

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Pattern Recognition Techniques and Financial Analysis

A thesis presented in fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Finance

at Massey University

Palmerston North, New Zealand

Suresh Kumar Oad Rajput

2014

“He is the source of light in all luminous objects. He is beyond the darkness of matter and is unmanifested. He is knowledge, He is the object of knowledge, and He is the goal of knowledge. He is situated in everyone’s heart”

-Bhagavad-Gita As It Is (13.18)

Copyright is owned by the Author of this thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. This thesis may not be reproduced elsewhere without the permission of the Author.

**I dedicate this thesis to my parents,
for teaching me the value of knowledge and my wife,
for supporting me to acquire that knowledge.**

ABSTRACT

The balance sheet statement is an essential feature of financial reporting, and is expected to convey complete information on firms' operating business decisions. Since these decisions are based on the manager's perception of the existing and future investment opportunities, they cannot be directly observed. This results in two major data analysis issues. First, it is difficult to observe directly the most common operating business decisions; secondly, these decisions may not have a same linear relation to all firms and all firm's performance measures. This thesis attempts to address these issues in three interconnected essays.

The first essay examines an outcome of the double-entry bookkeeping system when financial transactions simultaneously shift a firm's financial position, providing the special information to interpret the meaning of a transaction. Using the factor analysis model, this essay makes use of this information, and identifies the five fundamental factors (decisions) that can capture a firm's time-varying operating business status in a given year. These factors include: financial flexibility, short-term credit, long-term investment, convertible debt usage, and preferred stock usage. The method of extracting these factors controls for missing variable bias, account for limited attention, and provide true decomposition of accounting aggregates such as total asset growth. These factors subsist in predicting future stock returns, forecasting a firm's value (Tobin's Q), cash flows, and earnings beyond their well-known determinants.

The second essay explores the sources of return predictability contained in financial flexibility, which is the first factor identified in essay one. The horse races of the asset pricing versus mispricing tests find a significant positive premium on financial flexibility based return factor, and make it a candidate for a new priced factor. The evidence suggests that covariances dominate the characteristics, and it is non-redundant to well-established risk factors. This factor meets the new conservative minimum of *t*-statistics value of above 3.0 and is constructed using unobserved information.

The final essay addresses the second issue in the data analysis by employing the nonlinear firm grouping technique – the K-means clustering analysis method. Firms are grouped in their 12 natural groups using the five fundamental factors identified in the first essay, and firm size as the clustering criteria. This essay shows how firms differ on priority and the composition of their common operating decisions. This type of firm grouping suggests that operating business decisions are related to firm-specific health and structure instead of industry. This essay recommends the nonlinear firm grouping prior to employing the linear regression models in predicting future performance measures to improve the precision of business analysis.

ACKNOWLEDGEMENTS

This research was supported by the Overseas Scholarship Program Phase II through the Higher Education Commission of Pakistan. The financial support of this sponsor is gratefully acknowledged.

There are several additional people and organisations that I need to thank for their support and help throughout the PhD journey. All of their contributions made this research not only possible, but also enjoyable – and I am genuinely grateful to all of them.

First of all, I must thank my supervisors. Dr. Jianguo Chen has been my main supervisor throughout my PhD research at Massey University. Both students and faculty know him as JG in the School of Economics and Finance. JG is a down-to-earth personality and easy to reach and that is why students who either want to do a masters by research or PhD prefer him to be their supervisor. He is also well known for his high-end expertise in financial econometrics, which is why most of the faculty wants him to collaborate in their research. I truly appreciate his willingness to let me work on my research and introducing me to the importance of pattern recognition techniques in financial analysis.

My second supervisor, Dr. Udomsak (Jeff) Wongchoti, was the first person who spoke the most precious sentence of: *“Yes, I am ready to supervise you, but I want JG to be the main supervisor”*. I said *“Why not,”* as I knew both of them well as they had taught me the research methods in finance (125.785) paper during my Postgraduate Diploma in Finance. I must say, Jeff has always been supportive of the research and encouraged me to

be creative and to find new ways to refine research outputs so that it makes the right impact on readers. He also has a keen critical eye to ensure that the research problem and its results are theoretically sound and empirically correct. Finally, their long discussions over my research almost every other week, provided guidance, encouragement and constructive support throughout my PhD research.

This research was conducted at the School of Economics and Finance, Massey University, Palmerston North Campus. This situation was unrivalled and very favourable in terms of both the research setting and my personal situation. The office space, database, statistical software, weekly research seminar series, and exclusive IT support staff were vital for the research to grow and be refined. Special thanks to Fong Mee Chi for helping me with database availability and encouragement throughout my research. I also admire the support of Cameron Rhodes for IT support, and Maryke Bublitz and Kim Williams for their administrative supports. Palmerston North is famous as a student city and it does meet the expectations of such a label. I would say this is the only city in New Zealand where students have the luxury to live with a family on a meager scholarship stipend. I appreciate the New Zealand system of providing free medical services and free education for my children and free maternity services to my wife. I also appreciate the social welfare activities of the city council and the different organisations that help overseas people to settle in the city.

As I mentioned above, I have had fun throughout the PhD journey. This is due entirely to the wonderful people I shared the office with and my incredible Pakistani friends – both in New Zealand and Pakistan. All of the fellow PhDs have been amazingly fun

people to work with. There are a few people in particular that I would like to thank. Lee Hwei (Karren) Khaw, Wei Yi, Ahmed Raza, Pok Wei Fong, Azira, Norlida Mahussin, Xiaoqi Chen, Uzma Shehzad, Mui Kuen Yuen, and Q. (Sophie) Wang have been the best office friends I could have asked for. Karren and I used to have long discussions and fights over the research problems. Saqib Shareef, my senior PhD fellow and Pakistani scholar, has been an inspiration and great help right from the beginning of my studies here at Massey University. Mr. Zafar Hayat has been the best friend right from the start of our journey from Pakistan to New Zealand. And, we have been the harshest critics of each other. I would also extend my gratitude to Faisal Rana, economics PhD fellow and Pakistani scholar, who has always been ready to offer assistance at the crucial time in a most positive way. He is known in the community for his modesty and patience. I have also been lucky to find support from Dr. Zulifqar Butt, Dr. Muhammad Imran, Muhammad Ibrar Bari, Asif Zia, Sajjid Khan, Dilawar Arbab, Jai Forrester, Vikas Mital, Hamid Irshad, Islahudin Qazi, Aamir Ghafoor Bajwa, Zahid ur Rehman, Mahmood Ghaznavi, and Shujjat Khan.

I would like to thank Ben Marshall, Ralph A. Walkling, Terry Walter, Steven Cahan, Baljit Sidhu, Ronald J. Balvers, Dayoung Huang, H. Zafer Yuksel, Januj Juneja, Aslihan Cicek, Vitali Alexeev, Robert Faff, Leigh Roberts, Hai Lin, David Tripe, Kevin Davis, Petko Kalev, Paul Docherty, Md Humayun Kahir, Jing Liao; seminar, workshop, and conference participants at Massey University; the Second and Fourth Young Researcher Workshops, Sirca, 2013-14 in Sydney; the 2013 annual meeting of the Financial Management Association in Chicago; and the School of Economics and Finance Symposium, Victoria University in Wellington.

Finally, I express my gratitude to my wife Saroj for her unflinching support and backing me up all along this PhD journey. I admire her resolve and patience in taking care of me and raising our children. The success of my research would not have been possible without her support. She is the love of my life. I have always cherished the prayers of my loving father and mother during my studies. I fully appreciate their patience and fortitude while I have been away. I dedicate this thesis to my father Bhooro Mal Oad Rajput, mother Ajni, wife Saroj, and to my two beautiful children Bhoomika and Adarsh who consistently provided encouragement to accomplish this PhD project.

TABLE OF CONTENTS

ABSTRACT	v
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	xi
LIST OF TABLES	xvi
LIST OF FIGURES	xviii
CHAPTER ONE: INTRODUCTION.....	1
1.1 Introduction and Motivation	1
1.2 Essay One: Latent Accounting Growth, Corporate Financing Decisions, and Return Predictability	8
1.3 Essay Two: Financial Flexibility: A New Risk Factor?	10
1.4 Essay Three: Latent Accounting Growth, Firm Size, and Firm Grouping Dynamics	12
1.5 Research Outputs from the Thesis	15
1.6 Structure of the Thesis	17
CHAPTER TWO: ESSAY ONE: LATENT ACCOUNTING GROWTH, CORPORATE FINANCING DECISIONS, AND RETURN PREDICTABILITY	18
2.1 Introduction.....	20
2.2 Data Sample, Model, and Methodology	28
2.2.1 Data sample.....	28
2.2.2 Model for identifying common corporate financing decisions.....	29
2.2.2.1 Model setting	29
2.2.2.2 Factor analysis model	31

2.2.3	Methods	36
2.2.3.1	Fama-MacBeth (FM)	36
2.2.3.2	Fama-MacBeth with Newey-West (FM-NW)	37
2.2.3.3	OLS cluster-robust standard errors	38
2.2.3.4	WLS (weighted-least-squares) two-way cluster-robust standard error ..	39
2.3	Model Outputs and Descriptive Statistics.....	39
2.3.1	Factor analysis model outputs.....	39
2.3.2	New factors time-series stability.....	48
2.3.3	Descriptive statistics	48
2.4	Empirical Analyses	53
2.4.1	Does latent accounting growth matter to the fundamentals of business?53	
2.4.1.1	Q theory implications.....	55
2.4.1.2	Predicting firm value	58
2.4.1.3	Prediction of future cash flows and earnings.....	63
2.4.2	Forecasting future stock returns.....	71
2.4.3	Portfolio analysis	78
2.4.3.1	Time trends of new factor ranked extreme decile.....	78
2.4.3.2	Characteristics and risk adjusted long-run abnormal returns.....	86
2.4.3.3	Annual buy-and-hold returns	92
2.5	Conclusion	95
CHAPTER THREE: ESSAY TWO: FINANCIAL FLEXIBILITY: A NEW RISK		
	FACTOR?	97
3.1	Introduction.....	99
3.2	Data and Methodologies	107
3.2.1	Sample and data construction	107

3.2.2	Measurement and characteristics of FIN_FLEX	108
3.2.3	Financial flexibility return factor (FLEX)	114
3.2.4	Asset pricing methodologies.....	115
3.2.4.1	Two stage cross-sectional regressions (2SCSR).....	115
3.2.4.2	Characteristics versus covariance test.....	118
3.2.4.3	Fama and MacBeth’s (1973) cross-sectional regressions.....	118
3.2.4.4	Financial flexibility and information uncertainty	119
3.3	Descriptive statistics, correlation matrix, and relationship of financial flexibility with prominent factors	120
3.3.1	Descriptive statistics	120
3.3.2	FLEX: A non-redundant factor.....	125
3.4	Empirical Results.....	127
3.4.1	Is financial flexibility priced in the cross-sectional regressions?	127
3.4.2	Do covariances dominate characteristics?	130
3.4.2.1	Portfolio evidence based on Daniel and Titman’s (1997) test.....	131
3.4.2.2	Cross-sectional regressions evidence.....	135
3.4.2.3	Information uncertainty and financial flexibility.....	135
3.5	Conclusions.....	139
CHAPTER FOUR: ESSAY THREE: LATENT ACCOUNTING GROWTH, FIRM SIZE, AND FIRM GROUPING DYNAMICS		142
4.1	Introduction.....	144
4.2	Data Sample, Variable Estimation and Research Methods	153
4.2.1	Data sample.....	153
4.2.2	Estimation of clustering criteria.....	154
4.2.3	Descriptive statistics	155
4.2.4	Research methods	157

4.2.4.1	K-means cluster analysis	158
4.3	Results.....	161
4.3.1	Cluster number choice	161
4.3.2	Cluster summary	163
4.3.3	Cluster descriptive statistics.....	163
4.3.4	Clustering versus industry patterns.....	171
4.3.5	Predictive regression analyses across clusters	180
4.3.5.1	Predicting the future firm value	181
4.3.5.2	Predicting future cash flows	184
4.3.5.3	Predicting future earnings.....	185
4.4	Conclusion	186
CHAPTER FIVE: CONCLUSION		190
5.1	Major Findings and Implications.....	190
5.1.1	Essay One: Latent accounting growth, corporate financing decisions, and return predictability.....	190
5.1.2	Essay Two: Financial flexibility: A new risk factor?	192
5.1.3	Essay Three: Latent accounting growth, firm size, and firm grouping dynamics	194
5.2	Future Areas of Research.....	196
REFERENCES		198
APPENDIX A : ESSAY ONE....		213
Appendix A.1: Definitions, Data source, Mnemonics and Labels for the Original Input Dataset Items		213
Appendix A.2: Definition of variables.....		216
Appendix A.3: Robustness Tests		219

Appendix A.4: Full Sample Factor Ananalysis Output Factors based Return Predictability
Tests.....228

APPENDIX C: ESSAY THREE234

Appendix C.1: Fama and French Industry Specifications based on SIC Codes234

LIST OF TABLES

Table 2.1	Latent accounting growth factors identified by factor analysis.....	43
Table 2.2	Mean (median) values of selected characteristics for decile portfolios sorted by latent factors.....	50
Table 2.3	Pearson (Spearman) correlation coefficients between latent accounting factors and other characteristics	51
Table 2.4	Forecasting firm value.....	61
Table 2.5	Forecasting future firm cash flows	67
Table 2.6	Forecasting future firm earnings	69
Table 2.7	Relations between latent growth factors and the subsequent monthly excess stock returns	75
Table 2.8	Average monthly abnormal returns for latent factors decile portfolios 1, 2 and 3 years after portfolio formation (whole sample period).....	90
Table 3.1	Characteristics of financial flexibility portfolios.....	112
Table 3.2	Fama-MacBeth cross-sectional regressions	113
Table 3.3	Summary statistics and correlations for factors, factor loadings, and underlying variables	122
Table 3.4	Time-series regressions	126
Table 3.5	Two-pass cross-sectional regression on 25 size-book-to-market portfolios.....	128
Table 3.6	Characteristics versus covariances tests	132
Table 3.7	Fama-MacBeth cross-sectional regressions	136
Table 3.8	Information uncertainty and financial flexibility	138
Table 4.1	Descriptive statistics.....	156
Table 4.2	Cluster summary and variables statistics.....	164
Table 4.3	Descriptive cluster analysis	169
Table 4.4	Industry characteristics' pattern across clusters	177
Table 4.5	Predicting future fundamentals across the clusters	183
Table A.1	Definitions, Data source, Mnemonics and Labels for the Original Input Dataset Items	213

Table A.3.1	Relations between latent factors and subsequent monthly raw stock returns	220
Table A.3.2	Relations between return anomalies and latent growth factors	222
Table A.4.1	Fama-MacBeth Regressions of Annual Stock Returns on Latent Growth Factors and Other Variables	232
Table C.1.1	234
Table C.1.2	234
Table C.1.3	234

LIST OF FIGURES

Figure 2.1 (a-e): Q theory implications.....	57
Figure 2.3 (a-f): Time trends of ST_CREDIT ranked extreme decile	82
Figure 2.4 (a-f): Time trends of LT_INV ranked extreme decile	83
Figure 2.5 (a-f): Time trends of CVT_DEBT ranked extreme decile.....	84
Figure 2.6 (a-f): Time trends of PSK_USE ranked extreme decile	85
Figure 2.7 (a-j): Time-series average annual buy-and-hold returns spread	93
Figure 2.8 (a & b): Annual buy-and-hold returns	94
Figure 4.1: Choice of number of clusters and pseudo F-statistics and cubic clustering criteria.....	162

CHAPTER ONE

INTRODUCTION

An intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius, and a lot of courage to move in the opposite direction.

-Albert Einstein

This chapter presents the overall motivation for, and an overview of, this thesis. In particular, it outlines the motivation and the important contribution of the three essays contained in this thesis. The chapter concludes by drawing a structure for the remainder of the thesis.

1.1 Introduction and Motivation

The Statement of Financial Accounting Concept No.1 identifies that the balance sheet statement is a central part of financial reporting. The balance sheet is expected to provide a spatial and temporal snapshot of the firm's assets and liabilities to its potential users such as investors, creditors, and stakeholders. Consequently, it is assumed that a balance sheet provides all the necessary information with regards to a firm's operating business performance. However, it is observed that there are two major issues in the data analysis. First, it fails to provide information about how many fundamental (common) factors are crucial to map a firm's time-varying business status. Second, it cannot be automatically assumed that these factors have a simple and linear relation with the performance measures.

In the real world, all firms take certain common corporate finance decisions such as capital structure, liquidity management, and pay-out policy decisions (Denis, 2011). However, a firm's prioritisation and composition of such decisions is time-varying. For example, firms simultaneously use heterogeneous types, sources and priorities of debt (Rauh and Sufi, 2010). This debt heterogeneity is firm-specific, and is driven by the firm's financial health and investment opportunities. Correspondingly, financial flexibility is the first priority to real-world chief financial officers (CFOs) (Graham and Harvey, 2001); and academics find it as a missing link in capital structure theories (DeAngelo and DeAngelo, 2007). However, it is unclear how one estimates accurately this unobserved measure (Marchica and Mura, 2010); how do firms attain and give up flexibility? And to what extent is financial flexibility crucial in predicting the firm's performance? Thus, the pattern of financial flexibility differs between firms and across time. This heterogeneity of the balance sheet information also affects the perceptions of its potential users such as investors.

Investors' opinions and expectations are heterogeneous about pricing the information content they have, and asset returns they may obtain. These varied beliefs are not captured by close and pre-conceived asset pricing theories and models (Fernandez, 2014)¹. The traditional theories and testing models are linear; firms' and investor attitudes are not linear, and thereby financial asset valuation information is not simply linear. The

¹ Bateson (1972) in his book *Steps to an Ecology of Mind* (pp. 488 to 503) argues that, most of the ideas with a history of success have been conceived when the world has been relatively less complex. Back then, those ideas consistently produced positive outcomes because they were matching the dynamics of the world at that time. However, the modern world is excessively non-linear and extremely complex. The complexity of the world implies that there are too many variables involved in any given situation. Therefore, the application of previously successful, abstract, and generalised ideas without their proper feasibility evaluation results in ignoring many important links that exist with the system. This means narrowing down the focus to only certain linear relationships. Consequently, the application of outworn ideas fails to provide wide-ranging solutions that can resolve all the targeted problems. In order to overcome the issue of practice failure, a paradigm shift in ideas and practices from inflexibility to flexibility is essentially required.

financial information data is varied and vast and needs to be captured in its exact pattern instead of treating pieces of information in isolation or lumping them together in pre-determined assumptions. For example, there has been an increased use of balance sheet information to determine stock prices^{2,3}. The return predictability based on such a suboptimal construct gives rise to limited attention (Hirshleifer, et al., 2004). In fact, a single or pre-determined set of accounting items cannot be representative of certain decision types. In other words, an input of insufficient information would result in substandard output. The reason is that when firms make a decision, the result of the implementation could be recorded in many accounts of the balance sheet – double-entry bookkeeping system, thus potential users may miss out on interrelated information.

The double-entry bookkeeping system makes the investigation of information interaction difficult because, the financial transaction-related information may get concealed in various balance sheet accounts. In contrast, one account could be used to keep track of many different business activities. For example, a cash account balance is influenced by net business profits, which may also be caused by net investment activities. It is observed that double-entry bookkeeping may produce some supplementary information by separating the cash balance into different components. If the component of a cash account change is constantly, and jointly moving with operating business incomes (owner's equity), it might be an indication of business profits. And, if another part of the cash

² See, for example, Ou and Penman (1989), Holthausen and Larcker (1992), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), and Piotroski (2000) for evidence that various balance sheet ratios can be utilised to explain future return.

³ The prominent accounting information based return anomalies include accounting accruals (Sloan, 1996), net operating assets scaled by lagged total assets (cumulative accruals) (Hirshleifer et al., 2004), investment-to-assets (Lyandres et al., 2008), external financing (Bradshaw et al., 2006), net stock issuance (Daniel and Titman, 2006; Pontiff and Woodgate, 2008). Additionally, Cooper et al. (2008) document that the total asset growth anomaly is superior in predicting the average returns compared to other return anomalies.

balance change has regularly moved in conjunction with the change in working capital or fixed assets, it may be a superior signal of investment expenditure. These types of moving components together may serve as an excellent indication of types of business decisions. If that is the case, then the individual decision type can be represented by multiple accounting items, and all items are expected to follow similar movements and represent the same decision across time. Thus, this thesis on the one hand, contributes a way forward to overcome the problem of limited attention by recognising the exact nature of the unobserved common corporate finance decisions. On the other hand, these decisions have implications for deeper understanding of market and financial information such as return predictability, firm profitability, and firm value.

The resolution of these issues leads to the more complex question of how to map accurately the interactions of a large nonlinear information set; a jigsaw puzzle with no reference pattern. The key to solving the puzzle is to focus on the basic purposes of the multivariate analysis. In particular, the purpose of the structural simplification helps in reducing the array of information into a few understandable variables (factors); and then classifying the entities or objects based on the simplified factors (Kendall, 1975). These aims of data reduction and classification require one to employ the correlation structure of the information set⁴.

This thesis in three interconnected essays focuses on resolving the identified data analysis issues, as they have not been adequately addressed in the literature. Two main

⁴ Subrahmanyam (2010) in his survey of literature on the cross-section of expected returns argues that researchers need to employ the correlation structures amongst the candidate variables in predicting the stock returns and varying methodologies.

pattern recognition techniques are used – factor analysis and cluster analysis⁵. The first essay identifies the most important business activity factors using factor analysis; and the second essay provides further insights into the importance of one of these factors. The third essay identifies the different business management styles so that each business type is dealt with separately.

Overall, this study intends to contribute to the area of financial pattern recognition by means of exploring the common measures of operating business decisions that not only helps in improving the precision of business analysis, but also in terms of future performance prediction. This objective is accomplished using balance sheet accounting change rate information⁶. The change rate of information is crucial to this research, because the degree of investment and financing activities is not static but is dynamic and changes within time and space; every new transaction a firm makes brings change to the information set. And if that is so, then the pattern of simultaneous change through time may result in time-series regularities and accurate predictions thereafter in financial analysis. In particular, it can serve as a crucial to identifying commonalities across variables; the similarities and differences among and between firms.

⁵ Pattern means something that happens in a regular and repeated way. Source: <http://www.merriam-webster.com/dictionary/pattern>. In everyday language, the term pattern is associated with dresses, weather, events, and spending patterns. In this thesis, the term pattern is used to identify the firms' corporate finance decision trends. Then, the field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying data into different categories (Bishop, 2006).

⁶ Roberts (1959) argues that *level* can give an artificial appearance of price's *pattern* when compared to *changes*. Brown and Goetzman (1997) argue that the change rate of time-series of historical accounting information would free the data from units and scale questions.

