	863	474	219				
	0.41	0.38	0.39	0.41	0.45	0.35	0.38	0.43				
1994	822	1393	907	863	489	1102	1486	213	398	437	455	194
	0.35	0.33	0.34	0.37	0.39	0.39	0.38	0.38	0.37	0.33	0.34	0.37
1995	1122	1318	981	991	597	1157	1020	263	501	592	526	261
	0.41	0.38	0.39	0.42	0.40	0.36	0.39	0.40	0.38	0.37	0.41	0.39
1996	1365	1369	1043	1048	640	1187	1171	302	531	661	517	380
	0.36	0.36	0.37	0.47	0.40	0.35	0.36	0.41	0.38	0.37	0.38	0.40
1997	1324	1329	1192	1269	666	1262	1005	303	545	652	537	391
	0.39	0.37	0.36	0.37	0.42	0.36	0.37	0.43	0.37	0.36	0.37	0.40
1998	1426	1546	1197	1151	648	1163	918	400	566	726	612	373
	0.43	0.43	0.44	0.44	0.45	0.39	0.39	0.44	0.41	0.39	0.40	0.44
1999	1235	1428	1201	1203	703	1450	997	348	603	818	636	379
	0.49	0.46	0.48	0.50	0.48	0.40	0.41	0.47	0.44	0.41	0.42	0.45
2000	859	1443	844	1457	605	1372	820	477	321	579	783	372
	0.44	0.40	0.44	0.40	0.44	0.39	0.41	0.44	0.42	0.40	0.38	0.45
2001	770	1248	800	1119	565	1296	871	298	307	547	671	329
	0.41	0.38	0.42	0.36	0.40	0.37	0.37	0.44	0.39	0.37	0.37	0.41
2002	970	1254	771	951	654	1253	859	570	342	573	669	326
	0.38	0.36	0.38	0.37	0.42	0.37	0.39	0.40	0.39	0.37	0.36	0.40

clear strategy for SIC to use 3-digit peer groups, peer groups based on GICS can be selected from the 6- or 4-digit level.

I observe that the size and structure of a peer group have a significant influence on the valuation accuracy. If I use a very detailed level like 4-digit for SIC and 8-digit for GICS, the number of peer firms is very small, which leads to high valuation errors independently of the valuation method. On the other hand, if I use the 1- or 2-digit level the peer firms become relatively inhomogeneous, which leads to a high dispersion of multiples and also to a low valuation accuracy. The medium levels - 3-digit for SIC and 6- or 4-digit for GICS - generate a tradeoff between peer group size and homogeneity.

6.2 The Cluster Analysis Approach

I use the statistical method of cluster analysis to develop an independent classification scheme. This scheme will be used as a benchmark for the other classification systems. With this method I am able to calibrate the benchmark system. It is possible to determine the average size of groups and the homogeneity in terms of different financial measures. The idea is that matching on value drivers is a better way to identify comparable firms and to value firms with peer groups.[6] A portfolio of groups with similar firms in terms of value drivers should outperform a portfolio of pure industry matches.

Cluster analysis provides a useful statistical tool for the development of a new classification scheme for companies. Ketchen, Jr. and Shook (1996) analyze the use of cluster analysis in the field of strategic management. They cover 45 published studies. They point out that there is large space for these kinds of statistical methods. One of their main critics is the lack of validity. To ensure the validity of the results a conceptual basis for each of the examined research problems is required.

[6] Alford (1992) shows that combinations of return on equity and total assets are effective criteria for the selection comparable firms. Bhojraj and Lee (2002) demonstrate that a combination of industry membership with total assets and further value drivers produce a better valuation accuracy than industry membership alone.

Cluster analysis is a rather loose collection of different statistical methods that is used to assign single entities to specific groups. The main principle of cluster analysis techniques can be summarized in two basic formulations of this approach: First, the members within one cluster or group should be as close as possible together, which means in the case of this study that one cluster contains similar companies depending on the chosen value drivers and characteristics. Second, the distance or difference between the clusters should be as large as possible. To determine the relation between companies and clusters a distance function is required that measures the degree of correspondence between the characteristics of companies.

At first glance, cluster analysis is one of the techniques that seems to be closely related to classification schemes and the development of peer groups. Typically peer groups for company valuation are implicitly created by companies from the same industry. These companies often are similar in terms of characteristics like future earnings, profitability, risks or company size. But in some cases the dispersion within one industry is high. Cluster analysis on the other hand provides the possibility to use a method with a theoretical background that can be adjusted in various ways. Especially large samples of data and different combinations of companies can be grouped by an efficient procedure. Because of the huge number of possible selection criteria, this process will focus on common variables that have been proved to be successful in explaining company values.

Because it is not obvious, whether there is one similar group structure that holds for all company characteristics, different combinations of clustering algorithms and distance functions can be used to find the best way. Three different approaches are supposable:

1. Every company determines one single cluster in the beginning of the process. The fusion of these clusters based on some specific variables on different stages leads to bigger clusters with more companies inside. In subsequent steps, the two closest items and possibly some more are combined into a new aggregate cluster, reducing the number of clusters by at least one. Eventually all individual items are combined into one large cluster. The prediction error as one

central indicator for the validity of the method will rise so that there has to be a stop point, where the error is reasonable and the clusters have a homogeneous shape in terms of the number of companies. Additionally, other indicators for the optimal number of clusters can be used. In this context it has to be mentioned that there is not one optimal rule for the determination of the number of clusters. But one can see from the previous valuation analysis that there exists a significant tradeoff between the size of the industry peer group and the homogeneity within one group. This method is called an hierarchical approach.

2. One starts with some predetermined clusters based on industry membership or several other useful criteria. The first prediction error over all clusters can be calculated at this fixed stage. In the next step the companies will be changed between the clusters with the objective of an error reduction. It is possible to find a global minimum of the error value if all combinations are performed. One direct problem is that the number of clusters in this method will be unchanged at the end.

3. The divisive clustering method proceeds in the opposite direction, starting with one single cluster that contain all companies from the dataset and then in succeeding steps breaking this cluster apart into several separate clusters on the basis of their individual dissimilarity. This clustering method is very similar to the first one in terms of the statistical validity.

The cluster algorithm that is used in this study is based on the first approach and covers an agglomerative hierarchical method. The financial and statistical literature in this field provides some insights for this decision.

Gupta and Huefner (1972) examine the descriptive power of different key financial ratios to explain industry characteristics. They show that cluster analysis based on these ratios develops groups that correspondent highly with the observed industry characteristics. They use an hierarchical algorithm and define one cluster as one company in the beginning. Jensen (1971) uses cluster analysis based on an hierarchical approach with different financial ratios to develop a statistical classification technique that can be used for different purposes like performance comparison. He concludes from the results that cluster analysis is an appropriate statistical method to classify companies.

At the starting-point every company by itself forms one cluster with m characteristics. To control for different scalings between characteristics I use a z-transformation to normalize data. The distance between two different companies k and k' with several characteristics x within the multidimensional space R^m can be explained by the Euclidean distance, that is defined as

$$d_{k,k'} = \left[\sum_{j=1}^{m} (x_{kj} - x_{k'j})^2 \right]^{1/2} , \tag{4}$$

where $d_{k,k'}$ is one distance element of the m-dimensional distance matrix D, x_j is one characteristic and m is the total number of characteristics that are used. For this approach I apply the four variables: sales growth, return on assets, leverage and natural logarithm of total assets. During the clustering process single companies on the first stage and single companies or groups of companies on further stages are sequentially merged to new groups. At the end there is a disjunct structure of groups that contains all companies.

There are some types of hierarchical algorithm that differ in the use of the specific distance function. To limit the number of clustering methods it will only be focused on the procedures that seems to have the best performance to solve the grouping problem. Criteria that will have an impact are the shape of the clusters, the size, the dispersion and the impact of possible outliers. A lot of research covers these characteristics of different clustering procedures[7]. For valuation purposes an equal distribution of the number of companies within the clusters is useful. Therefore the clustering method of Ward (1963) will be used. This produces medium-sized and relatively similar clusters in terms of the number of members. A peer group that is not too small or too large shows the best valuation accuracy. On the other hand groups with high differences in terms of company members would reduce the advantages of this method for practitioners because it is harder to explain the combination of the companies.