The first essay examines the correlation structure of the balance sheet change rate information of US common firms from 1984 to 2009 in order to find the common factors that map the time-varying operating business. The essay hypothesises that the balance sheet information can be aggregated into certain stable and orthogonal latent factors. The research is initiated from the idea that business operating decisions and their results should be fundamentally reflected in financial statements, before or after the time when the decisions are made (Myers, 1974). The history of financial statements contains rich information about the status of a business, but it is not clear how many common operating business decisions there are. The real challenge to filling this gap is that the exact nature of accounting information interactions are unknown. As a result, this essay starts with a systematic quantitative search to identify commonalities in simultaneous shifts among comprehensive sets of balance sheet change rates across firms. The first essay documents the change in financial flexibility (FIN_FLEX) as a first factor outcome when factor analysis is employed. This factor explains the maximum variation of the complete balance sheet and shows greater stock return forecasting power.

The second essay in this thesis examines the source of the greater return predictability of FIN_FLEX. The literature on asset pricing is generally divided into two camps (Lin and Zhang, 2013). One camp believes that the firm-specific, information-driven return predictability is due to the characteristics itself. It is because an investor either underreacts or overreacts to such information (mispricing) (Daniel and Titman, 1997; Daniel et al., 2001). This happens due to problems of limited attention (Hirshleifer, et al., 2004). The other camp led by Fama and French (1993, 1996) believes that it is the covariances (loadings) on the factors that derive the returns and not the characteristics. The

existence of positive premium trading on the characteristics could be an outcome of chance. However, this essay finds that FIN_FLEX based on a zero-cost return mimicking portfolio (FLEX) is a new priced risk factor. This finding is consistent with the new yardsticks of t -statistics value above 3.0 for average positive premium. The positive premium on FLEX is robust to well-known pricing factors, insensitive to the information environment, and a compensation for the manager's perception (risk) about future growth opportunities. Thus, this essay shows that covariances dominate the characteristics.

Although the first essay identifies the most common factors representing the operating business, the subsequent operating performance predictions involve linear models. The operating business is time-varying and in turn the corporate finance decisions are not linear. Thus, the final essay of this thesis contributes a different style of grouping firms that helps in improving the precision of business analysis. In contrast to the traditional firm grouping⁷, this essay uses five latent factors (identified in the first essay) and firm size to cluster firms into 12 natural groups. The firms are homogeneous within the cluster and heterogeneous across the clusters. This essay finds the common corporate finance decisions as firm-specific, and heterogeneous to the health and structure of their industry or their stage of business development. The idea is that the corporate entities are not unique, but share common attributes. These attributes can be used to make natural groupings of firms that share similar business management styles. And, if that is so, then investors can construct better profitable investment strategies for a specific style. Next, this

⁷ The grouping of firms is usually done using firm attributes such as size, book-to-market ratio (Fama and French, 1993) or on industry affiliation. The prominent industry classification systems include Standard Industrial Classification (SIC), North American Industry Classification System (NAICS), Global Industry Classification Standards (GICS) and Fama and French's (1997) reclassification of firms into 48 industries based on the SIC codes.

study explores what attributes or clustering criteria make these new clusters unique. Lastly, the essay seeks to understand the performance of clustering criteria in predictive regressions of future firm fundamentals within each cluster in order to understand the different strategies and performance indicators among these clusters.

The remainder of this chapter provides a brief overview for each of the three essays in order, Section 1.6 presents the research outcomes, and Section 1.7 summarises the structure of the remainder of the thesis.

1.2 Essay One: Latent Accounting Growth, Corporate Financing Decisions, and Return Predictability

The balance sheet statement is a key feature of financial reporting and is outlined in the Statement of Financial Accounting Concept No.1 through the Financial Accounting Standards Board (FASB). Firms record both real and financial transactions (decisions) in the balance sheets using the double-entry bookkeeping system. This accounting system on the one hand, generates the problem of data duplication by concealing the information relating to a transaction in various accounts; and, on the other hand, recording the value of the same asset varies among firms. The distinctive accounting practices pose a challenge of analysing the interactive relations of this accounting information. The understanding of these relations is the key to knowing the proper number and meaning of operating business decisions as these decisions are not directly observed as they depend on the firms' perceptions of their future investment opportunities and current financial position. This aspect of data analysis is not adequately addressed in accounting and finance literature. This essay explores the correlation structure of a long list of balance sheet change rates to

obscure the most common factors that are fundamental to mapping the time-varying space of business operating decisions.

This essay employs factor analysis on 25 annual cross-sections separately to identify the data interactions and reduce it to the basic common corporate finance decisions that orthogonally synchronise through time. It is a statistical pattern recognition technique that takes into account common variance using the correlation structure of the input data variables. Although factor analysis is a linear technique, taking change rates information as an input helps to account for nonlinearity in the data to some extent. The factor analysis reduces the 37 dimensions of the change rate variables into five common and orthogonal latent factors every year^{8,9}. The common factors include: Financial Flexibility (FIN_FLEX), Short-term Credit (ST_CREDIT), Long-term Capital Investment (LT_INV), Convertible Debt Usage (CVT_DEBT) and Preferred Stock Usage (PSK_USE). The finding of these factors is a significant contribution to accounting and finance literature because these factors are not directly observed from the balance sheet information; even the exact number of fundamental factors is unknown. Moreover, this study is able to identify which and how many of the accounting change rates are necessary to measure a particular factor. Additionally, this essay makes another contribution by showing the importance of these factors in forecasting stock returns and firm fundamentals.

⁸ This study also applies the factor analysis on the full sample as a single cross-section. The full sample factor analysis factor structure is not provided in this thesis for the sake of brevity as it is similar to annual factor analysis shown in the essay one. However, the full sample based factors' implication for return predictability is shown in Appendix A. 4.

⁹ The five common factors are interchangeably named as a basic common corporate finance decision, operating business decisions, new factors, and unobserved or latent growth factors in this thesis.

First of all this essay finds that the new factors possess the strong explanatory power of cross-sectional stock returns compared to well-established return determinants in the finance literature. The latent factors based trading strategies controlling for pricing characteristics (e.g., size, book-to-market ratio, and price momentum) are shown to earn substantial incremental returns. Secondly, this essay examines the impact of the new measures of operating business decisions on firm value. It is surprising that these decisions (factors) are value adding activities that persist in creating value to their firms for at least up to three years. Moreover, this finding is consistent with the Q theory of investment (Cochrane, 2011) that an increase in investment and financing activities results in higher Tobin's Q values and *vice versa*. In particular, this essay argues that the long-run robustness of the new factors in predicting firm value is due to their intertemporal and forward-looking nature. In addition, this result also means that firms are able to correctly identify positive investment opportunities. Finally, this essay seeks to understand the usefulness of the new latent factors in forecasting performance measures such as firms' cash flows and earnings in the standard error clustering regressions. The evidence suggests that these factors robustly determine two performance measures.

1.3 Essay Two: Financial Flexibility: A New Risk Factor?

The second essay in this thesis examines whether information relating to financial flexibility is priced or not. Financial flexibility is perceived to be the first important consideration in capital structure decision-making by CFOs. Since investors judge every action of CFOs as a signal of new information relating to a firm's policy, then financial flexibility based decisions may be readily detected and incorporated into asset prices. As a

result, any premium earned trading on financial flexibility information would be the compensation for the risk associated with the manager's perception about future growth opportunities.

This essay focuses on the financial flexibility (FIN_FLEX) measure for several reasons. First, following the essay one, it is the first factor obtained using the factor analysis, explaining about 40% of the variance of the sample. Second, FIN_FLEX explains the cross-sectional variations in subsequent excess stock returns and the ability to earn 17% (13%) annual buy-and-hold returns on both equal (value) weighted basis. This return predictability is robust to the controls of prominent factors such as firm capitalisation, book-to-market ratio, price momentum, gross profitability, accounting accruals, net operating assets and total asset growth. Third, the first essay finds FIN_FLEX is an integrated measure of both investment and financial flexibility (Denis, 2011; Gamba and Triantis, 2008), and is forward looking (Ang and Smedema, 2011). Fourth, FIN_FLEX is found to be consistent with the definition extended by Marchica and Mura (2010) that flexibility is an unobservable factor which depends mainly on a manager's perceptions of future growth options. Fifth, the literature is silent on the underlying source of FIN_FLEX's return predictability.

The main contribution to the existing asset pricing literature is that this essay identifies a new risk factor (FLEX). FLEX is the return mimicking portfolio: the difference between stock return of a most negative (positive) shift in financial flexibility. The study employs multiple tests of both asset pricing and mispricing models that are well known in the literature. Using the two-stage cross-sectional regression framework (Cochrane, 2001),

this research finds FLEX as a priced factor, it bears a positive and significant premium. The findings are further verified using Daniel and Titman's (1997) characteristics and covariance tests and Fama-MacBeth's (1973) cross-sectional regression framework.

The second contribution of the research is that the value factor (HML) is a redundant factor to FLEX. The implications of this finding could be huge for the literature. The foremost implication requires detailed study to test whether FLEX is the underlying source of risk inherent in the HML. The third contribution this essay makes is that the FLEX meets the new pre-conditions to be a candidate of a risk factor; its average *t*-statistics value is 5.56, which is well above the new benchmark of 3.0 (Harvey et al., 2013). Lastly, this study also highlights that the FIN_FLEX return predictability is unrelated to the information environment. In addition to the above contributions, this research also sheds light on GPA (Gross Profitability) – a new measure of gross profitability defined by Novy-Marx (2013). The evidence suggests that GPA does not erode the forecasting power of FIN_FLEX, instead it gets subsumed in the Fama and MacBeth (1973) cross-sectional regressions.

1.4 Essay Three: Latent Accounting Growth, Firm Size, and Firm Grouping Dynamics

The first essay of this thesis is able to identify the five common corporate financing decisions using factor analysis – a linear factor forming technique. These decisions (factors) have been used in forecasting firms' performance measures using a simple linear framework. It is observed that such an assumption of linear relation does not fit the real-world situation, where firms differ in terms of priority and composition of these decisions.

For example, firms vary in their priority, composition, and type of leverage usage. This indicates that the behaviour of firms and the capital markets is heterogeneous, and cannot be captured by traditional linear models. Consequently, it cannot be automatically assumed that these five decisions have a simple and linear relation with firms' performance measures. Because these factors measure a certain type of business activity (or investment), one cannot make a simple statement or conclusion that such an increase or decrease will impact the firm's value in a certain way. Moreover, financial theory tells us that the same investment activity could make for very different results depending on the current economic situation of a company.

The third essay addresses this issue of the data analysis by showing how heterogeneous these common corporate financing decisions are in terms of priority and composition. The identification of such heterogeneity makes a major contribution to the literature on corporate finance, industry classification, portfolio diversification, asset pricing, and the new area of investment-specific shocks. This study identifies 12 different business management styles, by means of the five common corporate finance decisions (produced in our factor analysis result) and firm size.

This essay extends the data analysis to a new level in addressing the question of how firms differ in these common corporate financing decisions and why. The idea behind the research is based on a two conjectures. According to the first conjecture, the real corporate world is much more complicated than the way it is normally modelled in linear models or optimal formulae. The second conjecture recognises that existing firm managers have the sense to make an optimising marginal change while observing their financial

status and future investment opportunities. For example, to some firms, it is better to have a high level of leverage and for other firms, it is better to be in the lower leverage end. Thus, this research confirms the existence of multiple optimal structures corresponding to various firm-specific common situations. These situations could be expressed in the dimensions of the business life cycle, business risk nature, technical or technological components, and so forth. With several common business statuses, there could be multiple optimal business structures. The well-experienced manager will choose investment and financing decisions around the optimal structures. There exist many firm structure groups; firms within each group are close to homogeneous and generally heterogeneous across the groups. As a result, for varied business analysis, these groups should be treated separately. In such a context, the technique of clustering analysis is ideal, as it is a purely numerical technique that identifies the most populated clusters of firms.

Not all firms are unique or identical, and they may be grouped together into cohorts that share priorities and the composition of basic attributes. In this study, these attributes are the common corporate finance decisions from the first essay, coupled with firm size. The finance literature typically employs several firm classification methods such as SIC, NAICS, GICS and FF reclassification of firms in 48 industries based on the SIC codes. All of these traditional industry classifications are based on the product-market outputs, and on the assumption that corporate finance decisions are related to health prospects and the risk nature of the particular industries. In contrast, this study conjectures that firms can be classified into groups based upon individual business uniqueness and managers' investment decisions. There are a limited number of common business and financing styles, and these are reflected in clusters that reveal firm-specific overall quality. This study is the first to use

clustering analysis to classify firms into common investment styles for the purpose of investigating the quality of the within-cluster prediction model. This essay finds that the trial is a success. The R-square value increases from a level of 5% for all firms' regressions to the level of 50% for within-cluster regressions.

Specifically, this essay classifies firms into 12 internally homogeneous and externally heterogeneous groups using non-linear K-means clustering technique. The descriptive analysis of individual assets and industry breakups reveal interesting patterns. In particular, the non-linear relationship between firm size and age; size and book-to-market ratio is identified. The individual cluster appears to contain firms belonging to 10 industry groups. The result shows that five common corporate finance decisions are robust predictors of future firm value, earnings, and cash flow across the 12 clusters. However, this evidence contrasts sharply with the overall regression evidence reported in the first essay. The associated signs and magnitude of the relationship varies from cluster to cluster. This indicates that the classification in this essay provides an improved means of quantitative business analysis, for example, in terms of future performance prediction.

1.5 Research Outputs from the Thesis

Essay One:

The first essay contained in this thesis has been presented at the following forums:

- a) Rajput (2013) "Latent Accounting Growth, Corporate Finance Policies, and Return Predictability" Second Young Researcher Workshop, Sydney, July, 2013.

- b) Rajput (2013) “Latent Accounting Growth, Corporate Finance Policies, and Return Predictability” Seminar Series, School of Economics and Finance, Massey University, July, 2013.
- c) Rajput (2013) “Latent Accounting Growth, Corporate Finance Policies, and Return Predictability” Financial Management Association (FMA) International, Chicago, October, 2013

Essay Two:

The second essay contained in this thesis has been presented at the following forums:

- a) Rajput (2014) “Financial Flexibility: A new risk factor?” 5th SEF Research Symposium, School of Economics and Finance, Victoria University, Wellington, February, 2014.
- b) Rajput (2014) “Financial Flexibility: A new risk factor?” Seminar Series, School of Economics and Finance, Massey University, July, 2014.
- c) Rajput (2014) “Financial Flexibility: A new risk factor?” Fourth Young Researcher Workshop, Sydney, 10-11 July, 2014.
- d) Rajput (2014) “Financial Flexibility: A new risk factor?” Accepted for presentation at the Academy of Financial Services meeting, Nashville, October, 2014.
- e) Rajput (2014) “Financial Flexibility: A new risk factor?” Accepted for presentation at the Auckland Finance Meeting, 18-20 December, 2014.

1.6 Structure of the Thesis

The remainder of this thesis is structured as follows. The first essay uses the correlation structure to reduce the dimensions of the comprehensive accounting growth information and identifies five basic corporate financing decisions. The first essay is presented in Chapter 2. The supplementary information such as accounting item definitions, variable definitions, and further robustness test results are presented in appendix A.

The second essay shows the evidence that the return predictability of FIN_FLEX – the first factor output from the factor analysis – is driven by covariances rather than characteristics and is presented in Chapter 3.

The third essay examines how firms can be categorised into groups of distinct management styles base on five common corporate financing decisions and firm size as grouping criteria and is presented in Chapter 4. Appendix C provides the definitions of industry classification based on the SIC codes.

The chapter 5 concludes the thesis by summarising the major findings and implications of each of the three essays, along with suggestions for future research.

CHAPTER TWO

ESSAY ONE

LATENT ACCOUNTING GROWTH, CORPORATE FINANCING DECISIONS, AND RETURN PREDICTABILITY

This chapter presents the first essay that examines the interactions of simultaneous shifts in balance sheet items annually that can help measure the basic corporate financing decisions. A brief overview of the study is presented in Section 2.1 and the data sample, model setting, variable measurement, and research method used in this study are described in Section 2.2. The discussion on the results of the factor analysis model and descriptive data is provided in Section 2.3. The empirical findings are discussed in Section 2.4, and a conclusion is given in Section 2.5. The chapter's appendices and references are presented in Appendix A.

Latent Accounting Growth, Corporate Finance Decisions, and Return Predictability

Abstract

An annual simultaneous investigation of the interactions of the comprehensive balance sheet growth variables obscures unobserved fundamental factors crucial in mapping the firm-specific the time-varying operating business. Using a factor analysis model, this study identifies that five common shift (latent) factors consistently occur across the study period of 1984-2009. The five factors are interpreted to reflect five major decisions in businesses: financial flexibility, short-term credit, long-term capital investment, convertible debt usage, and preferred stock usage. The analysis has four relevant implications. First, the dimensions of firm-specific accounting information can be reduced by using a correlation structure. Second, the factors are robust in predicting long-run stock returns and generate incremental returns. Third, the new factors are consistent with the Q theory of investment to create firm value up to three years. Fourth, the new factors are strong negative predictors of future cash flows rather than earnings.

JEL classification: C38, G12, G32

Key words: asset pricing models, accounting growth, factor analysis, cross-section stock returns, corporate finance policies

Acknowledgements: Thanks go to the seminar, workshop, and conference participants at Massey University; the Sirca Second Young Researcher Workshop 2013 in Sydney; and the 2013 annual meeting of the Financial Management Association in Chicago.

2.1 Introduction

Corporate finance decisions lead a firm to engage in transactions in real or financial assets that effectively shift its financial position (Myers, 1974). This essay examines whether firms display systematic shifts in their corporate finance decisions observed in the balance sheets across time¹⁰. If so, what are the economic impacts of such decisions to the firm's future stock returns and firm value?

This essay derives motivation from the practice of double-entry bookkeeping in recording operating business activities. This accounting system records the transaction on both sides of the balance sheet by two or more offsetting entries. For example, an increase in assets of \$5,000 would result in an offsetting decrease in the cash account, if the asset is purchased using cash; or an offsetting increase in liabilities (short or long) if the asset is purchased via borrowing; and may increase equity, if the equity holder's stake is used. This example indicates that there could be a combination of internal and external capital sources to raise cash and the uses of this cash. Thus, the act of purchasing an asset brings changes in several balance sheet accounts such as inventories, current assets, fixed assets, total assets, accounts payable, debt in short-term liabilities, and long-term liabilities, among others. Overall, this financial reporting system lets several balance sheet accounts covary depending on the specific business transaction. Then, this study argues that if the structure of the covariances for the long list of accounting items can be identified, the actual fundamental factors mapping the time-varying firm's business can be identified optimally. This essay considers a transaction as an operating investment and financing decision.

¹⁰ Refer to the Statement of Financial Accounting Concept No. 1 that defines the objectives of financial reporting and identifies the balance sheet statement as being an important feature of financial reporting.

This study understands that the observed balance sheet contents do not directly convey complete information about operating business activities. In fact, the exact nature of the accounting item interactions produced by a financial transaction is not directly observed. This creates an important issue in data analysis: it is difficult to know how many fundamental factors represent operating business decisions or firms' financial status. Correspondingly, how do these factors impact on financial and market fundamentals? These questions have a basis in the fact that for the firm's management one of the prime goals is to maximise shareholder value. Managers are expected to show a good outlook for the firm in financial reporting instead of the exact nature of the business. Owing to this fact, financial reports can be misleading. This essay finds it is an important issue to be resolved because of the difficulties in extracting all the relevant parts of a business activity by the potential users (investors) of this information. Accordingly, this study argues that the investor using the partially observed information creates the problem of investor limited attention when pricing the same information (Hirshleifer et al., 2004).

This study argues that the issue of limited attention could be addressed by removing the missing variable biases. The missing variable biases are created due to a predetermined or mechanical way of constructing the accounting and financial measures, which may not be representative of certain types of business decisions because one accounting item could be reflective of several business activities, which can only be captured if one understands their interactive movements. Moreover, current research is using accounting aggregates such as total asset growth to understand the effects of investment in forecasting operating business and stock returns (Cooper et al.,2008). This study argues that aggregates may not be indicative of the different aspects of investment (short-term or long-term) and the ways

in which it is being financed. Thus, the natural decomposition of such an aggregate using the covariances of the components would result in factors that may show the true decomposition of the aggregates and impacts of the different investment activities on firm and market fundamentals. The cross-loadings of aggregates such as total asset growth and long-term debt shown in Table 2.1 indicating their natural decomposition into their natural investment and financing activities based factors.

The above issues in data analysis led to the second motivation for this study. Subrahmanyam (2010) in his survey evidence provides a rationale to study the systematic accounting shifts hidden in numerous items of the balance sheet. He concludes his survey by arguing that more than 50 variables have been used to predict the cross-section of average stock returns, but the research is still inconclusive. He recommends employing the correlational structure amongst variables and a comprehensive set of controls and varying methodologies to test their survival. Consequently, this essay seeks to address the above problems in data analysis. It shows that foremost attention be paid to understanding the correlational structure of the related variables to capture their causal variance. This method eliminates the measurement errors in accounting variables.

To the extent that corporate finance decisions entail simultaneous shifts in various balance sheet items – the real challenge of analysing the accounting information lies in the identification of the interactive and complex relations among them. The study finds that factor analysis is the best tool to identify such an interaction. The factor analysis is applied annually; a mathematical technique that identifies the “*most*” important common movements among the comprehensive list of variables by removing the duplications and

redundancies. These common movements indicate groups of accounting variables that orthogonally synchronise through time. Thus the application of factor analysis is intended to harvest the common variances among the annual change rates and reduce the dimension and not the complete variance¹¹. This study conjectures that these factors could represent certain types of shifts in corporate finance decisions that are common and meaningful among businesses.

Based on a comprehensive set of accounting variables¹² shown in US companies' balance sheets and supplementary balance sheet statements during the period 1984 to 2009, this essay in Table 2.1 identifies five factors (common movements in specific balance sheet items). These factors on average account for approximately 70% of a comprehensive set of accounting ratio changes in a year. The new factors are financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-term capital investment (LT_INV), convertible debt usage (CVT_DEBT) and preferred stock usage (PSK_USE).¹³

Extensive accounting literature, in the last two decades, has introduced several measures of investment and growth capable of predicting future stock returns. This literature separately considers cash or the accrual part of an accounting system and examines the value relevance and reliability side of the accounting numbers.¹⁴ Cooper et

¹¹ The factor analysis model takes into account the common variances and reduces the array of variables to a few factors and has been used in prior finance literature (Abarbanell et al., 2003; Bushee, 1998; Pinches and Mingo, 1973; Sorensen, 2000). See section 2.2 and 2.3 for a detailed discussion of factor analysis model and results.