[7]See for instance Jensen (1971), SAS/STAT User's Guide (2004), Jajuga, Sokolowski and Bock (2002).

Ward's method will minimize the absolute cluster internal variance and the sum of squares by merging two groups, respectively. The observed minimization criterion by comparison of two of all possible merged groups can be written by the following variance function:

$$V_g = \sum_{k=1}^{K_g} \sum_{j=1}^{m} (x_{kjg} - \overline{x}_{jg})^2, \tag{5}$$

where V_g is the variance of the new formed group, m contains the number of variables, x_{kjg} is the observed and transformed value of variable j=1...m from company k=1...K for all companies in group g, $\overline{x}_{jg}$ is the mean of the observed value x of the variable j in group g. Thus, the first sum covers the deviation for all characteristics from the corresponding mean characteristics for one company, and the second sum covers all companies within one group. The variance function will be minimized by the selection of one combination from all possible combinations on the specific stage.

Table 9 shows the distribution of firms over the developed groups of the classification approach. I use all firms that are available in Compustat and have a valid value for sales growth, return on assets, leverage and natural logarithm of total assets. The number of firms, number of different groups and the distribution of firms with these groups are reported. I develop separate firm portfolios for each year from 1990 to 2002.

From 1990 to 2002 the number of firms doubles. The ratio between the number of firms and number of groups stays almost constant. The average number of firms per group over all years is 13.64. An exception occurs in 1999 where the average number is 18. The minimum number of firms within the group ranges between 1 and 3. The maximum number is between 33 and 73. The standard deviation is also homogeneous and ranges between 7.2 and 11.6.

I test the valuation accuracy of several classification systems with the enterprise value to EBIT valuation method and use the artificial system as a benchmark. The multiple approach estimates the enterprise value of the firm by multiplying earnings with an enterprise value to EBIT multiple determined from a set of comparable companies. This method calculates the enterprise value to EBIT multiple for company

Table 9: Distribution of firms

This table reports the distribution of firms for a classification system based on value drivers and a cluster analysis approach. I show the total number of firms, the number of different groups and the distribution of firms within these groups from 1990 to 2002. Firm data for sales growth, return on assets, leverage and natural logarithm of total assets are from Compustat.

year	#firms	#groups	distribution of firms						
			mean	min	p25	med	p75	max	std
1990	1057	82	12.8	2	6	11	17	41	8.2
1991	1107	90	12.3	1	6	10	18	60	9.5
1992	1160	107	10.8	1	5	9	15	41	7.4
1993	1283	117	10.9	1	5	10	15	33	7.3
1994	1893	153	12.3	2	7	11	17	37	7.2
1995	2208	143	15.4	3	9	13	21	43	9.5
1996	2298	156	14.7	1	7	12.5	20	67	10.4
1997	2300	159	14.4	1	8	13	19	49	9.3
1998	2397	176	13.6	1	7	12	18	56	9.0
1999	2399	133	18.0	1	9	15	23	63	11.6
2000	2304	175	13.1	1	7	11	17	73	8.8
2001	2051	152	13.4	1	7	12	17	45	9.1
2002	2115	156	13.6	1	7	13	19	50	8.9

k as followed:

$$\frac{EV_k}{EBIT_k} = \frac{\text{market value}_k + \text{total debt}_k}{\text{earnings before interest and tax}_k}. \tag{6}$$

The estimation for firm k's enterprise value $\hat{EV_k}$ is given by

$$\hat{EV_k} = \left[median_{j \in C_k} \left[\frac{EV_j}{EBIT_j} \right] * EBIT_k \right], \tag{7}$$

where EV_j is the enterprise value of firm j defined by the sum of market value and total debt, $EBIT_j$ is EBIT for firm j and C_k is the set of comparable firms based on the underlying classification system used for valuing firm k. For all systems except the benchmark system I require at least 5 comparable companies. To do so, I start with the most detailed classification level and increase the level until I have at least

5 firms. The highest level represents the whole market. For the benchmark system I use the predefined groups without adjustments.

The valuation accuracy will be calculated by the deviation between the estimated firm value and the real firm value. I calculate the absolute prediction error APE_i for firm k by

$$APE_k = \left| \frac{\hat{EV_k} - EV_k}{EV_k} \right|, \tag{8}$$

where $\hat{EV_k}$ is the estimated enterprise value for firm k and EV_k is the observed market value for firm k.

Table 10 displays the median absolute valuation error for each classification system from 1990 to 2002. Compustat and Worldscope SIC are available from 1990 to 2002. GICS is available from 1994 to 2002. NAICS is available from 1997 to 2002. If the code is not available in the considered year I use the last available information. For GJGCS and WSIG timeseries data are not available. Therefore, I use the static code from 2002.

Table 10 documents that the benchmark system (bm) based on value driver firm portfolios outperforms other systems. This is an expected result because the firm value should depend primarily on value drivers and not on industry membership, which is only a proxy for value driver characteristics. The main information of this analysis is how close the different systems are related to the benchmark. I test the results by performing a t-test on the differences between the benchmark and the system itself. The SIC-based valuation with Worldscope data shows higher median errors in each year. One case shows a significant difference at the 1%-level, four at the 5%-level and four at the 10%-level. The other differences are negative but insignificant. The Compustat SIC-system shows similar results. The average difference between the Worldscope and Compustat SIC system is low. The largest deviation is in 2001 with 4%. In 7 out of 13 years the deviation between the two systems is only 1% and below. The Industry Group system from Worldscope has a similar background like the SIC system with 4-digit codes and a production-based categorization. The structure shows less groups. SIC has 760 4-digit categories, while WSIG has only

Table 10: Valuation accuracy

This table displays the median absolute valuation error for United States firms. Error is defined as the absolute value between the estimated enterprise value and the observed value. bm indicates the error for the benchmark system. The benchmark row for each year shows the difference between my system and the system in the column. A negative value indicates that a value driver based portfolio outperforms an industry classification system. I perform a t-test to show valuation differences. The last row shows the average error deviation from the benchmark system.

year	bm	Worldscope			Compustat		
		SIC	WSIG	DJGCS	SIC	NAICS	GICS
1990	0.24	0.25	0.31	0.30	0.26	0.28	0.24
benchmark		-0.01	-0.07***	-0.06**	-0.02	-0.04*	0.00
1991	0.23	0.25	0.32	0.30	0.25	0.27	0.28
benchmark		-0.02*	-0.09***	-0.07**	-0.02*	-0.04*	-0.05**
1992	0.21	0.24	0.30	0.28	0.23	0.24	0.24
benchmark		-0.03*	-0.09***	-0.07***	-0.02	-0.03*	-0.03*
1993	0.22	0.24	0.29	0.29	0.26	0.26	0.27
benchmark		-0.02	-0.07***	-0.07***	-0.04*	-0.04*	-0.05*
1994	0.21	0.21	0.29	0.27	0.22	0.23	0.23
benchmark		0.00	-0.08***	-0.06**	-0.01	-0.02	-0.02
1995	0.21	0.24	0.31	0.25	0.21	0.22	0.24
benchmark		-0.03**	-0.10***	-0.04	0.00	-0.01	-0.03**
1996	0.24	0.25	0.30	0.37	0.23	0.28	0.26
benchmark		-0.01	-0.06**	-0.13***	0.01	-0.04*	-0.02
1997	0.21	0.23	0.32	0.29	0.24	0.23	0.23
benchmark		-0.02*	-0.11***	-0.08***	-0.03*	-0.02	-0.02
1998	0.26	0.31	0.35	0.31	0.32	0.30	0.25
benchmark		-0.05**	-0.09***	-0.05	-0.06***	-0.04*	0.01*
1999	0.30	0.37	0.42	0.34	0.38	0.35	0.35
benchmark		-0.07***	-0.12***	-0.04*	-0.08***	-0.05**	-0.05**
2000	0.31	0.32	0.38	0.34	0.32	0.33	0.35
benchmark		-0.01*	-0.07**	-0.03	-0.01*	-0.02	-0.04**
2001	0.29	0.33	0.38	0.36	0.37	0.32	0.28
benchmark		-0.04**	-0.09***	-0.07**	-0.08***	-0.03*	0.01*
2002	0.29	0.32	0.36	0.36	0.34	0.30	0.30
benchmark		-0.03**	-0.07**	-0.07**	-0.05**	-0.01	-0.01
average		-2.5%	-8.5%	-6.5%	-3.2%	-3.0%	-2.3%