¹² The balance sheet-level items are converted to growth variables in this research following Chen et al. (2011) and Lyandres et al. (2008).

¹³ For simplicity, interchangeably we name the corporate decision types as 'new factors' that are not the level factors, but change variables. The discussion on new factors is contained in sub-section 2.3.1.

¹⁴ See, for example, Richardson et al. (2005), Bradshaw et al. (2006) and Dechow et al. (2008) among others.

al., (2008) find a total asset growth – an aggregate measure of overall growth that captures the return effects of other investment and financing measures. However, as noted by Lewellen and Resutek (2013), investment, external financing and accruals have dissimilar characters in predicting returns. In the same vein, this study argues that the new latent (unobserved) accounting growth measures are not a mere variation of asset growth measures. Instead, they are obtained using causal variance of the accounting items, which is irrelevant either to their prior mechanical relations or to the type of accounting system¹⁵. Furthermore, these factors do not bear high correlation with any of accruals measures – the highest significant spearman correlation of 0.043 exist between ST_CREDIT and ACCR (measure of accrual) – such as the new accrual measure extended by Richardson et al., (2005). Similarly, the highest spearman correlation of investment-to-asset ratio (I_A) as in Lyandres et al. (2008) and net operating assets scaled by lagged total assets (NOA) as in Hirshleifer et al. (2004) with LT_INV are 0.29 and 0.12, respectively. In addition, the sampling adequacy assumption of the factor analysis allocates the high loads to the highly reliable accounting items rather than the less reliable. Thus much more refined latent accounting growth measures are obtained.

After demonstrating the existence of the common corporate finance decisions, this study focuses on how firm value varies with variations in the five decisions. This focus on firm value helps to relate the Q theory of investments to the new factors (Cochrane, 2011). This study formally tests whether these new measures create or destroy firm value. This entails showing whether firms have high (low) Tobin's Q – a proxy for firm value, which is

¹⁵ Abarbanell and Bushee (1997) advocate the cautious use of mechanical means in combining the accounting items in the fundamental analysis.

constructed as in Hou and Robinson (2006) – values at times when there is an increase (decrease) in latent investment and financing decision measures. The main findings are shown in Figure 2.1 (a-e) and Table 2.4. As shown in Figure 2.1, all latent factors apart from but CVT_DEBT, are consistent with the Q theory of investment. Strikingly, this essay shows that no matter whether there are positive and negative shifts of latent decisions, the resulting effect is value creation. Table 2.4, under extensive Q controls in a standard clustered error regression framework, shows that latent factors possess robust positive associations with Tobin’s Q, except PSK_USE. The PSK_USE is negatively associated with Tobin’s Q. The above relationship exists over the long run that is up to three years at least¹⁶. In addition, the new factors are found to be robust negative predictors of the current and future firm cash flows, but positive predictors of firm earnings. The results are presented in Table 2.5 and 2.6. This study attributes negative association of cash flows to the accrual or real earnings management after controlling for several firm characteristics.

This study also sheds light on the latent factors’ ability to predict future stock returns. The findings are closely related to empirical literature that uses accounting information to predict stock returns¹⁷. This constitutes our second main finding as shown in Table 2.7. Table 2.7 shows that the new factors contribute strongly to the explanation of the cross-section of future excess stock returns beyond other factors known in the finance literature. It is critical to note that the newly introduced factors are not governed by stock

¹⁶ In unreported results, this study finds FIN_FLEX, ST_CREDIT, and PSK_USE robustly forecasting Tobin’s Q up to 10 lags.

¹⁷ See, for example, Ou and Penman (1989), Holthausen and Larcker (1992), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), Piotroski (2000), Sloan (1996), Spiess and Affleck-Graves (1999), Richardson and Sloan (2003), Fairfield et al. (2003), Baker et al. (2003), Hirshleifer et al. (2004), Titman et al. (2004), Daniel and Titman (2006), Bradshaw et al. (2006), Lyandres et al. (2008), Fama and French (2008), Pontiff and Woodgate (2008), Xing (2008), Cooper et al. (2008), Polk and Sapienza (2009), and Chen et al. (2011).

returns in their construction and thus are not subject to a data mining critique.¹⁸ This essay also makes sure that the information about the new factors is available prior to the return predictability testing periods. The return predictability of the new factors is economically and statistically significant. In terms of economic significance, a 1% point decrease in FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT factors increases future excess annual returns by about 3.2%, 2.4%, 2.9%, and 1%, respectively. The prominent base controls include firm capitalisation (LSIZE), the book-to-market ratio (LBTM), last six months (BHRET6) and 36 months' (BHRET36) price momentum (i.e., starting from the previous second month). The other return anomaly controls include total asset growth, cumulative accruals, accounting accruals, and investment-to-asset ratio.

In an attempt to provide further implications for investors, this study adopts a characteristics approach extended by Daniel and Titman (1997). In this test, the study controls for firm size, book-to-market ratio and momentum, and the asset pricing factors. The purpose of the controls is to obtain incremental returns after isolating the anomalous effects of controls. The findings in Table 2.8 show that portfolios based on the trading strategy of buying the lowest decile and selling short the highest decile, based on the five latent factors. The first three factors (FIN_FLEX, ST_CREDIT, and LT_INV) earn monthly characteristics-adjusted raw abnormal returns at the equal (value) weighted basis of 0.69% (0.42%), 0.14% (0.05%), and 0.63% (0.50%) respectively. The abnormal returns

¹⁸ This study applies factor analysis only to the balance sheet growth variables (X terms) and not to market values or potential dependent variables (Y terms). Factor analysis is used to produce annual factors (with available historical accounting information) instead of whole-period factors, which removes any data mining concerns in subsequent analyses that use the identified decision types as independent variables. In the return predictability test, the models are built on publicly available information about all common stocks. Moreover, there is no selection bias in the model as the study strictly follows the basic assumptions of the model (see section 2.2.2 for a detailed discussion of the research method).

are both statistically and economically significant, even after controls and up to three years into the future. Similar alphas are obtained for pricing factor models. The remaining two factors (CVT_DEBT and PSK_USE) earn abnormal returns and alphas, mainly for large stock portfolios. Furthermore, the mean equal-weighted annual buy-and-hold returns spread for FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE for the whole sample are 17%, 2.3%, 10.8%, -2.0%, and -0.3%, respectively. As for large stocks, the same is reported as 7.5%, -0.2%, 4.3%, 0.9%, and -1.2% respectively.

In summary, the empirical analysis highlights that the newly identified factors are the basic measures of corporate finance decisions. These decisions capture the management's perception of firms' financial health and future growth opportunities. The property of capturing growth opportunities translates into new factors forecasting the variations in future excess stock returns and earning abnormal returns (Kogan and Papanikolaou, 2014). Similarly, the new factors in adding value to firms fit into the manager's core objective of creating firm wealth (Myers and Majluf, 1984). This forecasting ability persists under controls of prominent accounting growth measures known in previous literature (see footnote 8). For example, this study argues that the ASSETG measure (Cooper, et al., 2008) may not take into account all data duplications and redundancies, an aggregate of the major items in the balance sheet, and may suffer from missing variable bias.¹⁹ However, this analysis captures these data shortcomings and the aggregates of all associated aspects relating to specific types of corporate financing

¹⁹ Fama and French (2008) argue that the total asset growth effect is of secondary importance as it cannot be observed in large cap firms. At the same time, Chan et al. (2007) show that asset growth takes a variety of forms like growth in cash, current assets, or long-term assets. And Cao (2011) shows that cash and operating liabilities do not contribute to the total asset growth effect.

decisions. Supporting this argument, the first three factors (FIN_FLEX, ST_CREDIT, and LT_INV) could present a true decomposition of the ASSETG. This essay further shows that FIN_FLEX is the most important component of ASSETG's return-generating effects. In conclusion, this essay provides evidence that investors' suboptimal attention to accounting information is exaggerated due to the substandard supply of the information; thus a correlational structure is the answer.

The rest of this study is organised as follows. In section 2.2, this essay discusses the data sample, model setting, and method. Section 2.3 provides a discussion of the model's outputs and descriptive statistics. Section 2.4 reports the empirical analyses where the essay relates the latent factors to firm fundamentals and stock returns using regression and portfolio tests. Section 2.5 provides the summary and conclusion.

2.2 Data Sample, Model, and Methodology

2.2.1 Data sample

This research is solely US-based and includes all firms listed on AMEX, NYSE and NASDAQ. The fiscal year balance sheet and income statement data are obtained from the Compustat database for all firms for the period 1984 to 2009. The monthly market data are obtained from the CRSP database for common stocks. In the case of balance sheet data, firms having no total-asset data are excluded from the analysis. Similarly, income statement data are subject to availability for sales (net). Both data sets are merged using PERMNO and CUSIP. The variables used in this essay are growth variables. The estimations of these variables require at least two years of data.

Consistent with the standard literature in this area, the sample excludes financial and utility firms (i.e., stocks with four-digit SIC Codes 6000-6999 and 4900-4999).²⁰ Firm size is estimated using market equity at June of the fiscal year $t-1$. The market equity in December is obtained to estimate the book-to-market ratio at the end of fiscal year $t-1$, which follows Fama and French (1992). The final data sample contains 876,348 firm-month observations for the research period.

2.2.2 Model for identifying common corporate financing decisions

2.2.2.1 Model setting

This study conjectures that there are certain corporate financing decisions that are typically embedded in a specific set of balance sheet entries (due to double-entry bookkeeping). This, in turn, causes them to share a common variance that can be identified and, meaningfully reveals the changes in corporate financing decisions. These decisions are fundamental or common to businesses.

Suppose there are $N=f\{I_1, I_2, I_3, \dots, I_n\}$ common shifts in corporate finance decisions among businesses. Each shift $I_n=f\{b_1, b_2, b_3, \dots, b_n\}$ is a function of b_n – a shift in the balance sheet items. Each b_n element is expected to have a significant correlation with other unknown items depending on a particular type of change in the corporate finance decisions and their offsetting entries on both sides of the balance sheet owing to the double-entry bookkeeping system. The elements of a balance sheet that represent a decision to adjust a certain type of corporate finance policy are expected to

²⁰ Cases in which the shares outstanding are negative or zero are also excluded. In order to calculate the book equity for the book-to-market ratio, firms with negative or zero book equity are deleted.

follow simultaneous co-movements systematically across time. Thus, the consistencies of co-movements would bring heterogeneity among corporate finance decisions.

The above assumed setting of the I_n orthogonality may appear to be violated by the balance sheet items that are aggregates of certain business activities; for example, total assets, total liabilities, long-term debt, and invested capital. In fact, it is not a violation, because aggregates are decomposed into their respective I_n decisions based on their potential uses or specific transactions. The aggregates if observed in isolation would only give the average effect of all of the components involved in their construction. The total asset is a typical example that consists of all short (long)-term investment, and financing components, whose impact would not be observed unless it is decomposed into its natural components. The decomposition of total assets into natural components is the second objective of this study.

To capture annual accounting investment growth, the individual b_n item is further converted into the growth variable $X_{i, t-1}$, that is, the ratio of change in b_n items from fiscal time ' $t-1$ ' and fiscal time ' $t-2$ ' and the lagged total assets ($Total Asset_{t-2}$).

$$X_{i, t-1} = (b_{i, t-1} - b_{i, t-2}) / Total Asset_{t-2} \quad (1)$$

Where, $b_{i, t-1}$ stands for the balance sheet item i at time $t-1$ and $b_{i, t-2}$ is the lagged balance sheet item for the same firm. The $X_{i, t-1}$ variables are winsorized annually at the 1st

and 99th percentile to remove outliers.²¹ These growth variables are then standardised by fixing their mean at zero and standard deviation as one. The purpose of standardisation is to align all the variables to the same scale. This removes any possibility of misinterpretation that may arise from the use of different units for the variables (Jobson, 1992a). In the subsequent regression analysis, standardised variables allow for the comparison of the independent variable to the dependent variable (Hair et al., 1998).

2.2.2.2 Factor analysis model

The application of the factor analysis model is an exact match to the motivation for this study. This study is motivated by a need to understand the common interactions or common causal variances of the balance sheet change rate information. Factor analysis is the right tool to understand these common interactions owing to its capabilities in isolating the common variance from the unique variance²². The commonalities exist among the balance sheet items due to the practice of the double-entry bookkeeping system. This accounting system brings offsetting changes in two or more accounts at the same time. These accounts can be on either side or both sides of the balance sheet. For example, firms use cash to run the current operations or to exploit future growth options, and is either raised via equity and debt financing or by the use of internal cash holdings. This kind of investment would generate the future cash flows that could be used for several options including: paying liabilities, distributing to shareholders as dividends, repurchasing of shares, or reinvesting in new projects. All these business activities are continuous and bring

²¹ For winsorization of 1% at both tails, Butler et al. (2011), Baxamusa (2011), Sullivan and Zhang (2011), and Teoh et al. (1998) among others are followed. The study opts for 1% winsorization instead of 5% to reduce the loss of information.

²² Factor analysis differs from the principal component analysis in its objective. The former takes into account common variances and later full variances among the original variables.

changes in several accounts, and thus they are expected to share variances. The identification of such a causal variance would actually help to identify the actual factors that reflect the basic operating business activities of these time-varying businesses. Factor analysis using the correlation matrix, reduces the original set of variables to a few common factors.

The study's motivation in using factor analysis is different from the prior literature that employs this model to the correlation matrix of stock returns to understand the multivariate behaviour of the stock price changes over time (King, 1966), and to test the arbitrage pricing theory (APT) (Lehmann and Modest, 1988). King (1966) extracts the seven factors that are in agreement with the categories suggested by the security exchange commission (SEC). Among other applications, factor analysis is used to predict the industrial bond ratings (Pinches and Mingo, 1973); to investigate the institutional investor's influence on research and development (R&D) investment behaviour of corporate managers (Bushee, 1988); and to investigate the characteristics of firms involved in mergers and acquisitions (Sorensen, 2000). Recently, Abarbanel et al. (2003) have employed factor analysis to examine the institutional investor preferences and price pressures around spin-offs (Abarbanel et al., 2003). However, this study aims to examine the correlations among the balance sheet change rates to identify the most common corporate finance decisions that represent the time-varying operating business.

The factor analysis model applied to the p observed balance sheet change rates observed variable X_1, X_2, \dots, X_p , and covariance matrix $\Sigma(p \times p)$; a group of r unobserved

variables are named common factors f_1, f_2, \dots, f_r ; and a group of p unique unobserved factors $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$, where $r \ll p$, (Jobson, 1992b). The model can be described as follows:

$$\begin{cases} X_1 = \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1r}f_r + \varepsilon_1 \\ X_2 = \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2r}f_r + \varepsilon_2 \\ \vdots \\ X_p = \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pr}f_r + \varepsilon_{1p} \end{cases}$$

The above model's notations can be re-written as:

$$X = \lambda f + \varepsilon \tag{2}$$

Where λ_{pr} is the coefficient of common factors or loading of the p th variable on the r th common factor f in the above set of equations; X is the vector of accounting variables (balance sheet growth variables); and ε is the vector of errors.

This study assumes the underlying factors are orthogonal in deriving the λ loadings. Then, equation (2) can be written in terms of the covariance matrix of standardised original variables:

$$\Sigma = \lambda' \lambda + \epsilon' \epsilon \tag{3}$$

The above equation splits the covariance matrix into two parts, $\lambda' \lambda$ are explained by the factors and $\epsilon' \epsilon$ are not explained by the factors. The covariance matrix is non-singular and symmetric, and can be reduced to a diagonal matrix Λ , when multiplied by an orthonormal matrix Ψ ;

$$\Sigma = \Psi \Lambda \Psi' \tag{4}$$

From the above equation, Λ is representing eigenvalues and Ψ the corresponding eigenvectors. Based on the principal factor method, for a given number of factors, this study solves the matrix in equation (2). Further, the orthogonal rotation method (Varimax) is adopted to better understand the meaning of the loadings on each factor. This technique designs the factor loading matrix in such a way that only one of the elements is close to 1 in a row and the rest are close to zero. This method is widely used to get simple factor structures and factor scores (Jobson, 1992b).

$$\lambda_p = \psi_p \Lambda_p^{1/2} \quad (5)$$

The application of the factor analysis is aimed at identifying the most fundamental common movements in accounting growth variables. The factor analysis is implemented using a SAS FACTOR procedure. At the start of the research, it is not obvious how many factors can be obtained from the given dataset. The exact number of factors to be extracted requires running the preliminary factor analysis model every year separately with no factor number defined.

First of all, from the preliminary factor analysis result, the basic assumptions of the model are checked. This study checks the factor analysis assumptions, including small partial correlations among the input variables; a sampling adequacy of around 0.60 and a correlation cut-off of 0.45 (Hair, et al., 1998). A sampling adequacy of 0.50 is suggested for both the overall test and individual variables and of 0.60 for successful factor analysis. Correlations below 0.30 make factor analysis inappropriate (Hair, et al., 1998). In line with the cited literature, variables that do not meet the aforementioned criteria are discarded

from the analysis. Thus, following the above assumptions the initial set of 50 variables is reduced to 37 variables. This removal of variables is followed by the identification of a number of factors.

Secondly, this study observes the *eigenvalue* – equal to one and above one – and the *scree* plot in order to determine the exact number of factors to be extracted. The *scree* plot is a plot of the eigenvalues of a correlation matrix against the ordinal number of the eigenvalues. The point of curve where it abruptly begins a linear decline is selected as the appropriate number of factors to be extracted from the factor analysis model. Moreover, the correct factor structure requires that at least three variables are grouped together with distinct loadings for individual factors. However, it is not certain that an accurate number of factors can be obtained following the above practice and so getting the right structure also depends on the researcher's judgement.

This study shows that in each year the five factors are consistently adequate in explaining most of the covariances among the original variables. Each of the five factors consists of the same components each year. This finding leads the study to the third step, where the five factors are formally extracted out of the 37 balance sheet change rates in the annual factor analysis models. Every year the five factors are given orthogonal rotation with a principal component factor analysis method, using SAS VARIMAX function, to create the final factor structure. In the last step, this study labels the factors based on the highest loadings of factors' constituent change rates. The next sub-section 2.3 discusses the outputs of the factor analysis model in detail.

2.2.3 Methods

The methods discussed under this section are aimed at understanding the implications of the common factors. Factors are identified using the factor analysis models (discussed in Section 2.3) that forecast the performance of both firm and market fundamentals. The literature survey in this study identifies a number of methods: Fama-MacBeth (FM), Fama-MacBeth with Newey-West (FM-NW), and one(two)-way standard errors clustered by the firm (CL-i), time (CL-t), intersection of firm and time (CL-2), cluster-robust errors (CLR-2), and industry (CLRi-2). These methods are discussed and compared for their performance in the accounting research by Gow et al. (2010). They find CLR-2 is necessary to produce valid inferences in accounting research. Their evidence is in contrast to Petersen (2009), who argues that clustering standard error by two-ways (both by firm and time) is unnecessary in finance research. However, this research which is a mixed bag of both accounting and finance research has two agendas. First, it intends to apply these approaches to test the implications of the new common factors in predicting performance measures. Second, this study reflects on whether the inferences of these methodological approaches are consistent with the findings of Petersen (2009) or with Gow et al. (2010).

2.2.3.1 Fama-MacBeth (FM)

Fama and MacBeth (1973) designed this approach with the aim of removing the cross-sectional correlations. Particularly, in predicting the stock returns in excess of risk-free rate, the cross-sectional regressions across each month are estimated and later the

standard errors and t -statistics are obtained by averaging the coefficient estimates across the complete sample period.

The general equation of the following form generates the monthly betas or coefficients:

$$R_{it} - R_{ft} = \alpha_0 + \sum_{n=1}^N \beta_{nt} X_{it} + \varepsilon_{it}. \quad (6)$$

The t -statistics are estimated as:

$$t = \frac{\bar{\beta}}{se(\beta)}, \quad \text{where} \quad \bar{\beta} = \frac{1}{T} \sum_{t=1}^T \hat{\beta}_t \quad \text{and} \quad se(\beta) = \frac{\sigma_{\beta}}{\sqrt{T}} \quad (7)$$

According to Gow et al. (2010) this approach is suitable when there is no time-series dependency. However, if there is cross-sectional dependence in errors and regressors, then it may cause FM errors to be understated.

2.2.3.2 Fama-MacBeth with Newey-West (FM-NW)

Following the general literature in accounting and finance, this study adds the approach extended by Newey and West (1987) (NW) to the FM approach. The NW approach is aimed at obtaining the covariance matrix that is robust to both heteroscedasticity and serial correlation. FM-NW combined would correct both for serial correlation and cross-sectional correlation. Following Gow et al., (2010), this study sets and reports the lag lengths of $T-1$. However, the results are robust to varying lag lengths. Both FM and FM-NW approaches are used in Section 2.4.2 for return predictability tests.

2.2.3.3 OLS cluster-robust standard errors

Following Huber (1967), Rogers (1983), Thompson (2011), and Petersen (2009), Gow et al. (2010) this study applies the cluster-robust standard errors approach in forecasting the firm and market fundamentals. In order to test the significance of the common factors identified in this research, the ordinary least squares (OLS) standard errors are clustered along a cross-sectional dimension (e.g., firm, labelled as CL-i) and along a time-series dimension (e.g., year, labelled as CL-t) or along both dimensions (e.g., the intersection of firm and time, labelled as CL-2). The clustering of standard error along the firm means that errors might have been correlated across time for a specific firm. The clustering of standard error along time means that errors might have been correlated across firms at a specific point in time.

The clustering of standard error either along the firm or time is referred to as a one-way cluster-robust standard errors approach. When the errors are clustered both along time and firm it is referred to as a two-way cluster-robust standard error approach. In the CL-2 approach, there is a possibility of the existence of some correlation between different firms in different time periods (also known as persistent common shocks). In an approach with no persistent common shocks (CLR-2) as given in Thompson (2011), this study estimated the two-way cluster-robust standard estimator as:

$$V = V_1 + V_2 - V_I \quad (8)$$

where, V_1 and V_2 are the variance estimates that are clustered by firm and time, respectively. The V_I is variance estimate obtained from the CL-2 approach. The

coefficients, t -statistics, standard errors, and probability values are estimated using V covariance matrix²³.

2.2.3.4 WLS (weighted-least-squares) two-way cluster-robust standard error

This technique constitutes a further robustness check of the above approaches, and the results are not reported in this thesis. This approach is applied just for return forecasting regressions. The results are very consistent with CLRI-2. This is an alternative approach for obtaining standard errors to account for group wise (cluster wise) heteroscedasticity. This heteroscedasticity may emerge due to the size difference in variance estimators. Thus, this issue can be addressed using firm size as a weighting criterion.

2.3 Model Outputs and Descriptive Statistics

2.3.1 Factor analysis model outputs

The factor analysis results are given in Table 2.1. The model produces five latent factors that are linked to corporate finance decision types labelled as: FIN_FLEX (financial flexibility), ST_CREDIT (short-term credit), LT_INV (long-term capital investment), CVT_DEBT (convertible debt usage), and PSK_USE (preferred stock usage)²⁴. The new factors are able to capture about 70% of the common variation in the aggregate balance sheet (including supplementary items) growth variables. These factors are extracted annually instead of for the whole sample period²⁵ to ensure that the identified factors are robust through time. This is also crucial for the suitability of the study's subsequent

²³ This study is using SAS macro posted at <http://acct.wharton.upenn.edu/~dtayl/code.htm> by Daniel J. Taylor for implementing these approaches.

²⁴ Labelling refers to the components with highest loads in list of each factor obtained via factor analysis.

²⁵ Please refer footnote 8 stating that whole sample period based factor analysis output is consistent with the annual sample result.

financial analyses. Based on the analysis, any increase in the number of factors beyond the five indicated is not suitable as it will add no more than 3% to 5% to the explanatory power of each additional factor annually. Furthermore, the additional factor will not satisfy the basic assumptions of the factor analysis model discussed above.

From Table 2.1, the first factor, FIN_FLEX, has the highest weight on shifts of common equity-total, common equity-liquidation, cash and equivalents, common equity-tangible (with weighting loads above 0.84). Other important elements are shifts of current assets-total, invested capital, total assets, cash, capital surplus, and common shares outstanding (with weighting loads above 0.45). The annual interactions of shifts in the above balance sheet items (capturing management's perceptions) create this factor. The factor is consistent with the definition of prior literature²⁶ and behaves as the aggregate of both financial and investment flexibility of a firm.²⁷ Over and above, this factor is always the first factor obtained annually and captures almost 40% of unobserved variations in the balance sheet information content. Thus it appears to be one of the most important decisions for managers.²⁸

The second most common annual shift in corporate finance policy among firms in the sample is a shift in short-term credit (ST_CREDIT) decision. Specifically, this factor places the highest weight on the current liabilities-total, followed by liabilities-total and

²⁶ Denis (2011) defines financial flexibility as the ability to respond in a timely and value-maximising manner to unexpected changes in the firm's cash flows or investment opportunity set. Marchica and Mura ((2010) argue "since there is no well-defined measure of flexibility in the literature, this is an unobservable factor that depends largely on managers' assessment of future growth options."

²⁷ Investment and financial flexibility are close substitutes, and the value of financial flexibility is not only dependent on the cost of external financing but also on the liquidation value of capitals (Faulkender and Wang, 2006; Gamba and Triantis, 2008).

²⁸ Graham and Harvey (2001) note that financial flexibility is the most important determinant of capital structure and DeAngelo and DeAngelo (2007) find it a missing link that is important to connect capital structure theory.

debt in current liabilities (with loads above 0.70). The remaining elements are accounts payable, total debt, notes payable, and current liabilities-others (all have weighting loads above 0.45). In the corporate finance context, this particular factor can be linked to the level of firms' operating business growth as reflected in shifts in the ability and use of short-term financing (the cost of such financing depends on the credit standing of firms).²⁹ This essay shows that firms use short-term credit as the second most important decision to fund positive NPV investments. The corporate finance literature observes that large firms have higher short-term debt borrowing than smaller firms because of their higher-credit ratings and lower information asymmetries (Abarbanell and Bushee, 1997; Barclay and Smith, 1995a; Chung, 1993; Diamond, 1991; Doukas and Pantzalis, 2003; Easterwood and Kadapakkam, 1994).

The third factor identified by the factor analysis is labelled as long-term capital investment (LT_INV). Its most important components are plant, property, and equipment gross and net (with weighting loads above 0.76). The remaining influential elements are long-term debt, total debt, invested capital and total assets (all have a load greater than 0.40)³⁰. This factor may appear similar to other known corporate capital investment proxies in the literature.³¹ However, it offers a more comprehensive measure. For example, other measures of long-term capital investment (capital investment, investment-to-assets) do not

²⁹ Contrary to Cooper et al. (2008)'s debt financing growth decomposition, ST_CREDIT is found specifically capturing short-term liabilities. Cooper et al. (2008) combine both short- and long-term debt together. In contrast, long-term debt (i.e., the part of third factor) has insignificant (-0.03) correlation with debt in short-term liabilities item, thus not part of the ST_CREDIT. The long-term debt has a high causal relationship with shifts in fixed assets (LT_INV).

³⁰ The other component include deferred taxes for showing timing differences in financial reporting for tax purposes; and investment tax credit component is generated due to new capital investment. Deferred taxes in most firms are generated by the depreciation of fixed assets, and thus, creates deferred tax liability. The invested capital and total asset indicate the investment in fixed assets to expand the firm.

³¹ See for example, Hirshleifer et al. (2004), Titman et al. (2004), Lyandres et al. (2008), and Cooper et al. (2008) among others.

take into account that the associated components of the investment in fixed assets that are hidden in different balance sheet items may be helpful in fine-tuning the measure. *LT_INV* is comparable to but different from the property, plant, and equipment growth (ΔPPE) sub-component given in Cooper et al. (2008). It is different to the investment-to-assets (*I_A*) measure extended by Lyandres et al. (2008) as shifts in inventory are not included due to the very weak correlation between the inventory and fixed assets.

The next most common factor obtained is convertible debt usage (*CVT_DEBT*). This represents the shift in the relative use of convertible debt in a firm's debt financing. The higher weights for this factor are debt convertible-total and debt convertible-preferred stock (with loads greater than 0.69). Others are debt convertible and subordinated and long-term debt (with loads about 0.46). The factor possibly reflects the financing and investment constraints of firms with unbalanced capital structure and financial distress. In such a situation, firms avoid issuing stocks and in order to raise equity resort to using convertible bonds, notes, debentures, and subordinated debt and preferred stock. These hybrid securities can also be helpful in minimising the cost of external financing while simultaneously normalising agency costs (Bodie and Taggart, 1978; Doukas and Pantzalis, 2003; Doukas and Pantzalis, 2003).

Firms that typically use convertible debt are small and micro firms with lower credit ratings, low information asymmetries, and high agency costs of debt. According to Hovakimian et al. (2001), convertible debt issuance is the second largest after equity issuance. Firms that issue convertible debt exhibit poor subsequent return performance

Table 2.1

Latent accounting growth factors identified by factor analysis

$$X_{i,t} = (b_{i,t-1} - b_{i,t-2}) / \text{Total Asset}_{t-2}$$

FACTOR 1		FACTOR 2		FACTOR 3		FACTOR 4		FACTOR 5	
Variable	Load	Variable	Load	Variable	Load	Variable	Load	Variable	Load
Common Equity (Item#A60)	0.89	Total Current Liabilities (Item#A5)	0.88	Plant, Property, & Equipment-Net (Item#A8)	0.79	Debt convertible preferred stock (Item#A39)	0.88	Preferred Stock (Item#A130)	0.89
Common equity liquidation (Item#A238)	0.88	Debt in current liabilities (Item#A34)	0.81	Plant, Property, & Equipment-Gross (Item#A7)	0.76	Total long term debt convertible (Item#A97)	0.69	Preferred Stock-Liquidation (Item#A10)	0.81
Cash and Equivalents (Item#A1)	0.86	Total Liabilities (Item#A181)	0.70	Long-Term Debt (Item#A9)	0.66	Debt convertible and subordinated (Item#A154)	0.63	Preferred Stock Nonredeemable (Item#A209)	0.58
Total current assets (Item#A4)	0.85	Notes Payable (Item#A206)	0.67	Total Debt (Item#A9+A34)	0.53	Long-term Debt (Item#A9)	0.46	Total long term debt convertible (Item#A97)	0.53
Common equity tangible (Item#A11)	0.84	Total Debt (Item#A9+A34)	0.59	Total Assets (Item#A6)	0.43	CVT_DEBT		PSK_USE	
Invested Capital (Item#A37)	0.83	Current Liabilities Others (Item#A72)	0.51	Invested Capital (Item#A37)	0.40				
Total Assets (Item#A6)	0.77	Accounts payable (Item#A70)	0.48	Intangibles (Item#A33)	0.38				
Cash (Item#A162)	0.74	Retained Earnings (Item#A36)	0.40	Deferred Taxes and investment credit (Item#A35)	0.25				
Capital Surplus (Item#A210)	0.68	Accounts receivable (Item#A2)	0.37	LT_INV					
Common shares outstanding (Item#A25)	0.46	Total Assets (Item#A6)	0.32						
FIN_FLEX		Total Inventory	0.32						
		ST_CREDIT							

This table presents the factor structure as the outcome of the application of the annual factor analysis model on the balance sheet growth rates for the period 1985-2009. It estimates the balance sheet growth rates using the following formula given in the header. Where, $b_{i,t-1}$ stands for the balance sheet item i at time $t-1$ and $b_{i,t-2}$ is the lagged balance sheet item for the same firm. Following factor analysis assumptions, from the initial set of 50 variables only 37 variables qualify as inputs to the factor analysis model. Among factor analysis assumptions, the minimum correlation of 0.30 of each variable to at least three other variables and sampling adequacy of 0.60 is targeted (Hair, et al., 1998). For each of the new latent growth factors, the table presents constituent variable names and their codes. The five factors are also given a label that reflects the constituent items that have a higher loading on that particular factor. The ‘load’ refers to the average common variable loading that appears in each factor across the whole sample period.

(Lewis et al., 2001; Stein, 1992). It is noted here that convertible debt usage (CVT_DEBT) is not a part of the total asset growth decomposition of Cooper et al. (2008). For example, CVT_DEBT's constituent variables have insignificant correlation with total asset growth (ASSETG). In other words, they are orthogonal. This study posits that this factor is a measure of the operational risk arising from the convertibility of bonds into equity.

The PSK_USE is the last factor obtained. It can be interpreted as shifts in the extent of preferred stock usage in a company. It constitutes changes in preferred stock-total, liquidation and nonredeemable (with loads greater than 0.58) (i.e., Compustat item #A130, A209 & A10) and to a lesser extent cross-loading of the total long-term debt convertible. Houston and Houston (1990) find that firms issuing preferred stock have a propensity for lower tax rates and firms investing in preferred stock are likely to have higher tax rates. Further, they find that industrial firms, not utilities, issue the majority of preferred stocks; and their major use is in the merger market by both target and acquiring firms. Target firms use it as an anti-takeover device and as an acquirer for tax benefits. Hovakimian et al. (2001) document that preferred stock issuers realise lower returns compared to non-issuers. Howe and Lee (2006) find that in the long run, this underperformance is transient and confined to small firms. Such firms have poor profitability and high debt ratios, thus they use preferred stock to offset other expensive securities.

The five factors are obtained annually, but collectively using the same input data set (the balance sheet growth items or components). This lets the components decompose and align with the respective causal factors. Thus, factor components are neither mutually exclusive nor are redundant. The same component appearing in two factors may not be

categorised as redundant in the components as the model separates the common components of the growth item and aligns it to other close causal components. The original variables that have cross-loadings shown in Table 2.1 are actually aggregates of several business activities. The dataset providers (Compustat) accumulate several items to construct an aggregate item such as total assets, long-term debt, and total liabilities. For example, total assets is an aggregate of cash, non-cash, and fixed assets, thus appear to decomposed in first three factor outputs. Similarly, long-term debt could be raised for paying the current liabilities, purchasing fixed assets, and for risk management (hybrid assets).

Alternatively, long-term debt could be straight or convertible and both may be used by firms for different investment options. The straight source of long-term debt may be employed to fund the fixed assets and the convertible debt can be converted to the company's common or preferred stocks, or it may be used for altogether different considerations. Similarly, in the case of the second factor, the total liabilities and its various components appearing together may appear as redundant, but it is not. This factor takes into account mostly the current liabilities related components of the total liabilities, and not the long-term. The long-term obligations are closely aligned with the fixed assets and non-current liabilities, so they appear in factor three. Consequently, the cross-loading observed in Table 2.1 for total assets and long-term debt growth variables is not shocking. In fact, it indicates the true decomposition of these accounting aggregates in line of their potential type and use. It is also providing support to motivation in this study that factor analysis is capable to capture the causal relations (correlations) of balance sheet items and categorize them into their natural groups or factors.

Furthermore, the factor analysis output given in Table 2.1 may be confused with cash and accrual accounting literature. The balance sheet components in the five factors may appear to overlap with the accrual measure defined in Richardson et al. (2005) but it is different because factor in this study reflect the investment and financing activities not the accruals.³² According to Lewellen and Resutek (2013) investment, external financing, and accruals all have dissimilar characteristics. Even the accrual measure as in Richardson et al. (2005) have no significant high correlation with any of the new latent factors. It is because of a predetermined mechanical way to obtain accrual measure, which creates missing variables biases. The factor analysis approach captures the missing variables and resultant factors are different than measures of investment, financing, and accruals in prior literature.

Similarly, Bradshaw et al. (2006) construct a measure of companies' external financing activities that is negatively related with subsequent stock returns. Again, this study does not distinguish in advance whether the particular transaction or decision is either a financing or an investment activity and just lets the model take into account the correlation structure of information content and make unobserved associations. Thus, a predetermined or mechanical construction of a particular measure that may cause selection biases is avoided. Alternatively, the study does not discriminate whether the information is internal or external, as it is the causal or interactive structure of the information content that helps to decompose the information into the natural groupings. This grouping regularly

³² Their measure expands beyond the working capital accruals. The expanded measure of accruals robustly predicts the subsequent stock returns and this predictability stems from the mispricing of the accrual components that are a function of reliability. The subjective measures of accruals are less reliable and less persistent.

persists in annual cross-sections and has a particular hierarchy of decisions in the order of importance to the managers or firms.

Although it is known that factor analysis is a pure statistical technique, the study of outcomes suggests that the accounting concepts of both value relevance and reliability are covered in the real sense. The balance sheet items that are direct measures bear high loading and the items that are subjective have lower loadings. For example, accounts receivables and inventory in factor two (ST_CREDIT), as well as payables have high loadings because the model considers the adequacy of estimates as one of the basic assumptions, which is consistent with Richardson et al. (2005). It is again interesting to observe that growth in accounts receivables and inventory are part of ST_CREDIT, but not part of the FIN_FLEX, when FIN_FLEX predominately consists of current assets.

Apart from the accrual accounting, Dechow et al. (2008) decompose the cash component of earnings into its three sub-components, (i) change in the cash balance, (ii) distributions to debt holders, and (iii) distributions to equity holders. They find that both equity and debt distributions are priced, whereas the change in the cash balance is mispriced. Though this study does not decompose the cash component of earnings, the model does group the components that are highly causal to the cash component. Dechow et al. (2008) show that the greater persistence of the cash component of earnings is solely attributable to cash that is distributed to equity holders. The first factor FIN_FLEX, exclusively consists of shifts in cash balance and equity components, that add to Dechow et al.'s (2008) findings by showing the evidence directly observing the covariance structure. However, it is noted that the part of cash components common to equity makes the

FIN_FLEX factor, but the retained earnings part is more aligned with ST_CREDIT. This is why in the unreported stability analysis, cash balance and equity components are found to have linear relations with each other and with FIN_FLEX, and retained earnings have a non-linear relation with FIN_FLEX, but a linear one with ST_CREDIT factor and its components. Then, it could be argued that financing activity based cash is made part of FIN_FLEX and operating activity-based cash makes part of ST_CREDIT.

2.3.2 New factors time-series stability

It is critical to ensure that the factors are stable across time. To do this, the components of each factor in each of the 25 years covered in the study are checked and the factor components are indeed persistent. For example, the top five components of the financial flexibility and short-term credit factors are always the same when their components are listed. The top four components of the other three factors are always the same when the components of those factors are listed. The ranking of all factors except the financial flexibility (FIN_FLEX) factor displays minor variations through time. In another stability check, the study finds that all the components are linearly related to their own factor and, at the same time, are non-linearly related to external factors (unreported). The stability of the identified factors reveals the real economic forces that explain the changes in the corporate finance policy of firms over time.

2.3.3 Descriptive statistics

For each fiscal year from 1985 to 2009, the study allocates NYSE, Amex, and NASDAQ firms into deciles by ranking variables. The ranking variables are the new latent growth factors of financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-

term investment (LT_INV), convertible debt use (CVT_DEBT), and preferred stock usage (PSK_USE). The data sample consists of 876,348 non-financial and non-utilities firm-month observations starting from July 1986 and continuing until December 2010. Table 2.2 reports the time-series means and medians of accounting and market characteristics that can define the nature of firms belonging to the lowest and highest deciles. Table 2.3 reports both Pearson and Spearman correlation coefficients with relative p-values of the latent factors and other prominent characteristics.³³

Firms in the high FIN_FLEX decile have higher values for total asset growth, net operating assets, investment-to-assets, profitability, market size, cash flows, firm value, cash balances, and dividend-to-capital ratio; and lower values for accounting accruals, leverage, book-to-market ratio, earnings, and low information uncertainty than those in the low decile.³⁴

In the case of short-term credit (ST_CREDIT), firms that increase short-term credit show a high level of total asset growth, greater firm value (Tobin's Q), a higher dividend to capital ratio, and higher cash balances than do firms that decrease short-term credits. Firms in the high ST_CREDIT decile are also marginally higher in net operating assets, accounting accruals, investment-to-assets, profitability, leverage, earnings, and market size. The same firms exhibit lower book-to-market ratios and cash flows. Firms that increase (decrease) short-term credit have similar levels of information uncertainty.

³³ The definitions of the complete set of market and accounting characteristics (variables) examined in Tables 2.2 and 2.3 are in appendix A.

³⁴ Here and after, high decile refers to most positive change and low decile refers to most negative change.

Table 2.2**Mean (median) values of selected characteristics for decile portfolios sorted by latent factors**

FLEX_rank		<i>FIN_FLEX</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>	TA (\$m)	MV (\$m)	Earnings	CFlows	Cash/TA _{t-1}	Div/K	Inv_mv
Low	Mean	-0.62	0.09	0.63	-0.32	0.14	-0.23	0.24	4.18	0.45	2.78	672	664	-0.14	0.18	0.20	0.33	0.07
	Median	-0.54	-0.12	0.48	-0.16	0.03	-0.12	0.19	4.04	0.38	1.55	49	57	-0.11	0.06	0.13	0.00	0.02
High	Mean	1.85	3.27	1.46	-0.71	0.40	-0.02	0.09	4.97	0.31	4.27	345	955	-0.26	0.45	1.93	1.63	0.03
	Median	1.10	1.16	0.74	-0.12	0.11	0.05	0.02	4.91	0.25	2.68	57	135	0.10	0.17	0.88	0.00	0.01
STC_rank		<i>ST_CREDIT</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>	TA (\$m)	MV (\$m)	Earnings	CFlows	Cash/TA _{t-1}	Div/K	Inv_mv
Low	Mean	-0.91	0.95	1.01	-0.51	0.16	-0.03	0.22	4.43	0.49	2.54	970	1009	-0.12	0.39	0.62	0.18	0.11
	Median	-0.72	0.06	0.66	-0.20	0.02	0.05	0.18	4.22	0.41	1.54	60	68	0.05	0.24	0.10	0.00	0.08
High	Mean	1.02	1.94	1.25	-0.42	0.37	-0.02	0.26	4.62	0.41	2.81	971	1247	-0.09	0.34	0.82	1.34	0.11
	Median	0.71	0.49	0.84	-0.11	0.12	0.05	0.24	4.40	0.34	1.65	73	82	0.08	0.15	0.13	0.00	0.07
LTI_rank		<i>LT_INV</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>	TA (\$m)	MV (\$m)	Earnings	CFlows	Cash/TA _{t-1}	Div/K	Inv_mv
Low	Mean	-0.90	0.76	0.54	-0.62	-0.04	-0.06	0.21	4.47	0.49	2.75	711	699	-0.13	0.48	0.89	0.42	0.12
	Median	-0.75	0.00	0.52	-0.22	-0.04	0.01	0.16	4.39	0.39	1.54	69	81	0.01	0.23	0.15	0.00	0.08
High	Mean	1.88	2.31	1.86	-0.43	0.62	0.03	0.31	5.19	0.44	2.34	1382	1570	-0.07	0.36	0.69	1.22	0.09
	Median	1.34	0.65	1.16	-0.20	0.33	0.06	0.31	5.08	0.38	1.56	164	160	0.11	0.30	0.10	0.00	0.06
CVT_rank		<i>CVT_DEBT</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>	TA (\$m)	MV (\$m)	Earnings	CFlows	Cash/TA _{t-1}	Div/K	Inv_mv
Low	Mean	-1.06	1.27	1.15	-0.44	0.36	0.00	0.22	5.07	0.42	2.94	1032	1265	-0.07	0.37	0.64	0.46	0.11
	Median	-0.65	0.31	0.80	-0.19	0.12	0.06	0.19	4.97	0.35	1.76	106	144	0.08	0.26	0.14	0.00	0.07
High	Mean	1.45	1.54	1.18	-0.45	0.15	-0.03	0.31	5.04	0.44	2.48	1370	1590	-0.13	0.32	0.78	0.25	0.11
	Median	0.47	0.32	0.75	-0.17	0.04	0.04	0.31	4.96	0.36	1.55	139	142	0.05	0.20	0.14	0.00	0.07
PSK_rank		<i>PSK_USE</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>	TA (\$m)	MV (\$m)	Earnings	CFlows	Cash/TA _{t-1}	Div/K	Inv_mv
Low	Mean	-1.23	1.53	1.09	-0.26	0.31	-0.03	0.28	5.04	0.41	2.88	1149	1256	-0.15	0.11	1.00	1.38	0.11
	Median	-0.37	0.46	0.83	-0.14	0.09	0.05	0.27	4.94	0.35	1.68	107	140	0.07	0.18	0.15	0.00	0.07
High	Mean	0.68	1.79	1.21	-0.53	0.24	-0.05	0.18	4.81	0.39	3.21	801	1098	-0.12	0.42	0.87	0.54	0.13
	Median	0.26	0.37	0.68	-0.18	0.05	0.04	0.11	4.68	0.32	1.91	72	108	0.05	0.21	0.26	0.00	0.09

Mean (median) values of selected characteristics for decile portfolios ranked by latent factors (*FIN_FLEX*, *ST_CREDIT*, *LT_INV*, *CVT_DEBT*, *PSK_USE*). *LBTM* is a logarithm of book-to-market ratio. *LSIZE* is the log of the market value of equity. *ACCR* is the accounting Accruals estimated using Cooper, et al. (2008). *NOA* represents cumulative Accruals as defined by Hirshleifer, et al. (2004). *ASSETG* is the year-to-year change in total assets as defined by Cooper, et al. (2008). The measure of profitability (*ROA*) is net income divided by the total assets. *LEV* is the sum of long-term debt and debt in current liabilities, scaled by total assets as defined by Cooper, et al. (2008). *I_A* is the sum of change in inventories and change in gross property, plant, and equipment (*PPE*) scaled by lagged total assets. *Q* is the Tobin's *Q* a proxy for firm value. *TA* (\$m) is the total assets in million dollars. *MV* (\$m) is the market value of equity in million dollars. *Earnings* is the measure of firm earnings defined as in Hirshleifer et al. (2004). *CFlows* is the measure of firms' cash flows estimated as the difference between *Earnings* and *Accruals*. *Cash/TA_{t-1}* is the proxy for cash holdings, the ratio of the cash and equivalents to the lagged total assets. *Div/K* is the ratio of the dividends-common to the net plant, property, and equipment. *Inv_mv* is a proxy for the information uncertainty, measured as the inverse of the market value of equity. The latent factors are obtained annually via application of a factor analysis model; the list of components of each latent factor is given in Table 1. The detailed description of characteristics is given in appendix A.