***, **, * indicates significance at 1%, 5%, 10% level

170 categories. This leads to peer groups that are more inhomogeneous. This is one explanation for the weak performance of this system. The average deviation from the benchmark is 8.5%. The highest difference is 0.12 in 1999. The third system that is provided by Worldscope is the Dow Jones Global Classification Standard. This system covers 89 industry groups. The valuation errors are similar to the Industry Group system. The average deviation from the benchmark is 8.5%. The North American Industry Classification system is the successor of the SIC-system. The average deviation over all years is 3.0%. This value is similar compared to Worldscope's and Compustat's SIC-systems. The Global Industrial Classification System has a strong performance. The average deviation from the benchmark is 2.3%. This is similar to the results provided by Bhojraj and Lee (2003). They compare this system with SIC and NAICS from Compustat as well as Fama and French and show that it is superior in explaining valuation multiples with industry peer groups.

The figures demonstrate a clear pattern over time. Between 1990 and 1997 I find similar valuation errors over time. The benchmark values range between 21% and 24%. In 1998 I see an increase of this error and a constant higher level between 26% and 29% with a peak of 31% in 2000. Higher valuation errors can be explained by the internet bubble and unusual stock market reactions. The other classification systems are coherent with the results of the benchmark systems.

Overall, I find no significant differences between the systems provided by the Worldscope and Compustat database. Both SIC methodologies are equal. The NAICS system which is the successor of the SIC system shows also similar values. The GICS system has also a good performance while WSIG and DJGCS are significantly below the benchmark.

7 Conclusion

This paper provides a comprehensive investigation of industry classification systems. I document that the use of industry classifications is high in accounting and corporate finance journals. About 30% of the papers published in leading finance and accounting journals require an industry classification system. The analysis of classification systems over time and the comparison between different systems provides some interesting insights. First, the popular classification systems SIC, GICS, FF and NIACS show a high variation in terms of methodology, structure, allocation of firms and concordance. Second, empirical investigations based on classification systems show significant differences depending on the underlying systems. I perform two empirical valuation procedures. I show that there exists a tradeoff between the size of an industry peer group and the homogeneity within one group. I also point out that GICS leads to lower valuation errors than the other systems. These results are robust over time. I also document that Worldscope and Compustat SIC produce similar results. The best results are based on an artificial classification system that is based on value driver combinations.

References

- Alford, Andrew W. (1992): The effect of the set of comparable firms on the accuracy of the price-earnings valuation method, Journal of Accounting Research, 30, pp. 94-108.
- Baker, Malcolm and Richard S. Ruback (1999): Estimating industry multiples, working paper, Harvard University.
- Beatty, Randolph P., Susan M. Riffe, and Rex Thompson (1999): The method of comparables and tax court valuations of private firms: An empirical investigation, Accounting Horizons, 13, pp. 177-199.
- Berger, Philip G., and Eli Ofek (1995): Diversification's effect on firm value, Journal of Financial Economics, 37, pp. 39-65.
- Bhojraj, Sanjeev, and Charles M. C. Lee (2002): Who is my peer? A valuation-based approach to the selection of comparable firms, Journal of Accounting Research, 40, pp. 407-439.
- Bhojraj, Sanjeev, Charles M. C. Lee, and Derek K. Oler (2003): What's my

line? A comparison of industry classification schemes for capital market research, Journal of Accounting Research, 41, pp. 745-774.

- Cheng, C. S. Agnes, and Ray McNamara (2000): The valuation accuracy of the price-earnings and price-book benchmark valuation methods, Review of Quantitative Finance and Accounting, 15, pp. 349-370.
- Clarke, Richard N. (1989): SIC as delineators of economic markets, The Journal of Business, 62, pp. 17-31.
- Fama, Eugene F., and Kenneth. R. French (1997): Industry costs of equity, Journal of Financial Economics, 43, pp. 153-193.
- Fan, Joseph P. H., and Larry H. P. Lang (2000): The measurement of relatedness: An application to corporate diversification, The Journal of Business, 73, pp. 629-660.
- Fertuck, Leonard (1975): A test of industry indices based on SIC codes, Journal of Financial and Quantitative Analysis, 10, pp. 837-848.
- Gebhardt, William R., Charles M. C. Lee, and Bhaskaran Swaminathan (2001): Toward an implied cost of capital, Journal of Accounting Research, 39, pp. 135-176.
- Gupta, Manak C., and Ronald J. Huefner (1972): A cluster analysis study of financial ratios and industry characteristics, Journal of Accounting Research, 10, pp. 77-95.
- Herrmann, Volker, and Frank Richter (2003): Pricing with performance-controlled multiples, Schmalenbach Business Review, 55, pp. 194-219.
- Jajuga, Krysztof, Andrzej Sokolowski, and Hans-Hermann Bock (2002): Classification, clustering and data analysis, Springer, Berlin.
- Jensen, Robert E. (1971): A cluster analysis study of financial performance of selected business firms, Accounting Review, 46, pp. 36-56.
- Kahle, Kathleen M., and Ralph A. Walking (1996): The impact of industry classifications on financial research, Journal of Financial and Quantitative Analysis, 31, pp. 309-335.

- Ketchen JR., David J., and Shook, Christopher L. (1996): The application of cluster analysis in strategic management research: an analysis and critique, Strategic Management Journal, 17, pp. 441-458.
- Krishnan, Jayanthi, and Eric Press (2003): The North American Industry Classification System and its implications for accounting research, Contemporary Accounting Research, 20, pp. 685-717.
- Loughran, Tim, and Jennifer Marietta-Westberg (2005): Divergence of opinion surrounding extreme events, European Financial Management, forthcoming.
- Kim, Moonchul, and Jay R. Ritter (1999): Valuing IPOs, Journal of Financial Economics, 53, pp. 409-437.
- Lee, Charles M. C., James N. Myers, and Bhaskaran Swaminathan (1999): What is the intrinsic value of the Dow?, Journal of Finance, 54, pp. 1693-1741.
- Lie, Erik, and Heidi J. Lie (2002): Multiples used to estimate corporate value, Financial Analysts Journal, 58, pp. 44-54.
- Lins, Karl, and Henri Servaes (1999): International evidence On The value Of corporate diversification, Journal of Finance, 54, pp. 2215-2239.
- Liu, Jing, Doron Nissim, and Jacob Thomas (2002a): Equity valuation using multiples, Journal of Accounting Research, 40, pp. 135-172.
- Liu, Jing, Doron Nissim, and Jacob Thomas (2002b): International equity valuation using multiples, working paper, University of California at Los Angeles.
- Purnanandam, Amiyatosh K., and Bhaskaran Swaminathan (2003): Are IPOs really underpriced?, Review of Financial Studies, 17, pp. 811-848.
- Ramnath, Sundaresh (2002): Investor and analyst reactions to earnings announcements of related firms: an empirical analysis, Journal of Accounting Research, 40, pp. 1351-1376.
- SAS/STAT 9.1 User's Guide (2004), SAS Institute Inc.
- Saunders, Norman C., (1999): The North American Industry Classification System: change on the horizon, Occupational Outlook Quarterly, pp. 34-37.

- Villalonga, Belen (2004): Does diversification cause the diversification discount?, Financial Management, 33, pp. 5-27.
- Ward, J.H. (1963): Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association, 58, pp. 236-244.