Table 2.3

Pearson (Spearman) correlation coefficients between latent accounting factors and other characteristics

	<i>FIN_FLEX</i>	<i>ST_CREDIT</i>	<i>LT_INV</i>	<i>CVT_DEBT</i>	<i>PSK_USE</i>	<i>ASSETG</i>	<i>NOA</i>	<i>ACCR</i>	<i>I_A</i>	<i>ROA</i>	<i>LEV</i>	<i>LSIZE</i>	<i>LBTM</i>	<i>Q</i>
<i>FIN_FLEX</i>		0.058	-0.061	0.013	-0.151	0.209	0.091	0.018	0.161	0.017	-0.190	0.002	-0.147	0.152
		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0443	<.0001	<.0001
<i>ST_CREDIT</i>	0.026		0.066	0.006	-0.012	0.063	0.037	0.043	0.140	0.017	0.079	-0.026	-0.061	0.019
	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
<i>LT_INV</i>	-0.210	-0.051		0.011	0.000	0.099	0.117	0.000	0.294	0.049	0.194	0.060	-0.048	-0.013
	<.0001	<.0001		<.0001	0.644	<.0001	<.0001	0.7366	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
<i>CVT_DEBT</i>	0.016	0.012	-0.212		0.014	0.026	0.014	0.003	0.001	-0.020	0.158	0.012	-0.023	-0.008
	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	0.011	0.5556	<.0001	<.0001	<.0001	<.0001	<.0001
<i>PSK_USE</i>	0.194	-0.036	-0.169	-0.206		0.003	0.037	-0.008	0.005	0.030	0.026	-0.011	0.032	-0.047
	<.0001	<.0001	<.0001	<.0001		0.0056	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
<i>ASSETG</i>	0.468	0.275	0.385	-0.086	0.000		0.785	-0.077	0.397	-0.018	-0.012	0.003	-0.029	0.026
	<.0001	<.0001	<.0001	<.0001	0.9888		<.0001	<.0001	<.0001	<.0001	<.0001	0.0059	<.0001	<.0001
<i>NOA</i>	0.078	0.088	0.422	-0.079	-0.089	0.495		-0.054	0.368	0.016	0.029	0.001	0.000	-0.007
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	0.5966	0.7727	<.0001
<i>ACCR</i>	0.095	0.119	0.053	-0.015	-0.060	0.205	0.094		0.113	0.012	-0.003	0.007	-0.020	0.008
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.0013	<.0001	<.0001	<.0001
<i>I_A</i>	0.096	0.171	0.623	-0.262	-0.077	0.682	0.515	0.219		0.014	0.030	0.003	-0.035	0.019
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.0023	<.0001	<.0001
<i>ROA</i>	0.279	0.021	0.118	-0.101	0.061	0.295	0.252	0.024	0.238		0.068	0.271	0.020	-0.169
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
<i>LEV</i>	-0.213	0.051	0.129	0.120	-0.149	-0.035	0.422	-0.081	0.046	-0.002		-0.011	0.131	-0.163
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.1037		<.0001	<.0001	<.0001
<i>LSIZE</i>	0.112	-0.012	0.131	-0.069	-0.012	0.161	0.010	-0.040	0.107	0.359	0.008		-0.344	0.153
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001
<i>LBTM</i>	-0.185	-0.082	-0.070	0.109	-0.059	-0.289	0.108	-0.124	-0.179	-0.178	0.162	-0.342		-0.334
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
<i>Q</i>	0.227	0.052	0.061	-0.132	0.068	0.289	-0.099	0.110	0.176	0.212	-0.277	0.365	-0.852	
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

The table reports both Spearman (Pearson) correlation coefficients in lower (upper) triangles. Bold correlation coefficients indicate the significance equal to or less than the 5% level. The p-values are given in italic form below the correlation coefficients. The detailed description of the variables is given in appendix A.

On average the mean and median statistics across LT_INV ranked deciles show that the firms in the high LT_INV decile have higher values in relation to: total asset growth, net operating assets, accounting accruals, investment-to-assets, profitability, leverage, earning, dividend-to-capital ratio, and greater accounting and market size as compared to the low decile LT_INV firms. Additionally, high LT_INV decile firms exhibit lower book-to-market ratio, firm value, and cash balances than low LT_INV decile firms exhibit.

For convertible debt usage, the low (high) decile stocks have similar *mean (median)* values for total asset growth, net operating assets, accounting accruals, size, book-to-market ratio, firm value, cash flows, cash balances and information uncertainty. The two deciles differ across the other characteristics; for example, investment-to-assets, profitability, earnings, and dividend-to-capital ratio is higher for high decile stocks, and leverage is the only characteristic to have a lower value in high decile stocks than in low decile stocks. Finally, for the preferred stock usage ranked portfolios, except the total asset growth, net operating assets, leverage, Tobin's Q, earnings, cash flows, and information uncertainty; the other characteristics have higher mean and median values in the low decile stocks compared to high decile stocks. However, variation in the values of characteristics is marginal across the two deciles.

Table 2.3 shows the Pearson (lower triangle) and Spearman (upper triangle) correlations for the five latent growth factors and the other prominent characteristics. The correlation coefficients are not big enough to create a multicollinearity problem. The significant correlation between total asset growth and the first three latent factors shows the true decomposition of the total asset growth variable. The large correlation values observed

among the latent factors and other characteristics are Pearson correlations; for example, FIN_FLEX has a correlation of 0.47 with ASSETG and LT_INV has a correlation of 0.42 with NOA, 0.39 with ASSETG, and 0.62 with I_A.

Overall, the characteristic differences across the two extreme deciles are distinct for FIN_FLEX and LT_INV but are marginal for ST_CREDIT, CVT_DEBT, and PSK_USE. Thus, the study expects to see large variations in returns across the two extreme portfolios based on FIN_FLEX and LT_INV as compared to the other three latent factors. A similar pattern may be observed in regression analysis for return predictability.

The next important question is whether these factors have predictive power for equity returns and firm fundamentals. The existence of such predictability will endorse the economic implications of these latent factors. Such implications are empirically investigated in the next section.

2.4 Empirical Analyses

2.4.1 Does latent accounting growth matter to the fundamentals of business?

This section tests whether the new factors (corporate financing and investment decisions) create or destroy firm value. A firm's decision to undertake positive net present value (NPV) projects in the absence of agency issues (Myers and Majluf, 1984) and financing constraints (Almeida et al., 2011), should be a value-creating activity. The factors under consideration capture the systematic changes in investment and financing decisions across two consecutive periods, where decisions are based on several characteristic states

(Gamba and Triantis, 2008). Although the objective of this research is not to test why firms take such decisions but to look into the shifts of particular decisions, the characteristics' uncertainties (Gamba and Triantis, 2008) and agency and informational asymmetric issues (Myers and Majluf, 1984); financial constraints (Almeida, et al., 2011) are assumed to be captured by the new latent decision measures. Thus, this study is intended to observe the resulting effect on the value of the firm. However, the latent factors are expected to be wealth creating activities based on the Q theory that the valuations (Tobin's Q) have a positive relationship with investment (Cochrane, 2011). Similar to overall investment, the distinct segments of investment (five latent factors) are function of the Q as given in Cochrane (2011), $1 + \alpha \frac{i_t}{A_t} = \frac{\text{market value}_t}{\text{book value}_t} = Q_t$, where i_t = investment and A_t = assets.

In the following two sub-sections, this essay tests two conjectures. The first to be tested is the Q theory implication that on average firms have higher Q values at times when there is an increase in latent investment and financing decision measures (positive change values) than at times when there is a reduction in latent investment and financing decision measures (negative change values). Second, on average these latent investment and financing decision measures have a positive and long-term relationship with firm value. It is because of the intertemporal and forward-looking nature of the measures. The new factors capture information about firm decisions between time $t-2$ and time $t-1$ with the implications of these decisions for time t . The Tobin's Q is used as a proxy for the market value of the firm in empirical tests and is measured as market equity plus total assets minus book equity scaled by total assets, as defined in Hou and Robinson (2006).

2.4.1.1 Q theory implications

Figure 2.1 (a-e) provides a comparison across the two samples. The first sample has only positive values (greater than zero) of the respective latent growth measures (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE) and the second sample has negative values (less than zero). The figure shows the average annual Tobin's Q for firms in these two samples for the period 1985 to 2009. The difference between positive and negative values is shown as a spread, which is expected to be positive. Apart from the annual trend, averages over the sample are also checked along with their respective *t*-value for significance tests (unreported table).

Figure 2.1 (a & b) shows that the average value of Tobin's Q in each of the 25 years for positive financial flexibility (Pos_FLEX) and positive short-term credit (Pos_STC) is always higher than that for negative financial flexibility (Neg_FLEX) and negative short-term credit (Neg_STC). Thus the spread is also significantly positive. Across the sample period, the average Q values for Pos_FLEX (Neg_FLEX) are 3.24 (1.76) with a spread value of 1.48, showing that firms with positive financial flexibility have on average 84% higher firm value (Qs) than those that have negative financial flexibility. The sample average Qs for Pos_STC (Neg_STC) are 2.3 (2.0) with a spread of 0.4, showing on average 19% higher firm value to firms that increase short-term credit over those that decrease it.

The Q value trend for the remaining three latent growth factors (LT_INV, CVT_DEBT, and PSK_USE), shown in figure 1 (c & d), does not exactly follow the Q theory of investment predictions: in certain years, the spread is negative. For LT_INV, we observe: that both negative long-term investments (Neg_LTI) and positive long-term

investments (Pos_LTI) have similar levels of Q values. However, firms in the Neg_LTI sample have higher Q values than do firms in the Pos_LTI sample for the years 1995, 1997, 1999 and 2006. The full sample average Q for Pos_LTI is 2.1; and for Neg_LTI, it is 2.0, with a spread of only 6%. In the case of debt convertible use (CVT_DEBT) the resulting effects of investment decisions on firm value are contrary to the Q theory of investment. In 17 years out of the 25 years firms that increase convertible debt investment (Pos_CVT) show lower Q values than do firms that decrease the debt convertible investment (Neg_CVT). Across the sample, the average Q value spread of the two samples is -0.2, indicating that firms in the Pos_CVT sample have 8% less firm value than firms in the Neg_CVT.

In regard to the investment decision to use preferred stock (PSK_USE), shown in figure 2.1 (e), firms in the Neg_PSK sample have higher firm values in seven years, but on average firms in the Pos_PSK sample have higher firm values. This trend is also observed in the whole sample. The average Q value is 2.28 for Pos_PSK and 2.01 for Neg_PSK, with a spread of 14%. Over the sample Q values for positive and negative samples and the spread for the five latent accounting growth factors is significant at the 1% level. Consequently, the tests suggest that on average, the decision to increase investment (latent factors) increases firm value, except for CVT_DEBT.

Finally, figure 1 (a-e) shows that from 1999-2000, there are peaks across all factors. These peaks are indicating the telecom bubble decline back to their normal levels in the year 2001 and after. Similarly, there are dips from 2007-2008 owing to the global financial crisis. Interestingly, in both cases firm values remain positive.

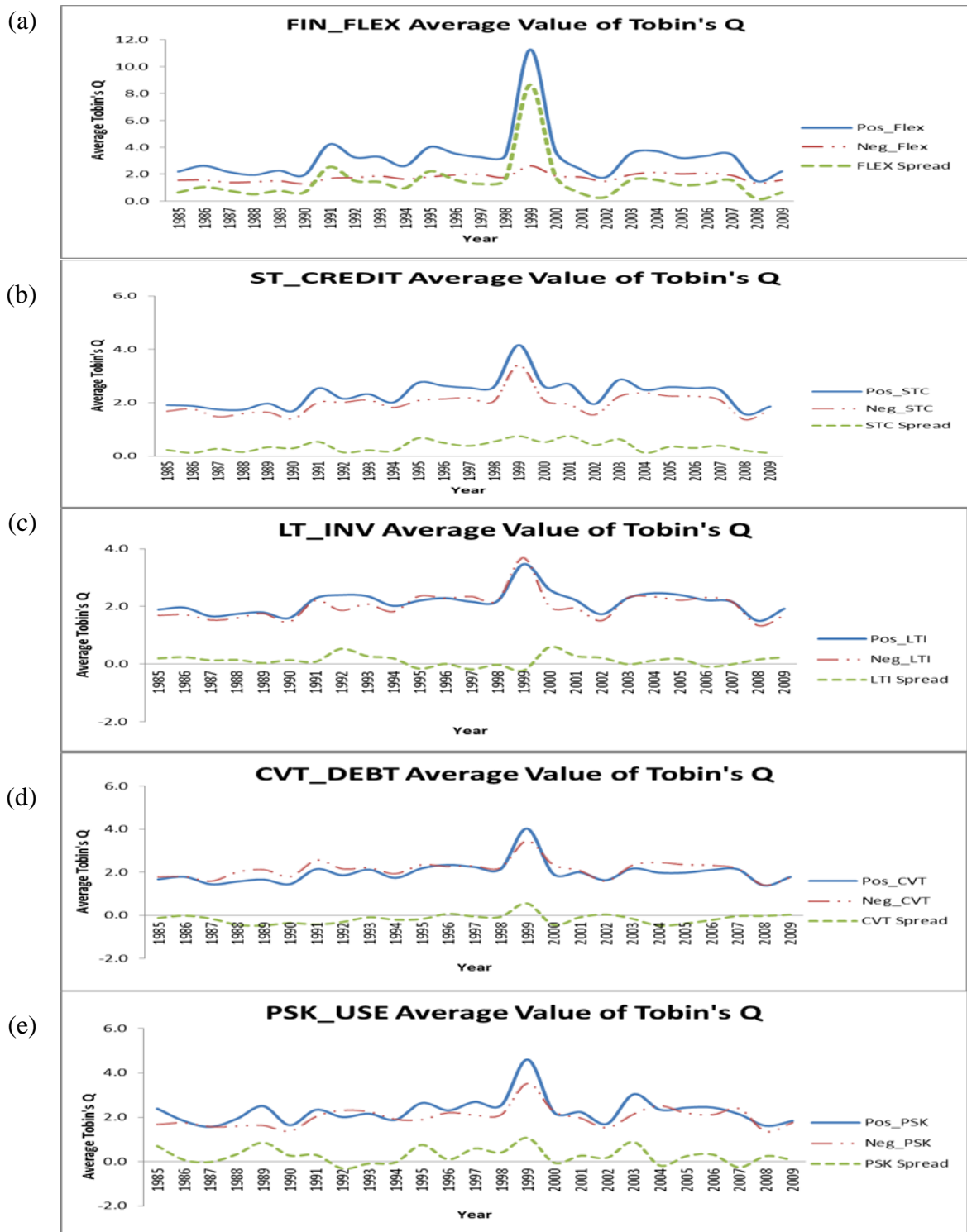


Figure 2.1 (a-e): Q theory implications

The figures show the average annual Tobin's Q for firms that have negative (positive) values for latent accounting growth factors (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE) separately. Also shown is the difference between positive and negative values as a spread. For the period 1985 to 2009, each year the sample is divided into the negative sample if firms have values less than zero and positive sample if firms have values greater than zero.

2.4.1.2 Predicting firm value

This sub-section tests the conjecture that on average these latent investment and financing decisions, being intertemporal in nature and forward looking, have a positive and long-term association with firm value. To infer whether latent factors increase firm value or not, the study employs the OLS cluster-robust standard error approach as described in Section 2.2.3. In addition to the clustering of the standard errors along cross-sectional (e.g., firm), and time-series (e.g., time) dimensions, this study also takes into account the persistent common shocks (e.g., both firm and time), and industry effect.

The dependent variable that measures firm value is the Tobin's Q. The independent variables (IVs) are five latent growth factors (LatentGrowth) and set of control variables (Controls). LatentGrowth refers to the set of new latent factors FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE. Controls refers to the control variables that are constructed following Allayannis and Weston (2001). Among the control variables, Inv_growth is R&D expenditure to sales, to show that firm value depends on future investment growth opportunities. Access_fm is a dummy for access to financial markets, which is equal to 1 if the firm paid a dividend in the current year. The dividend dummy indicates that firms that pay a dividend are less financially constrained and will have lower firm value. Size is a log of total assets that controls for the effect of size. LEV is the sum of long-term debt and debt in current liabilities scaled by total assets, as defined by Cooper et al. (2008). ROA is net income (Compustat item #A178) divided by total assets (Compustat item #A6).

The data sample covers the period from 1985 to 2009 with 876,348 firm-month observations excluding utilities and financial firms. The firms with net sales less than \$10 million are also deleted. The results are presented in Table 2.4. The study estimates the following regressions:

$$Q_{it+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{it+1} \quad (9)$$

$$Q_{it} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{it} \quad (10)$$

$$Q_{it} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-n} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{it} \quad (11)$$

Panel A, Table 2.4 shows results from estimating the model given in equation (9) using one-year ahead Q_{t+1} as a dependent variable. The LatentGrowth and Controls are observed at year t . In the first five models (M-1 to M-5), each of the five factors is included separately with the controls in forecasting the firm's value. The third column titled Coeff. is presenting parameter coefficients and in the next four columns the t -statistics are presented for firm-fixed effect (CL-i), time-fixed effect (CL-t), the intersection of firm and time fixed effect (CL-2), and CLR-2 showing t -statistics when persistent common shocks are avoided. Columns 8 and 9 present the parameter coefficients and t -statistics for the sixth model (M-6), which include all five latent factors along with the control variable set. Finally, when M-6 is augmented with the dummy variables representing the nine industries (CLRi-2), the resulting statistics of this model (M-7) are presented in columns 10 and 11. The sample tested in Panel A, constitutes 710,951 firm-month observations, 7,536 firm clusters, 24 time-series, and 58,425 clusters of firm-time intersections.

Similar to Panel A, Table 2.4 presents results for the same seven models across Panels B to E. In Panel B, both dependent and IVs are observed concurrently. In Panel C,

this study introduces the one-year lag in the IVs. Similarly, the two-year lag is introduced in IVs in Panel D, and the three-year lag in Panel E. The details, such as number of firms, firm cluster, number of time-series, and number of intersection clusters, are shown in each panel³⁵. Three lag values are used to test the long-term effects of these latent decisions on firm value. The table reports results only for the LatentGrowth variables.

Panel A, Table 2.4 shows in M-1 to M-7 that except for LT_INV all other factors have robust firm value forecasting power. The FIN_FLEX, ST_CREDIT, and CVT_DEBT are positively related with Tobin's Q, whereas PSK_USE is negatively related. The result is robust even to industry effect. Panel B, Table 2.4 indicates that in the concurrent time regressions all latent corporate financing and investment decisions robustly increase firm value, except PSK_USE. The coefficients (*t*-statistics) of FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE in CLR-2 are 0.681 (4.30), 0.229 (6.40), 0.089 (4.11), 0.044 (1.43), and -0.243 (-2.47) respectively, all significant at the 1% level. The results for controls show that all have expected signs and significance except Size, which is insignificant in explaining firm value. In the subsequent panels, the CVT_DEBT becomes insignificant to two-year lag in Panel D, and in Panel E, both the LT_INV and CVT_DEBT become insignificant to three-year lag. The other three factors FIN_FLEX, ST_CREDIT, and PSK_USE are persistent in explaining the firm-value up to the three-year lag. The evidence suggests that management's investment and financing decisions as a whole are value-creating activities when considered at the unobserved level. Moreover, these results show that two-way standard error clustering is unnecessary as in Petersen (2009, 473) even in accounting as CLR-2 results are very similar to CL-t.

³⁵ This study ran a similar regression by introducing 10 annual lags in IVs. The unreported results suggest that FIN_FLEX, ST_CREDIT, and PSK_USE consistently explain firm value up to 10 years.

Table 2.4

Forecasting firm value

$$Q_{it+1} = \alpha + \beta_1 LatentGrowth_{i,t} + \beta_2 Controls_{i,t} + \varepsilon_{it+1}$$

$$Q_{it} = \alpha + \beta_1 LatentGrowth_{i,t} + \beta_2 Controls_{i,t} + \varepsilon_{it}$$

$$Q_{it} = \alpha + \beta_1 LatentGrowth_{i,t-n} + \beta_2 Controls_{i,t} + \varepsilon_{it}$$

Panel A										
Dependent: Q_{t+1} (n= 710951, Cli=7536, CLt=24, CL2=58425, t=1985-2009)										
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6	Coeff.	CLRi-2 (t-stat) M-7
M-1	FIN_FLEX	0.3315	10.16***	4.21***	10.22***	4.20***	0.3110	3.88***	0.2897	3.77***
M-2	ST_CREDIT	0.1352	9.09***	5.44***	10.43***	5.22***	0.0819	3.54***	0.0694	3.12***
M-3	LT_INV	0.0066	0.74	0.37	0.84	0.36	0.0157	1.16	0.0178	1.59
M-4	CVT_DEBT	0.0397	4.61***	3.58***	4.79***	3.50***	0.0284	2.79***	0.0160	1.72*
M-5	PSK_USE	-0.1049	-7.21***	-6.28***	-6.88***	-6.53***	-0.0514	-4.42***	-0.0455	-3.87***
Panel B										
Dependent: Q_t (n= 777518, Cli=8659, CLt=25, CL2=67848, t=1985-2009)										
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6	Coeff.	CLRi-2 (t-stat) M-7
M-1	FIN_FLEX	0.6812	9.72***	4.29***	9.63***	4.30***	0.6434	4.23***	0.6215	4.13***
M-2	ST_CREDIT	0.2287	9.72***	6.57***	10.37***	6.40***	0.1101	3.40***	0.0985	3.02***
M-3	LT_INV	0.0896	4.86***	4.86***	5.09***	4.11***	0.0970	5.77***	0.0941	6.53***
M-4	CVT_DEBT	0.0437	2.70***	1.44	2.78***	1.43	0.0194	0.69	0.0066	0.28
M-5	PSK_USE	-0.2432	-6.60***	-2.46***	-6.57***	-2.47***	-0.1374	-1.83***	-0.1327	-1.81***
Panel C										
Dependent: Q_t (n= 722609, Cli=7676, CLt=24, CL2=59281, t=1985-2009)										
Models	IV_{t-1}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6	Coeff.	CLRi-2 (t-stat) M-7
M-1	FIN_FLEX	0.3169	8.86***	3.77***	9.52***	3.72***	0.3072	3.60***	0.2807	3.34***
M-2	ST_CREDIT	0.1375	7.32***	5.16***	8.08***	4.94***	0.0823	2.63***	0.0722	2.33**
M-3	LT_INV	0.0369	3.39***	1.82*	3.85***	1.82*	0.0402	2.26**	0.0350	2.33**
M-4	CVT_DEBT	0.0201	2.02**	2.46**	2.06**	2.40**	0.0121	1.78*	-0.0003	-0.05
M-5	PSK_USE	-0.1152	-5.75***	-4.82***	-5.53***	-4.82***	-0.0622	-3.51***	-0.0607	-3.56***

Table 2.4 (Continued)

Panel D		Dependent: Q_t (n= 632259, Cli=6815, CLt=23, CL2=51801, t=1985-2009)									
Models	IV_{t-2}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6	Coeff.	CLRi-2 (t-stat) M-7	
M-1	FIN_FLEX	0.2791	7.24***	6.33***	7.28***	6.31***	0.2707	5.95***	0.2538	5.92***	
M-2	ST_CREDIT	0.1092	5.01***	4.47***	5.50***	4.20***	0.0652	3.11***	0.0569	2.69***	
M-3	LT_INV	0.0387	3.44***	3.25***	3.88***	2.98***	0.0435	3.85***	0.0391	4.15***	
M-4	CVT_DEBT	0.0003	0.03	0.03	0.03	0.029	-0.0028	-0.30	-0.0149	-1.57	
M-5	PSK_USE	-0.0729	-4.21***	-3.96***	-4.06***	-4.11***	-0.0261	-1.28	-0.0234	-1.14	
Panel E		Dependent: Q_t (n= 556658, Cli=5983, CLt=22, CL2=45555, t=1985-2009)									
Models	IV_{t-3}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6	Coeff.	CLRi-2 (t-stat) M-7	
M-1	FIN_FLEX	0.2352	8.16***	3.66***	8.57***	3.62***	0.2222	3.59***	0.2039	3.46***	
M-2	ST_CREDIT	0.1213	4.63***	3.46***	4.82***	3.38***	0.0873	2.83***	0.0819	2.87***	
M-3	LT_INV	0.0222	1.87*	1.00	2.05**	0.976	0.0278	1.30	0.0247	1.28	
M-4	CVT_DEBT	0.0068	0.7	0.91	0.64	1.1346	0.0032	0.5203	-0.0068	-1.10	
M-5	PSK_USE	-0.0872	-3.49***	-3.54***	-3.48***	-3.54***	-0.0511	-2.11**	-0.0491	-2.07**	

The table reports the cluster-robust standard error regressions of firm value proxy by Tobin's Q. It is the market value of the firm that is market equity plus total assets minus book equity and divided by total assets as defined in Hou and Robinson (2006). The header of Table 2.2 defines the LatentGrowth that includes FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE. The Controls include Inv_growth, Leverage, ROA, Access_fm, and Size variables that are constructed following Allayannis and Weston (2001). Inv_growth is R&D expenditure to sales. Access_fm is a dummy for access to financial markets, which is equal to 1 if the firm paid a dividend in the current year. Size is a log of total assets that control for the effect of size. LEV is the sum of long-term debt and debt in current liabilities, scaled by total assets as defined by Cooper, et al. (2008). ROA is net income (Compustat item #A178) divided by the total assets (Computstat item #A6). The study period is spread over the 25 years of 1985 to 2009 with 876,348 firm-month observations for a non-financial data sample. The table reports the coefficients and *t*-statistics for seven models across Panels A-E using standard clustered errors along firm (CL-i), along time (CL-T), the intersection of both firm and time (CL-2), CLR-2 which controls for persistent common shocks, and CLRi-2 which controls for industry effect. The CL-2, CLR-2, and CLRi-2 control for cross-sectional and time-series correlation. *, **, *** Denotes significance at the 10%, 5%, and 1% level, respectively. IV refers to the independent variables.

2.4.1.3 Prediction of future cash flows and earnings

This section enhances the understanding of the relationship between corporate fundamentals and the new latent factors. Following striking results indicated in Table 2.4, this study tries to relate these new latent factors with other firm fundamentals such as cash flows (Cashflows) and earnings (Earnings). This study looks at both earnings and cash flows separately because they are not equivalent measures (Sloan, 1996). In Table 2.5 and Table 2.6, this essay provides evidence that the new latent growth factors are robust predictors of current and future cash flows and earnings. Using the similar standard error clustering approaches along firm (CL-i), time (CL-t), the intersection of both time and firm (CL-2), and approaching it without persistent common shocks (CLR-2), this study runs equations of the following form:

$$X_{i,t+1} = \alpha + \beta_1 LatentGrowth_{i,t} + \beta_2 Controls_{i,t} + \varepsilon_{i,t+1} \quad (12)$$

$$X_{i,t} = \alpha + \beta_1 LatentGrowth_{i,t} + \beta_2 Controls_{i,t} + \varepsilon_{i,t} \quad (13)$$

$$X_{i,t} = \alpha + \beta_1 LatentGrowth_{i,t-n} + \beta_2 Controls_{i,t} + \varepsilon_{i,t} \quad (14)$$

The dependent variables $X_{i,t+1}$, and $X_{i,t}$ are Cashflows and Earnings measured in fiscal year $t+1$, and t . Earnings are measured by dividing income from continuing operations (Compustat item #178) by lagged total assets (Compustat item #6). Cashflows are the difference between Earnings and Accruals (ACCR). The Cashflows, Earnings, and ACCR are measured as in Hirshleifer et al. (2004). The independent variable LatentGrowth includes the latent accounting growth factors FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE and Controls that are other firm characteristics. The Controls variable set is similar to that used in the previous section for forecasting firm value and includes Inv_growth, Access_fm, Size, ROA, and LEV. In the models where the dependent

variable is Earnings, this study excludes the ROA control owing to the perfect correlation. However, all control variables are included in the regression models, where Cashflows are a dependent variable. The equation (14) introduces the annual lag in the LatentGrowth variable set, but not in the controls. This essay reports the results up to the three year lag, thus the long-term relationship of LatentGrowth is observed with Cashflows and Earnings.

The results in Table 2.5 indicate the implications of the investment and financing decisions (LatentGrowth). Generally the literature explores the impact of free cash flow on firms' investment decisions. In Modigliani and Miller's (1958) framework firm level investment and financing should not be related to cash flows. However, when this study examines the changes in investment and financing decisions, it encounters several financial frictions, which makes this relationship worth researching. The relationship between investment and cash flows is expected to be positive (Hubbard, 1997) owing to an agency problem and costly external financing. The agency issue indicates the misuse of internally generated cash, when it is in excess of the requirement for exploiting the current investment opportunity set (Jensen, 1986; Stulz, 1990). In contrast, costly external finance compels the firm to use internally generated cash flows and to expand the investment opportunity set (Fazzari et al., 1988; Hubbard, 1997).

Contrary to prior literature, evidence in this study presented in Panels A to E of Table 2.5 shows that overall FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT factors have a significantly negative relation with Cashflows. The PSK_USE financing is the only factor that is consistently positively related with current and future cash flows. This finding has four possible explanations. First, this negative association reflects that the current over-

investment in an unfeasible large investment opportunity set is resulting in poor future performance in generating cash flows from operations. This is consistent with the literature that relates investment activity with poor future stock returns (Titman et al., 2004; Fairfield et al., 2003; Dechow et al., 2005).

Second, intensive earnings management through accruals would attach the larger portion of the earnings with accruals (DeFond and Jiambalvo, 1994; Rangan, 1998; Subramanyam, 1996; Teoh, et al., 1998). The extent of this earnings management is such that earnings attached to accruals are most of the time larger than earnings itself. Then, the resulting cash flows which are the difference in earnings and accruals become negative. When the majority of cash flows are negative, the only possibility for a positive relationship between Cashflows and LatentGrowth to exist is if firms are able to raise *cheap* capital and squander it (Richardson, 2006). However, the second explanation requires this relationship to be positive in the long term; the declining coefficients of LatentGrowth are an indication of a future positive relationship. Third, the under-performance of the investment and financing activities (LatentGrowth) may be attributed to the increasing inclination of managers for real activities management (Gunny, 2005; Roychowdhury, 2006; Zang, 2006) compared to accrual-based earnings management. This essay argues that both real and accruals-based earnings management activities result in negative Cashflows and LatentGrowth relationships. Fourth, negative cash flow would mean external equity or zero dividend in those years or cash investment in value adding projects by liquidating assets that generate negative free cash flow rather positive. To observe the positive cash flows the long-forecast horizon needs to be observed (Penman, 2001; Hou et al., 2014b).

Table 2.6 in Panel A to E reports the results for firm earnings (Earnings) as a dependent variable. The variables of interest LatentGrowth are found as having a weak, but positive association with Earnings. The result is consistent with the findings in Section 2.4.1.2 that LatentGrowth is a value creating activity. Decisions such as FIN_FLEX, ST_CREDIT, LT_INV, and PSK_USE have significantly positive future earnings forecasting power. The CVT_DEBT is the only decision that has a negative relationship with future earnings. Current earnings appear to be unrelated with FIN_FLEX and CVT_DEBT. The evidence for the two-year and three-year lag suggests that the FIN_FLEX's long-term relationship with earnings is reversed, but remains insignificant. However, ST_CREDIT and LT_INV are persistently strong determinants of future earnings. The table 2.5 and 2.6 reveals that the coefficients for FIN_FLEX, ST_CREDIT, and LT_INV change owing to management activities and their interdependence on firm earnings and cash flows and signs of CVT_DEBT and PSK_USE do not change as not governed by earnings and cash flows.

This evidence is consistent with the literature that argues that investment decisions and earnings are interdependent (McNichols and Stubben, 2008). Investment and financing decisions have their roots in positive outcomes (earnings and revenue) that are again dependent on expectations about the future growth opportunities and product demand. This study shows that on aggregate whether or not firms manage earnings and report better earnings for a more positive outlook of the firm to potential users of such information, the basic operating business activity decisions (FIN_FLEX, ST_CREDIT, LT_INV, and PSK_USE) would have a positive impact on overall future earnings for the firm. Overall, the evidence indicates that latent measures of corporate financing and investment are more associated with firms' future cash flows than with earnings.

Table 2.5**Forecasting future firm cash flows**

$$X_{i,t+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t+1}$$

$$X_{i,t} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t}$$

$$X_{i,t} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-n} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t}$$

Panel A		Dependent: Cashflows_{t+1} (n= 710951, Cli=7536, CLt=24, CL2=58425, t=1985-2009)						
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0567	-25.94***	-12.52***	-27.79***	-12.33***	-0.0504	-9.98***
M-2	ST_CREDIT	-0.0584	-20.12***	-17.65***	-21.75***	-16.7467***	-0.0494	-14.66***
M-3	LT_INV	-0.0140	-5.89***	-4.28***	-7.68***	-3.88***	-0.0145	-4.21***
M-4	CVT_DEBT	-0.0150	-10.62***	-7.53***	-10.76***	-7.48***	-0.0137	-7.15***
M-5	PSK_USE	0.0128	6.83***	3.88***	6.40***	3.97***	0.0033	1.25
Panel B		Dependent: Cashflows_t (n= 777518, Cli=8659, CLt=25, CL2=67848, t=1985-2009)						
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0984	-3.89***	-2.30**	-3.89***	-2.30**	-0.0836	-1.85*
M-2	ST_CREDIT	-0.1043	-3.72***	-2.92***	-3.74***	-2.92***	-0.0907	-2.46**
M-3	LT_INV	0.0042	0.30	0.28	0.31	0.28	0.0064	0.41
M-4	CVT_DEBT	-0.0110	-3.47***	-4.73***	-3.44***	-4.82***	-0.0083	-2.52**
M-5	PSK_USE	0.0322	5.16***	2.14**	5.15***	2.14**	0.0175	1.63
Panel C		Dependent: Cashflows_t (n= 722609, Cli=7676, CLt=24, CL2=59281, t=1985-2009)						
Models	IV_{t-1}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0512	-20.79***	-11.11***	-21.94***	21.11***	-0.0456	-8.98***
M-2	ST_CREDIT	-0.0567	-15.62***	-15.75***	-16.15***	-15.26***	-0.0483	-12.31***
M-3	LT_INV	-0.0085	-3.37***	-2.31**	-4.09***	-2.15**	-0.0077	-2.03**
M-4	CVT_DEBT	-0.0129	-9.57***	-7.18***	-9.54***	-7.19***	-0.0120	-6.95***
M-5	PSK_USE	0.0087	4.61***	3.44***	4.34***	3.57***	0.0003	0.14

Table 2.5 Continued

Panel D		Dependent: Cashflows_t (n= 632259, Cli=6815, CLt=23, CL2=51801, t=1985-2009)						
Models	IV_{t-2}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0462	-16.05***	-9.93***	-17.08***	-9.72***	-0.0428	-8.95***
M-2	ST_CREDIT	-0.0463	-16.39***	-12.69***	-17.29***	-12.32***	-0.0400	-10.60***
M-3	LT_INV	-0.0012	-0.48	-0.38	-0.57	-0.35	-0.0011	-0.36
M-4	CVT_DEBT	-0.0127	-8.80***	-7.08***	-8.90***	-7.03***	-0.0122	-7.26***
M-5	PSK_USE	0.0066	3.41***	2.90***	3.16***	3.09***	-0.0011	-0.53
Panel A		Dependent: Cashflows_t (n= 556658, Cli=5983, CLt=22, CL2=45555, t=1985-2009)						
Models	IV_{t-3}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0371	-18.77***	-14.28***	-21.19***	-13.46***	-0.0339	-12.02***
M-2	ST_CREDIT	-0.0433	-16.55***	-16.05***	-17.46***	-15.33***	-0.0388	-13.59***
M-3	LT_INV	0.0026	1.04	0.64	1.27	0.60	0.0024	0.58
M-4	CVT_DEBT	-0.0111	-7.20***	-5.32***	-7.15***	-5.34***	-0.0103	-5.45***
M-5	PSK_USE	0.0043	2.35**	2.26**	2.25**	2.35**	-0.0014	-0.71

The table reports the cluster-robust standard error regressions of firm cash flows $X_{i,t+1}$ proxy by Cashflows. Cashflows are a difference of Earnings and Accruals (ACCR) as defined in Hirshliefer et al., (2004) (see appendix A). The header of Table 2.2 defines the LatentGrowth that include FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE. The Controls include Inv_growth, Leverage, ROA, Access_fm, and Size variables that are constructed following Allayannis and Weston (2001). Inv_growth is R&D expenditure to sales. Access_fm is a dummy for access to financial markets, which is equal to 1 if the firm paid a dividend in the current year. Size is a log of total assets that control for the effect of size. LEV is the sum of long-term debt and debt in current liabilities, scaled by total assets as defined by Cooper, et al., (2008). ROA is net income (Compustat item #A178) divided by the total assets (Computstat item #A6). The study period is spread over 25 years that is 1985 to 2009, with 876,348 firm-month observations for a non-financial data sample. The table reports the coefficients and t -statistics for seven models across Panels A-E using standard clustered errors along firm (CL-i), along time (CL-T), the intersection of both firm and time (CL-2), and CLR-2 which controls for persistent common shocks. The CL-2, CLR-2, and CLRi-2 control for cross-sectional and time-series correlations. *, **, *** Denotes significance at the 10%, 5%, and 1% level, respectively. IV refers to the independent variables.

Table 2.6

Forecasting future firm earnings

$$X_{i,t+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t+1}$$

$$X_{i,t} = \alpha + \beta_1 \text{LatentGrowth}_{i,t} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t}$$

$$X_{i,t} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-n} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{i,t}$$

Panel A		Dependent: Earnings_{t+1} (n= 710951, Cli=7536, CLt=24, CL2=58425, t=1985-2009)						
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	0.0194	10.08***	2.90***	10.3***	2.90***	0.0200	3.13***
M-2	ST_CREDIT	0.0133	8.19***	5.82***	8.93***	5.60***	0.0097	4.06***
M-3	LT_INV	0.0056	6.11***	3.67***	6.54***	3.58***	0.0059	3.57***
M-4	CVT_DEBT	-0.0024	-2.88***	-2.78***	-2.84***	-2.81***	-0.0030	-3.43***
M-5	PSK_USE	0.0027	2.01**	0.82	1.97**	0.83	0.0062	2.19**
Panel B		Dependent: Earnings_t (n= 777518, Cli=8659, CLt=25, CL2=67848, t=1985-2009)						
Models	IV_t	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	0.0154	0.91	0.31	0.92	0.31	0.0176	0.35
M-2	ST_CREDIT	0.0372	2.39**	2.39**	2.40**	1.87*	0.0333	1.50
M-3	LT_INV	0.0207	3.24***	1.96**	3.26***	1.95*	0.0198	1.83*
M-4	CVT_DEBT	-0.0002	-0.07	-0.09	-0.07	-0.09	-0.0006	-0.20
M-5	PSK_USE	0.0261	4.81***	1.24	4.80***	1.24	0.0296	1.77*
Panel C		Dependent: Earnings_t (n= 722609, Cli=7676, CLt=24, CL2=59281, t=1985-2009)						
Models	IV_{t-1}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	0.0112	4.64***	1.35	4.72***	1.35	0.0113	1.42
M-2	ST_CREDIT	0.0115	6.57***	5.13***	7.04***	4.94***	0.0093	3.68***
M-3	LT_INV	0.0041	3.98***	2.27**	4.23***	2.23**	0.0040	2.15**
M-4	CVT_DEBT	-0.0015	-1.64	-1.42	-1.59	-1.46	-0.0017	-1.76*
M-5	PSK_USE	0.0031	1.92*	0.86	1.89*	0.86	0.0052	1.74*

Table 2.6 (Continued)

Panel D		Dependent: Earnings_t (n= 632259, Cli=6815, CLt=23, CL2=51801, t=1985-2009)						
Models	IV_{t-2}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0013	-0.65	-0.29	-0.66	-0.29	-0.0014	-0.33
M-2	ST_CREDIT	0.0073	4.48***	3.24***	4.64***	3.19***	0.0075	3.49***
M-3	LT_INV	0.0019	2.06**	1.2	2.24**	1.17	0.0016	0.98
M-4	CVT_DEBT	-0.0011	-1.35	-1.22	-1.31	-1.26	-0.0011	-1.22
M-5	PSK_USE	0.0027	1.86*	1.11	1.88*	1.11	0.0026	1.26
Panel E		Dependent: Earnings_t (n= 556658, Cli=5983, CLt=22, CL2=45555, t=1985-2009)						
Models	IV_{t-3}	Coeff.	CL-i (t-stat)	CL-t (t-stat)	CL-2 (t-stat)	CLR-2 (t-stat)	Coeff.	CLR-2 (t-stat) M-6
M-1	FIN_FLEX	-0.0030	-1.80*	-0.88	-1.85*	-0.87	-0.0029	-0.87
M-2	ST_CREDIT	0.0060	3.67***	3.39***	3.65***	3.40***	0.0063	3.71***
M-3	LT_INV	0.0032	3.42***	2.21**	3.73***	2.14***	0.0029	1.89*
M-4	CVT_DEBT	-0.0009	-1.01	-0.8	-1.01	-0.81	-0.0008	-0.78
M-5	PSK_USE	0.0027	1.80*	1.74*	1.85*	1.70*	0.0025	1.66*

The table reports the cluster-robust standard error regressions of firm earnings $X_{i,t+1}$ proxy by Earnings. Earnings are measured by dividing the Income from continuing operations (Compustat item #178) by lagged total assets as defined in Hirshliefer et al., (2004) (see appendix A). The header of Table 2.2 defines the LatentGrowth that includes FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE. The Controls include Inv_growth, Leverage, Access_fm, and Size variables that are constructed following Allayannis and Weston (2001). Inv_growth is R&D expenditure to sales. Access_fm is a dummy for access to financial markets, which is equal to 1 if the firm paid a dividend in the current year. Size is a log of total assets that control for the effect of size. LEV is the sum of long-term debt and debt in current liabilities, scaled by total assets as defined by Cooper, et al., (2008). The study period is spread over 25 years that is 1985 to 2009 with 876,348 firm-month observations for a non-financial data sample. The table reports the coefficients and t -statistics for seven models across Panels A-E using standard clustered errors along firm (CL-i), along time (CL-T), the intersection of both firm and time (CL-2), and CLR-2 which controls for persistent common shocks. The CL-2, CLR-2, and CLRi-2 control for cross-sectional and time-series correlations. *, **, *** Denotes significance at the 10%, 5%, and 1% level, respectively. IV refers to the independent variables.

2.4.2 Forecasting future stock returns

In this section, the study attempts to illustrate the effect of the new latent growth factors on future stock returns in excess of the risk-free rate. Since section 2.4.1.1 show that latent growth is consistent with the Q theory of investment, similarly this theory links these factors and stock returns. This study borrows an economic model from Hou et al. (2014a) – $E_0 [r_{i1}^S] = \frac{E_0[\Pi_{i1}]}{1+a(I_{i0}/A_{i0})}$, where, r_{i1}^S are future stock returns, A_{i0} and I_{i0} are assets and investment at time 0, and Π_{i1} is a profitability at time 1 – and just focus on investment channel to show that the latent growth variables have negative relation with the subsequent stock returns. The literature in this area documents the negative future return and investment and financing relation (Berk et al., 1999; Titman et al., 2004; Fairfield et al., 2003; Hirshliefer et al., 2004; Cooper et al., 2008; Lyandres et al., 2008; Kogan and Papanikolaou, 2013).