SFB 649 Discussion Paper Series

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Nonparametric Risk Management with Generalized Hyperbolic Distributions" by Ying Chen, Wolfgang Härdle and Seok-Oh Jeong, January 2005.
002 "Selecting Comparables for the Valuation of the European Firms" by Ingolf Dittmann and Christian Weiner, February 2005.
003 "Competitive Risk Sharing Contracts with One-sided Commitment" by Dirk Krueger and Harald Uhlig, February 2005.
004 "Value-at-Risk Calculations with Time Varying Copulae" by Enzo Giacomini and Wolfgang Härdle, February 2005.
005 "An Optimal Stopping Problem in a Diffusion-type Model with Delay" by Pavel V. Gapeev and Markus Reiß, February 2005.
006 "Conditional and Dynamic Convex Risk Measures" by Kai Detlefsen and Giacomo Scandolo, February 2005.
007 "Implied Trinomial Trees" by Pavel Čížek and Karel Komorád, February 2005.
008 "Stable Distributions" by Szymon Borak, Wolfgang Härdle and Rafal Weron, February 2005.
009 "Predicting Bankruptcy with Support Vector Machines" by Wolfgang Härdle, Rouslan A. Moro and Dorothea Schäfer, February 2005.
010 "Working with the XQC" by Wolfgang Härdle and Heiko Lehmann, February 2005.
011 "FFT Based Option Pricing" by Szymon Borak, Kai Detlefsen and Wolfgang Härdle, February 2005.
012 "Common Functional Implied Volatility Analysis" by Michal Benko and Wolfgang Härdle, February 2005.
013 "Nonparametric Productivity Analysis" by Wolfgang Härdle and Seok-Oh Jeong, March 2005.
014 "Are Eastern European Countries Catching Up? Time Series Evidence for Czech Republic, Hungary, and Poland" by Ralf Brüggemann and Carsten Trenkler, March 2005.
015 "Robust Estimation of Dimension Reduction Space" by Pavel Čížek and Wolfgang Härdle, March 2005.
016 "Common Functional Component Modelling" by Alois Kneip and Michal Benko, March 2005.
017 "A Two State Model for Noise-induced Resonance in Bistable Systems with Delay" by Markus Fischer and Peter Imkeller, March 2005.
018 "Yxilon – a Modular Open-source Statistical Programming Language" by Sigbert Klinke, Uwe Ziegenhagen and Yuval Guri, March 2005.
019 "Arbitrage-free Smoothing of the Implied Volatility Surface" by Matthias R. Fengler, March 2005.
020 "A Dynamic Semiparametric Factor Model for Implied Volatility String Dynamics" by Matthias R. Fengler, Wolfgang Härdle and Enno Mammen, March 2005.
021 "Dynamics of State Price Densities" by Wolfgang Härdle and Zdeněk Hlávka, March 2005.
022 "DSFM fitting of Implied Volatility Surfaces" by Szymon Borak, Matthias R. Fengler and Wolfgang Härdle, March 2005.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