In a stock level analysis, this study employs a Fama-MacBeth (1973) (FM) for addressing cross-sectional correlations; a Fama-MacBeth-Newey-West (FM-NW) to address both cross-sectional and serial correlations; and OLS standard error clustering approaches for addressing firm (CL-i), time (CL-t) or both firm and time effects (CL-2) along with any persistent common shocks (CLR-2). In addition, this study also takes into account the industry effect (CLRi-2) associated with any of the new latent growth factors. In addition the study uses the WLS (weighted least squares) standard error clustering for taking into account the cluster size difference or any residual correlation. The results of the WLS standard error clustering tests are not reported for the sake of brevity. The data sample is spread over a period of 25 years (July 1986 to June 2011). Table 2.7 reports the

coefficients and t -statistics for the eight models for all the above methods. The study employs regression models of the following form:

$$ERET_{i,t+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \beta_2 Controls_{i,t-1} + \varepsilon_{i,t+1} \quad (15)$$

LatentGrowth in equation (15) refers to the set of new latent factors (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE) measured at the end of fiscal year $t-1$. The dependent return variable is monthly future stock returns in excess of risk-free return that starts from July of year $t+1$ and continues to June of $t+1$. This study skips the year t for accounting conservatism leaving a sufficient time-window for data availability. The new latent factors are created from our systematic search method (factor analysis) that has the potential to synthesise all the elements relative to a fundamental mechanism. For example, a company's investment policy is usually separated into long-term considerations and short-term considerations. Even within the long-term consideration type, the result could be shown in several accounting variables. This essay argues that the latent factors created by factor analysis have this characteristic, and they are better represented than in the simple asset growth measure used in the previous literature. Additionally, factor analysis processes employ the correlational structure of the potential input information that coincides with Subrahmanyam's (2010) argument that researchers need to employ correlation structures amongst the variables that are candidates for predicting stock returns. Based on the above notions, this study measures these latent accounting factors and conjectures that these factors contain information about future average stock returns.

Following the standard literature on accounting growth and the cross-section of expected returns, the set of control variables are selected to include in the regression model in equation (4). Following Fama and French (1992) and Li and Sullivan (2011), these determinants include log of firm size (LSIZE), the log of book-to-market ratio (LBTM), the last six months (BHRET6) and the last three years' (BHRET36) price momentum, where the most recent month is skipped as a control for microstructure biases. ASSETG is a measure of total asset growth, ACCR is accounting accruals, NOA cumulative accruals, and ROA is the measure of profitability, and I_A is the measure for investment to assets. The detailed definitions of the control variables are given in appendix A.

For the full sample, in Table 2.7, the first model (M-1) indicates that all latent growth factors except PSK_USE have significant negative associations with future stock returns after controlling for size, and book-to-market. In the remaining seven models, this study keeps the size and book-to-market ratio fixed, and keeps adding other well-known return determinants. The coefficients (*t*-statistics) of FIN_FLEX, ST_CREDIT, LT_INV and CVT_DEBT factors in M-1 for CLRi-2 approach are -0.0029 (-3.60), -0.0012 (-2.11), -0.0027 (-4.53), and -0.0011 (-2.88) respectively. Models (M-2) to (M-8) indicate that the regression coefficients of the latent factors remain significant when model (M-1) is augmented by ASSETG, NOA, ACCR, I_A, ROA, BHRET6, and BHRET36.³⁶

The empirical results support the argument of this study that the new factors better represent a range of accounting variables than does the simple asset growth measure

³⁶ In the analysis after accounting for the effects captured by all the control variables, the residuals of some of the controls do not possess sufficient information about future stock returns (see appendix A and Tables A.3.2 and A.3.3).

(ASSETG) used in the previous literature. In other words, these factors better classify and specify the investment activities than the aggregates like ASSETG, NOA, ACCR, and I_A. The new factors are able to either subsume the average returns-related effects of ASSETG, NOA, ACCR, ROA, and I_A or have reduced their significance considerably; however, the opposite is not true. This evidence supports the motivation of this study that the accounting information is not being employed adequately. In fact, the pre-determined way of constructing accounting measures is unable to account for covariances among the accounting items arising from the double-entry bookkeeping system. Thus, on the one hand, these constructs miss out on some of the related information unintentionally, and on the other hand, unrelated items are being combined to represent some financial activities. The latter means that these items do not share any variance. Overall, the existing accounting measures tested in this section for their return predictability lack the potential to fully reflect the specific aspects of operating business activities; instead they are aggregates and suboptimal and they lead to the problems of limited attention. This study provides a new set of factors for measuring the time-varying operating business, which would reduce the problem of limited attention and better predict the future excess stock returns³⁷. In other words, this study eliminates the measurement errors in constructing the accounting variables.

³⁷ In capitalisation groups based cross-sectional regressions (see appendix A, Table A.3.1), the latent factors maintain a negative association with the future stock returns except PSK_USE.

Table 2.7

Relations between latent growth factors and the subsequent monthly excess stock returns

$$ERET_{i,t+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \beta_2 Controls_{i,t-1} + \varepsilon_{i,t+1}$$

Dependent: Excess Returns(ERET _{t+1}) (n= 876348, Cli=9822, CLt=25, CL2=76706, t=1985-2009)											
ERET _{i,t+1}	IV _{i,t-1}	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-1	FIN_FLEX	-0.0026	-4.03***	-3.81***	-0.0027	-8.43***	-3.08***	-8.70***	-3.07***	-0.0029	-3.60***
	ST_CREDIT	-0.0020	-2.80***	-2.77***	-0.0013	-3.13***	-2.22**	-3.16***	-2.20***	-0.0012	-2.11**
	LT_INV	-0.0024	-4.59***	-4.18***	-0.0027	-10.87***	-3.93***	-11.01***	-3.93***	-0.0027	-4.53***
	CVT_DEBT	-0.0009	-3.20***	-3.07***	-0.0010	-4.42***	-2.71***	-4.54***	-2.68***	-0.0011	-2.883***
	PSK_USE	0.0001	0.32	0.32	-0.0001	-0.47	-0.30	-0.48	-0.29	-0.00003	-0.06
	LSIZE	-0.0024	-3.84***	-3.69***	-0.0018	-17.4***	-2.40**	-17.71***	-2.40**	-0.0016	-2.29**
	LBTM	0.0046	1.95*	1.82*	0.0106	14.26***	2.33**	14.33***	2.33**	0.0125	3.15***
ERET _{i,t+1}	IV _{i,t-1}	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-2	FIN_FLEX	-0.0029	-2.52**	-2.72***	-0.0026	-7.99***	-3.10***	-8.25***	-3.08***	-0.0028	-3.65***
	ST_CREDIT	-0.0021	-2.82***	-2.90***	-0.0013	-3.06***	-2.18**	-3.10***	-2.17**	-0.0012	-2.07**
	LT_INV	-0.0026	-3.57***	-3.55***	-0.0027	-10.64***	-3.92***	-10.77***	-3.91***	-0.0027	-4.54***
	CVT_DEBT	-0.0009	-3.21***	-3.05***	-0.0010	-4.38***	-2.69***	-4.50***	-2.67***	-0.0011	-2.87***
	PSK_USE	0.0001	0.26	0.26	-0.0001	-0.42	-0.26	-0.42	-0.26	-0.00001	-0.02
	ASSETG	0.0004	0.49	0.58	-0.00005	-1.55	-1.56	-1.55	-1.56	-0.00005	-1.54
	LBTM	0.0046	1.93*	1.80*	0.0106	14.27***	2.33**	14.35***	2.33**	0.0125	3.15***
ERET _{i,t+1}	IV _{i,t-1}	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-3	FIN_FLEX	-0.0019	-2.89***	-2.68***	-0.0025	-7.81***	-2.88***	-8.05***	-2.87***	-0.0028	-3.42***
	ST_CREDIT	-0.0014	-1.94*	-1.89*	-0.0012	-2.94***	-2.13**	-2.98***	-2.12**	-0.0012	-2.03**
	LT_INV	-0.0013	-2.11**	-1.99*	-0.0026	-9.61***	-3.73***	-9.72***	-3.72***	-0.0025	-4.28***
	CVT_DEBT	-0.0008	-2.70***	-2.56**	-0.0010	-4.36***	-2.67***	-4.49***	-2.64***	-0.0011	-2.85***
	PSK_USE	0.0003	0.77	0.76	-0.0001	-0.27	-0.17	-0.28	-0.17	0.00003	0.06
	NOA	-0.0030	-2.56**	-2.43**	-0.0003	-1.67	-2.04**	-1.67	-2.04**	-0.00030	-2.20**
	LSIZE	-0.0023	-3.82***	-3.67***	-0.0018	-17.4***	-2.40**	-17.7***	-2.40**	-0.0016	-2.29**
	LBTM	0.0052	2.30**	2.14***	0.0107	14.33***	2.35**	14.42***	2.35**	0.0126	3.17***

Table 2.7 (Continued)

$ERET_{i,t+1}$	$IV_{i,t-1}$	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-4	FIN_FLEX	-0.0025	-3.99***	-3.74***	-0.0027	-8.43***	-3.08***	-8.71***	-3.07***	-0.0029	-3.60***
	ST_CREDIT	-0.0019	-2.69***	-2.70***	-0.0013	-3.13***	-2.20**	-3.17***	-2.19**	-0.0012	-2.10**
	LT_INV	-0.0024	-4.61***	-4.20***	-0.0027	-10.87***	-3.93***	-11.01***	-3.93***	-0.0027	-4.53***
	CVT_DEBT	-0.0009	-3.21***	-3.07***	-0.0010	-4.42***	-2.71***	-4.54***	-2.68***	-0.0011	-2.88***
	PSK_USE	0.0001	0.26	0.26	-0.0001	-0.47	-0.3	-0.48	-0.29	-0.00003	-0.06
	ACCRUALS	-0.0026	-2.14**	-1.98**	0.0000	0.35	0.21	0.35	0.21	0.00003	0.21
	LSIZE	-0.0024	-3.85***	-3.69***	-0.0018	-17.4***	-2.40**	-17.71***	-2.40**	-0.0016	-2.29**
	LBTM	0.0044	1.88*	1.76*	0.0106	14.26***	2.33**	14.34***	2.33**	0.0125	3.15***
$ERET_{i,t+1}$	$IV_{i,t-1}$	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-5	FIN_FLEX	-0.0025	-3.84***	-3.47***	-0.0026	-8.12***	-3.02***	-8.39***	-3.01***	-0.0029	-3.56***
	ST_CREDIT	-0.0018	-2.44**	-2.36**	-0.0012	-3.01***	-2.15**	-3.03***	-2.14**	-0.0012	-2.05**
	LT_INV	-0.0020	-3.24***	-3.09***	-0.0027	-9.79***	-3.67***	-9.90***	-3.66***	-0.0026	-4.18***
	CVT_DEBT	-0.0009	-3.28***	-3.11***	-0.0010	-4.42***	-2.71***	-4.55***	-2.69***	-0.0011	-2.89***
	PSK_USE	0.0001	0.36	0.37	-0.0001	-0.44	-0.28	-0.45	-0.28	-0.00002	-0.04
	I_A	-0.0020	-1.06	-1.12	-0.0002	-0.69	-0.69	-0.69	-0.69	-0.00018	-0.54
	LSIZE	-0.0024	-3.84***	-3.69***	-0.0018	-17.41***	-2.40**	-17.72***	-2.40**	-0.0016	-2.29**
	LBTM	0.0046	1.95*	1.82*	0.0106	14.26***	2.33**	14.34***	2.33**	0.0125	3.15***
$ERET_{i,t+1}$	$IV_{i,t-1}$	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-6	FIN_FLEX	-0.0022	-3.67***	-3.44***	-0.0025	-8.03***	-2.61***	-8.25***	-2.60***	-0.0028	-3.05***
	ST_CREDIT	-0.0021	-2.91***	-2.90***	-0.0011	-2.78***	-2.16**	-2.81***	-2.15**	-0.0011	-2.08**
	LT_INV	-0.0023	-4.48***	-4.09***	-0.0026	-10.34***	-3.95***	-10.49***	-3.94***	-0.0026	-4.54***
	CVT_DEBT	-0.0009	-3.42***	-3.29***	-0.0011	-4.72***	-2.98***	-4.86***	-2.95***	-0.0011	-3.05***
	PSK_USE	0.0001	0.15	0.15	0.0000	-0.01	-0.01	-0.01	-0.01	0.00010	0.16
	ROA	-0.0091	-2.10**	-1.92*	-0.0138	-10.64***	-1.61	-9.77***	-1.62	-0.01150	-1.48
	LSIZE	-0.0020	-3.93***	-3.89***	-0.0013	-13.16***	-1.95*	-13.28***	-1.95*	-0.0013	-1.87*
	LBTM	0.0056	2.80***	2.61***	0.0118	15.96***	2.86***	16.21***	2.86***	0.0132	3.49***

Table 2.7 (Continued)

Dependent: Excess Returns(t+1) (n= 746544, Cli=8279, CLt=25, CL2=67589, t=1985-2009)											
ERET_{i,t+1}	IV_{i,t-1}	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-7	FIN_FLEX	-0.0045	-4.69***	-4.57***	-0.0040	-7.16***	-3.42***	-7.39***	-3.39***	-0.0044	-3.69***
	ST_CREDIT	-0.0021	-2.57	-2.48	-0.0018	-3.51***	-2.78***	-3.51***	-2.78***	-0.0017	-2.66***
	LT_INV	-0.0026	-4.72***	-4.39***	-0.0026	-8.76***	-3.04***	-8.73***	-3.04***	-0.0026	-3.46***
	CVT_DEBT	-0.0007	-2.38**	-2.26**	-0.0009	-3.37***	-2.05**	-3.42***	-2.04**	-0.0009	-2.20**
	PSK_USE	0.0003	0.40	0.41	-0.0003	-0.65	-0.46	-0.65	-0.46	-0.00030	-0.43
	LSIZE	-0.0023	-3.98***	-3.75***	-0.0018	-16.04***	-2.31**	-16.19***	-2.31**	-0.00150	-2.23**
	LBTM	0.0030	1.31	1.18	0.0101	12.07***	2.03**	12.22***	2.03***	0.0123	2.91***
	BHRET6	-0.0051	-1.54	-1.48	-0.0088	-13.37***	-2.28**	-13.44***	-2.28**	-0.0091	-2.40***
ERET_{i,t+1}	IV_{i,t-1}	Coeff.	FM	FM-NW1	Coeff.	CL-i t-stat	CL-t t-stat	CL-2 t-stat	CLR-2 t-stat	Coeff.	CLRi-2 t-stat
M-8	FIN_FLEX	-0.0034	-3.54***	-3.60***	-0.0026	-4.74***	-2.20**	-4.90***	-2.18**	-0.0030	-2.52**
	ST_CREDIT	-0.0020	-2.38**	-2.30**	-0.0015	-2.98***	-2.28**	-2.98***	-2.28**	-0.0014	-2.19**
	LT_INV	-0.0024	-4.38***	-4.00***	-0.0022	-7.48***	-2.66***	-7.54***	-2.65***	-0.0022	-3.03***
	CVT_DEBT	-0.0006	-2.29**	-2.19**	-0.0009	-3.42***	-2.10**	-3.48***	-2.09**	-0.0009	-2.23**
	PSK_USE	0.0002	0.24	0.25	-0.0004	-0.77	-0.57	-0.77	-0.57	-0.00040	-0.54
	LSIZE	-0.0021	-3.51***	-3.31***	-0.0016	-15.4***	-2.22**	-15.53***	-2.22**	-0.00140	-2.13**
	LBTM	0.0018	0.77	0.71	0.0073	9.00***	1.55	8.96***	1.55	0.0094	2.32**
	BHRET36	-0.0030	-2.85***	-2.46**	-0.0021	-10.02***	-3.03***	-11.24***	-3.00***	-0.0021	-3.01***

For a sample period of July, 1986 to December, 2011, this study ran Fama and MacBeth's (1973) cross-sectional regressions (FM), Fama-MacBeth with Newey-West (FM-NW1) with a lag length of 1, and standard clustered errors regressions along firm (CL-i), along time (CL-T), the intersection of both firm and time (CL-2), CLR-2 which controls for persistent common shocks, and CLRi-2 controls for industry effect. The CL-2, CLR-2, and CLRi-2 control for cross-sectional and time-series correlation. The regressions are run for the individual monthly future stock returns in excess of risk-free rate ($ERET_{i,t+1}$) on the LatentGrowth factors and the control characteristics ($IV_{i,t-1}$). The header of Table 2 defines both the LatentGrowth (include FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE) and the Controls (include LBTM, LSIZE, BHRET6, BHRET36, ASSETG, NOA, ACCR, I_A, and ROA). The table shows average slope coefficients with significance (***) for 1%, ** for 5% and * for 10%) based on t -statistics. See appendix A for the variable definition.

2.4.3 Portfolio analysis

2.4.3.1 Time trends of new factor ranked extreme decile

Following Cooper et al. (2008), reported in figures 2.2 to 2.6 are the long-run effects of earnings, firm value, and average stock return for two extreme decile portfolios. Hirshliefer et al. (2004) present similar evidence to show the limited attention problem relating the cumulative accruals. This analysis implying that investor showing limited attention in learning from pre-event period trends of firm earnings, value, and stock returns. However, this research shows that such an effect is lessen in post-event for latent five factors as compared to other well-known investment based factors. The reduction in limited attention issue can be attributed to the removal of missing variable bias and estimation errors with application of factor analysis.

The study first sorts and ranks the data sample on the lagged latent growth factors (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE) into decile growth portfolios every t year at the end of June 1985 to June 2009. Using the June t growth cutoffs, the portfolios are formed that are held for one year from July of year t to June of $t+1$ and then rebalanced. This figure plots the mean monthly earnings, mean values for Q, and mean monthly raw stock returns on both an equal(value)-weighted basis for the two extreme decile portfolios at an event time (five-years pre and post portfolio formation). The earnings are measured by dividing the income from continuing operations (Compustat item #178) by lagged total assets. Tobin's Q is measured as the market equity plus total assets minus book equity and divided by total assets as defined in Hou and Robinson (2006). The returns are 12 months accumulated monthly raw returns. The earnings, Q, and the returns are estimated starting four months after the fiscal year end.

In Figure 2.2 (a) for the two extreme FIN_FLEX deciles (deciles 1 and 10), the earnings of firms in the low financial flexibility decile are always lower and negative than the firms in the high financial flexibility decile in the five years pre and post ranking period. On average in Figure 2.2 (c), the value of Q is positive for both extreme deciles; the mean Q for the low decile is always lower than that for the high decile in the five years pre and post ranking periods. The values for Q remain stable across the event periods for the low decile but for the high decile it starts declining after year -3. On aggregate, investment in both extreme deciles has value creation implications. Then, from Figure 2.2 (e), the high FIN_FLEX decile firms appear to earn higher mean monthly returns in years -5 to -1. We see a crossover after the ranking period when high FIN_FLEX earns lower mean monthly returns than the low FIN_FLEX decile. The trend shows that although firms in the high decile have higher earnings and firm value in a post-ranking period than in prior years, they are still unable to sustain abnormal returns. Alternatively, it could be said that firms that have the most negative financial flexibility shifts and negative profitability produce a significant mean monthly return for at least the next three years compared to firms that are large sized with positive earnings and that show the most positive change in financial flexibility. A similar trend is observed for FIN_FLEX, shown in Figure 2.2 (b, d, & f) on a value-weighted basis for mean earnings and Tobin's Q but mean returns.

On an equal-weighted basis for ST_CREDIT (see Figure 2.3 (a)), firm earnings for both extreme deciles start declining from year -5 and become negative in year -1. After the ranking period the earnings of both deciles jump to positive figures in five post ranking years. On average the high decile possesses higher earnings as compared to the low decile. The average values of Q are positive for both deciles as shown in Figure 2.3 (c), but differ

marginally. However, on a value-weighted basis the trend is the same (see Figure 2.3 (d)), but the spread becomes significant. In terms of the capability to generate mean monthly returns as shown in Figure 2.3 (e), firms in the high ST_CREDIT decile earn higher returns during five prior event years compared to the post-ranking years. However, in post-ranking years the return spread is marginal and mixed. The investors are not sure what available information actually implies. The return spread between high (low) ST_CREDIT decile is corrected in the post-event period.

Figure 2.4 (a) shows that firms in the high LT_INV decile have positive earnings as compared to the earnings of low LT_INV decile firms. Firms in the low LT_INV decile have negative earnings in years -4 to 4. The mean Q in Figure 2.4 (c) is positive across the 10 years, but in the prior ranking period both deciles have similar values for Q. In the case of returns, in Figure 2.4 (e) there is a crossover in year -1, the firms in the high LT_INV decile earn higher returns prior to -1 and lower returns in later years, until they converge in year 4. It appears that despite high earnings and firm value levels, firms in the high LT_INV decile are unable to sustain the generation of higher returns in post-event years.

The evidence for LT_INV is quite consistent with the evidence from NOA (net operating assets) (Hirshleifer, et al., 2004). On a value-weighted basis the trend of both high and low deciles for earnings and firm value cannot be differentiated after the ranking year, as shown in Figure 2.4 (b & d), but there is significant variation in prior years. However, as shown in Figure 2.4 (f), the return effects remain the same. The ability of the investors (or potential users) to receive and interpret a firm's fundamentals and latent decisions appears to be limited.