023 "Towards a Monthly Business Cycle Chronology for the Euro Area" by Emanuel Mönch and Harald Uhlig, April 2005.
024 "Modeling the FIBOR/EURIBOR Swap Term Structure: An Empirical Approach" by Oliver Blaskowitz, Helmut Herwartz and Gonzalo de Cadenas Santiago, April 2005.
025 "Duality Theory for Optimal Investments under Model Uncertainty" by Alexander Schied and Ching-Tang Wu, April 2005.
026 "Projection Pursuit For Exploratory Supervised Classification" by Eun-Kyung Lee, Dianne Cook, Sigbert Klinke and Thomas Lumley, May 2005.
027 "Money Demand and Macroeconomic Stability Revisited" by Andreas Schabert and Christian Stoltenberg, May 2005.
028 "A Market Basket Analysis Conducted with a Multivariate Logit Model" by Yasemin Boztuğ and Lutz Hildebrandt, May 2005.
029 "Utility Duality under Additional Information: Conditional Measures versus Filtration Enlargements" by Stefan Ankirchner, May 2005.
030 "The Shannon Information of Filtrations and the Additional Logarithmic Utility of Insiders" by Stefan Ankirchner, Steffen Dereich and Peter Imkeller, May 2005.
031 "Does Temporary Agency Work Provide a Stepping Stone to Regular Employment?" by Michael Kvasnicka, May 2005.
032 "Working Time as an Investment? – The Effects of Unpaid Overtime on Wages, Promotions and Layoffs" by Silke Anger, June 2005.
033 "Notes on an Endogenous Growth Model with two Capital Stocks II: The Stochastic Case" by Dirk Bethmann, June 2005.
034 "Skill Mismatch in Equilibrium Unemployment" by Ronald Bachmann, June 2005.
035 "Uncovered Interest Rate Parity and the Expectations Hypothesis of the Term Structure: Empirical Results for the U.S. and Europe" by Ralf Brüggemann and Helmut Lütkepohl, April 2005.
036 "Getting Used to Risks: Reference Dependence and Risk Inclusion" by Astrid Matthey, May 2005.
037 "New Evidence on the Puzzles. Results from Agnostic Identification on Monetary Policy and Exchange Rates." by Almuth Scholl and Harald Uhlig, July 2005.
038 "Discretisation of Stochastic Control Problems for Continuous Time Dynamics with Delay" by Markus Fischer and Markus Reiss, August 2005.
039 "What are the Effects of Fiscal Policy Shocks?" by Andrew Mountford and Harald Uhlig, July 2005.
040 "Optimal Sticky Prices under Rational Inattention" by Bartosz Maćkowiak and Mirko Wiederholt, July 2005.
041 "Fixed-Prize Tournaments versus First-Price Auctions in Innovation Contests" by Anja Schöttner, August 2005.
042 "Bank finance versus bond finance: what explains the differences between US and Europe?" by Fiorella De Fiore and Harald Uhlig, August 2005.
043 "On Local Times of Ranked Continuous Semimartingales; Application to Portfolio Generating Functions" by Raouf Ghomrasni, June 2005.
044 "A Software Framework for Data Based Analysis" by Markus Krätzig, August 2005.
045 "Labour Market Dynamics in Germany: Hirings, Separations, and Job-to-Job Transitions over the Business Cycle" by Ronald Bachmann, September 2005.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

046 "Paternal Uncertainty and the Economics of Mating, Marriage, and Parental Investment in Children" by Dirk Bethmann and Michael Kvasnicka, September 2005.
047 "Estimation and Testing for Varying Coeffcients in Additive Models with Marginal Integration " by Lijian Yang, Byeong U. Park, Lan Xue and Wolfgang Härdle, September 2005.
048 "Zeitarbeit in Deutschland: Trends und Perspektiven" by Michael C. Burda and Michael Kvasnicka, September 2005.
049 "Courtesy and Idleness: Gender Differences in Team Work and Team Competition" by Radosveta Ivanova-Stenzel and Dorothea Kübler, September 2005.
050 "Do Factor Shares Reflect Technology?" by Benjamin Bental and Domonique Demougin, September 2005.
051 "Optimal investments for risk- and ambiguity-averse preferences: A duality approach" by Alexander Schied, September 2005.
052 "Relational Contracts and Job Design" by Anja Schöttner, September 2005.
053 "Explicit characterization of the super-replication strategy in financial markets with partial transaction costs" by Imen Bentahar and Bruno Bouchard, October 2005.
054 "Aid Effectiveness and Limited Enforceable Conditionality" by Almuth Scholl, August 2005.
055 "Limited Enforceable International Loans, International Risk Sharing and Trade" by Almuth Scholl, August 2005.
056 "Stock Markets and Business Cycle Comovement in Germany before World War I: Evidence from Spectral Analysis" by Albrecht Ritschl and Martin Uebele, November 2005.
057 "An empirical test of theories of price valuation using a semiparametric approach, reference prices, and accounting for heterogeneity" by Yasemin Boztuğ and Lutz Hildebrandt, November 2005.
058 "Integrable e-lements for Statistics Education" by Wolfgang Härdle, Sigbert Klinke and Uwe Ziegenhagen, December 2005.
059 "What does the Bank of Japan do to East Asia?" by Bartosz Maćkowiak, December 2005.
060 "Portfolio Value at Risk Based On Independent Components Analysis" by Ying Chen, Wolfgang Härdle and Vladimir Spokoiny, December 2005.
061 "How much of the Macroeconomic Variation in Eastern Europe is Attributable to External Shocks?" by Bartosz Maćkowiak, December 2005.
062 "The Impact of Industry Classification Schemes on Financial Research" by Christian Weiner, December 2005.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".