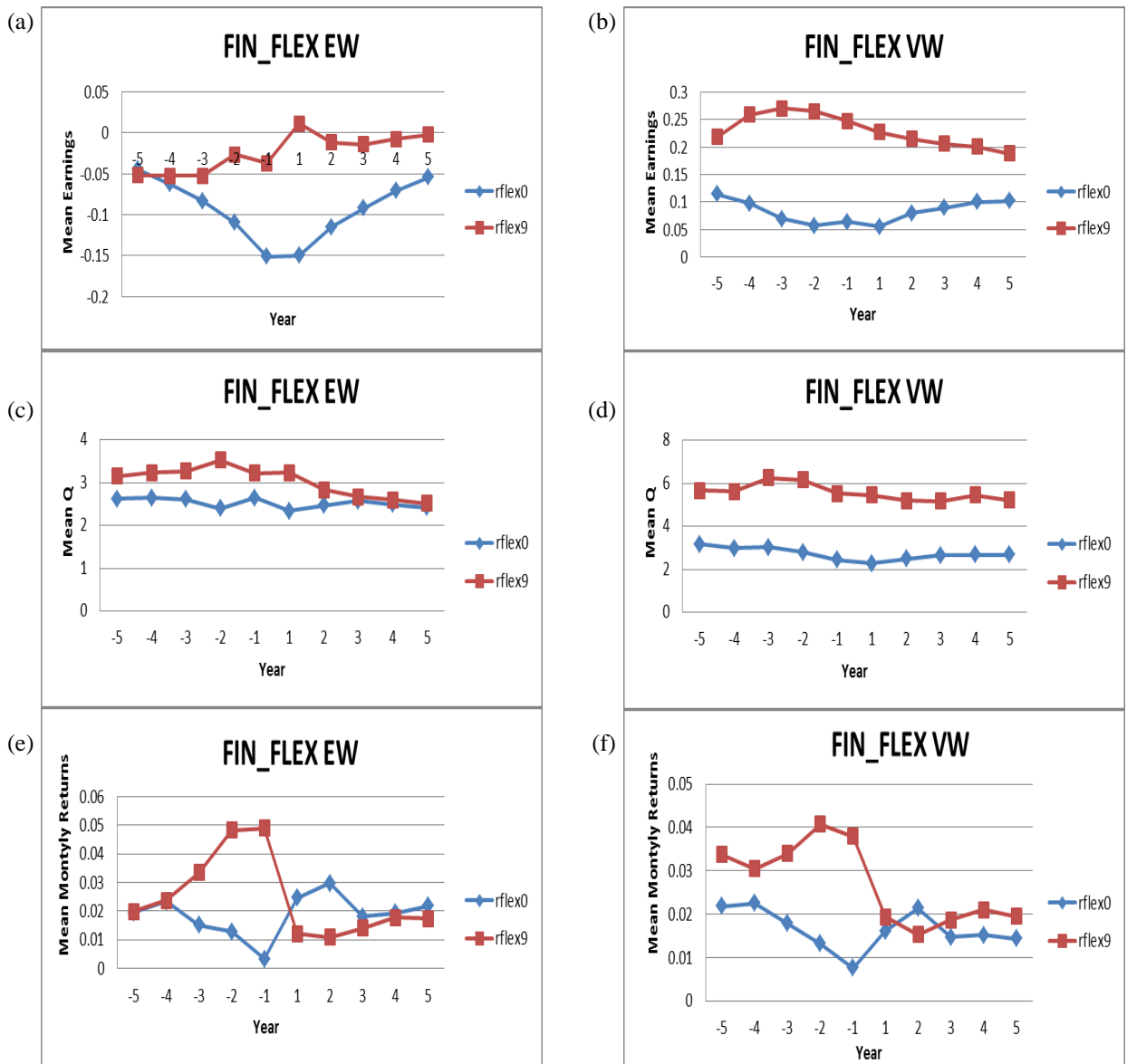


Figure 2.2 (a-f): Time trends of FIN_FLEX ranked extreme decile

The figures show the mean monthly returns, mean earnings, and mean Tobin's Q for FIN_FLEX sorted portfolios five years prior to the ranking year and five years post ranking period. The raw returns are accumulated monthly starting at one year after fiscal year end. First the data sample is sorted and ranked on the lagged FIN_FLEX rate into decile portfolios every t year at the end of June from 1985 to 2009. Using the June t FIN_FLEX growth cutoffs, portfolios are formed that are held for one year from July of t year to June of $t+1$ and then rebalanced.

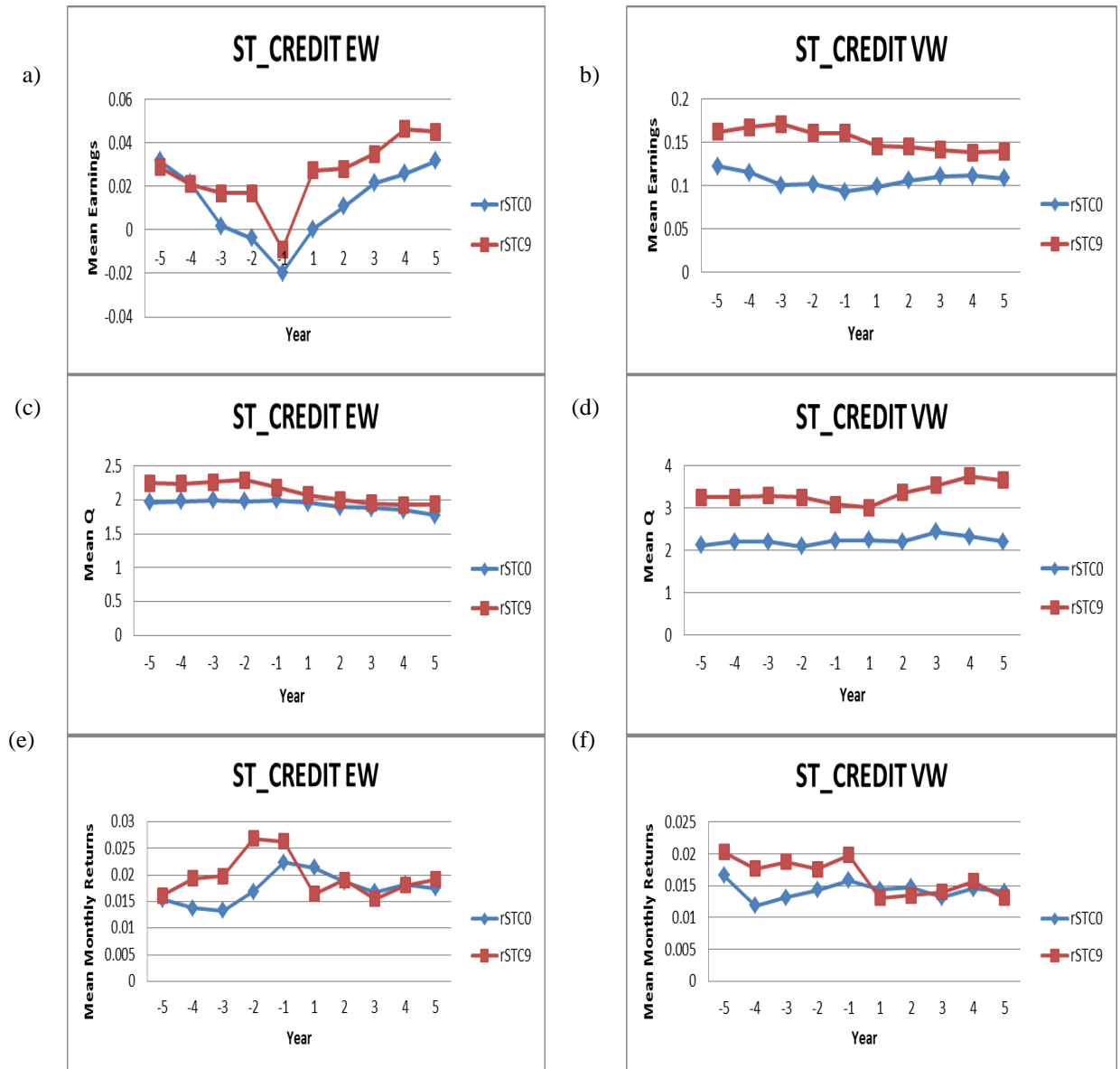


Figure 2.3 (a-f): Time trends of ST_CREDIT ranked extreme decile

The figures show the mean monthly returns, mean earnings, and mean Tobin's Q for ST_CREDIT sorted portfolios five years prior to the ranking year and five years post ranking period. The raw returns are accumulated monthly starting at one year after fiscal year end. First the data sample is sorted and ranked on the lagged ST_CREDIT rate into decile portfolios every t year at the end of June from 1985 to 2009. Using the June t ST_CREDIT growth cutoffs, portfolios are formed that are held for one year from July of t year to June of $t+1$ and then rebalanced.

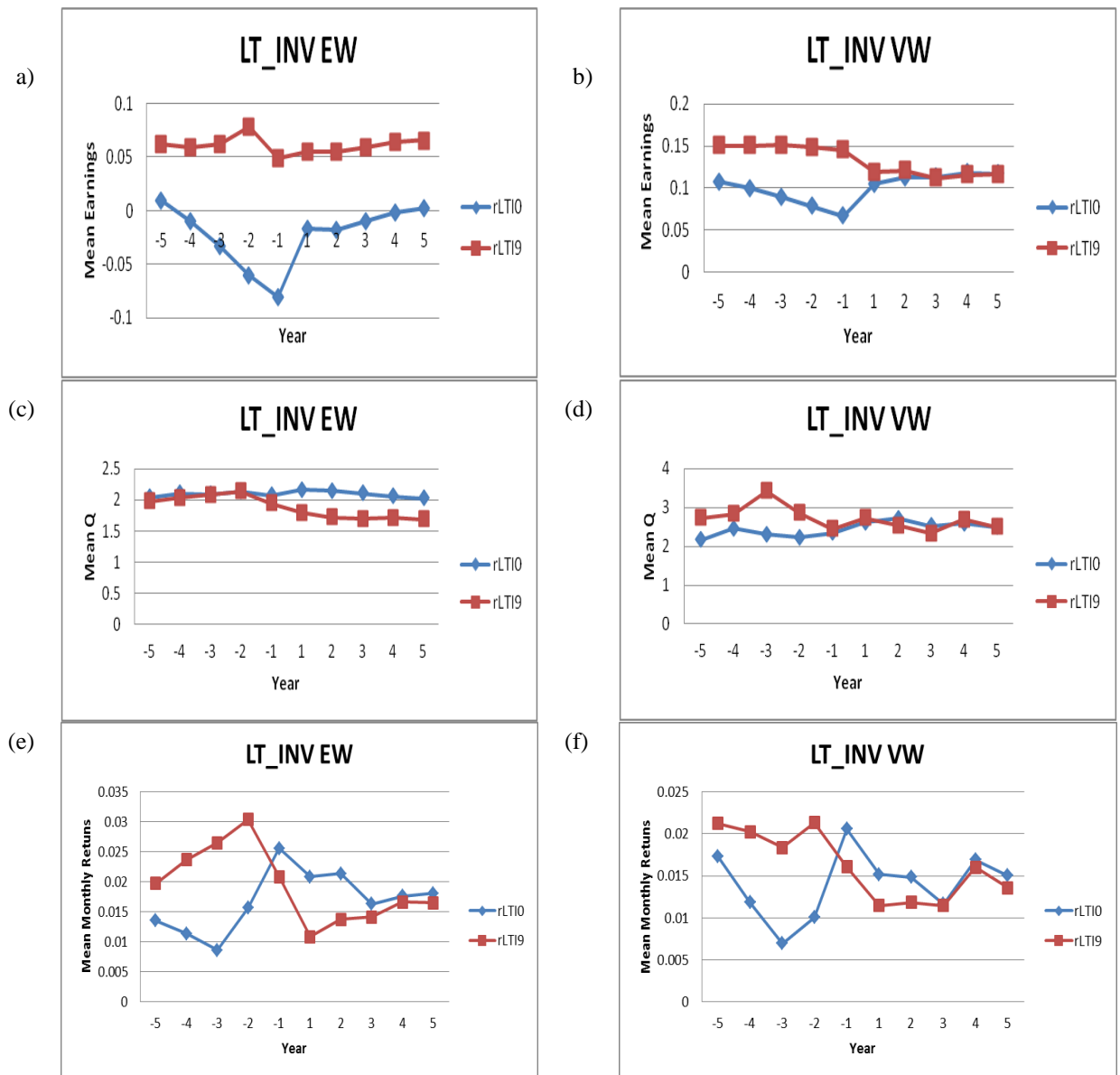


Figure 2.4 (a-f): Time trends of LT_INV ranked extreme decile

The figures show the mean monthly returns, mean earnings, and mean Tobin's Q for LT_INV sorted portfolios five years prior to the ranking year and five years post ranking period. The raw returns are accumulated monthly starting at one year after fiscal year end. First the data sample is sorted and ranked on the lagged LT_INV rate into decile portfolios every t year at the end of June from 1985 to 2009. Using the June t LT_INV growth cutoffs, portfolios are formed that are held for one year from July of t year to June of $t+1$ and then rebalanced.

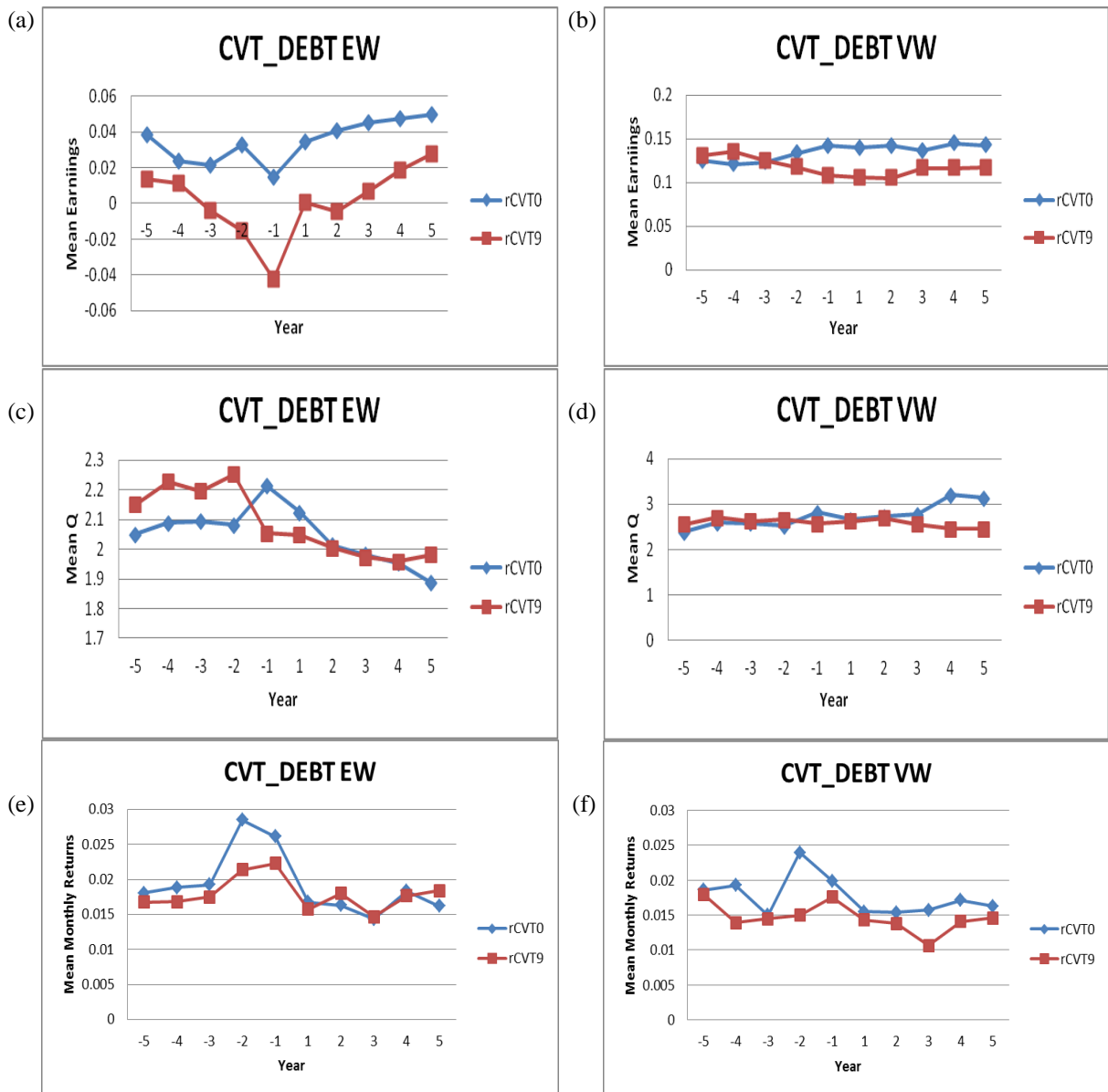


Figure 2.5 (a-f): Time trends of CVT_DEBT ranked extreme decile

The figures show the mean monthly returns, mean earnings, and mean Tobin’s Q for CVT_DEBT sorted portfolios five years prior to the ranking year and five years post ranking period. The raw returns are accumulated monthly starting at one year after fiscal year end. First the data sample is sorted and ranked on the lagged CVT_DEBT rate into decile portfolios every t year at the end of June from 1985 to 2009. Using the June t CVT_DEBT growth cutoffs, portfolios are formed that are held for one year from July of t year to June of $t+1$ and then rebalanced.

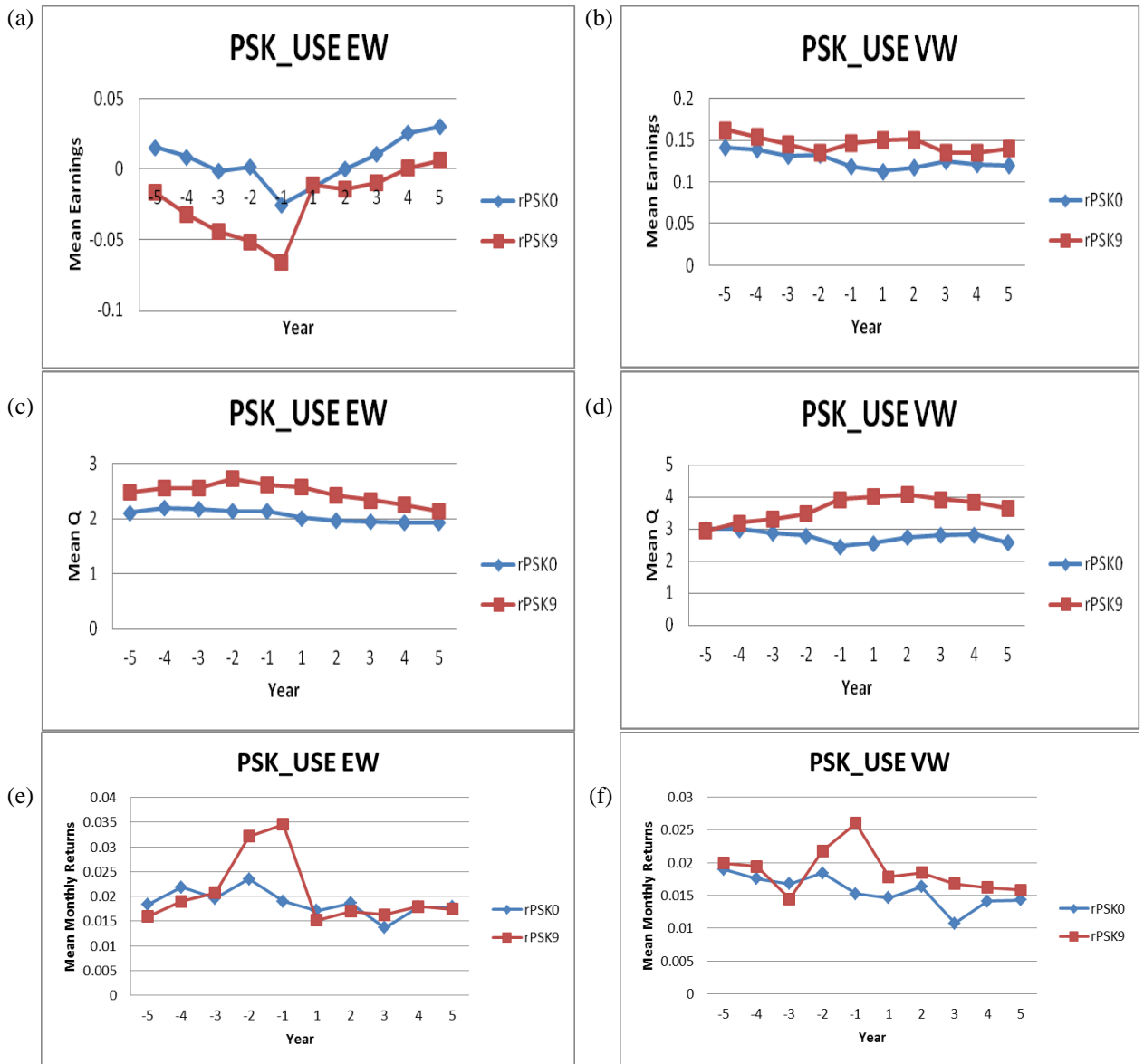


Figure 2.6 (a-f): Time trends of PSK_USE ranked extreme decile

The figures show the mean monthly returns, mean earnings, and mean Tobin’s Q for PSK_USE sorted portfolios five years prior to the ranking year and five years post ranking period. The raw returns are accumulated monthly starting at one year after fiscal year end. First the data sample is sorted and ranked on the lagged PSK_USE rate into decile portfolios every t year at the end of June from 1985 to 2009. Using the June t PSK_USE growth cutoffs, portfolios are formed that are held for one year from July of t year to June of $t+1$ and then rebalanced.

Next, Figure 2.5 (a) shows that firms in the low CVT_DEBT decile have, on an equal-weighted basis, higher earnings across the 10 years than do firms in the high CVT_DEBT decile. Mean Q values for firm value as shown in Figure 2.5 (c) are higher for firms in the high decile prior to year -1 and cross over later, remaining significant up to year 2. However, firms in the low CVT_DEBT decile earn higher mean monthly returns, as shown in Figure 2.5 (e), than do firms in the high CVT_DEBT decile prior to year 1. In later periods, the spread becomes meaningless, due to the similar level of capital investment, firm size, and the book-to-market ratios.

Finally, for PSK_USE (see Figure 2.6 (a & c), firms in the low PSK_USE decile always have higher positive earnings, but lower firm value across the 10 years than do the firms in the high PSK_USE decile. In the case of mean monthly stock returns shown in Figure 2.6 (e) on an equal-weighted basis, the trend prior to year 1 shows that firms in the high decile earn higher returns than firms in the low decile; but after year 1 the trend is indistinguishable. However, on a value-weighted basis, see Figure 2.6 (f), after year -3 firms in the high decile earn higher returns than firms in the low decile; but again the spread has no meaning owing to the similarities in firm size and other fundamentals across firms in both deciles.

2.4.3.2 Characteristics and risk adjusted long-run abnormal returns

This sub-section finds that there are incremental returns over and above the characteristics that mimic risk and the abnormal alphas controlling for asset-pricing models such as the capital asset pricing model (CAPM) (Sharpe, 1964), the Fama and French (1993) three-factor model (FF3) and Carhart's (1997) four-factor model (FF4). Table 2.8 reports the average monthly abnormal returns on the five factors sorted by extreme decile

portfolios in the first, second and third years subsequent to portfolio formation. For the period July 1986 to December 2010, monthly decile portfolios are formed by ranking on the previous year's latent factor. There is a 1-year lag between the fiscal year end and the month in which the portfolio is formed. The time-series average of monthly equal(value)-weighted abnormal returns (the abnormal returns are the difference between the stock returns and the size, book-to-market, and momentum matched benchmark portfolio returns) are reported for each decile and for a hedge portfolio. The hedge portfolio is created by going long in the lowest ranked latent factor portfolio and short in the highest ranked latent factor portfolio. Controls are made for risk (size and book-to-market) and other important return anomalies and the alpha estimates (t -statistics) are reported for CAPM, FF3, and FF4 for both equal and value-weighted hedge portfolios. Note that for reasons of brevity only four factor alphas are reported in Table 2.8.

Apart from the full sample, results are reported for three capitalisation-ranked portfolios. Following Fama and French (2008), the stocks under the 20th percentile of market cap at the financial year-end are grouped as micro stock. The stocks between the 20th and 50th percentiles are grouped as small stocks and those above the median are grouped as large stocks.

The results for characteristic adjusted hedge returns and under pricing factor controls are similar. Panel A of Table 2.8 reports results on an equal-weighted basis when firms are ranked by FIN_FLEX. For the full sample, the average monthly hedge returns for raw returns in year $t+1$ and adjusted hedge returns for years $t+1$, $t+2$, and $t+3$ are 0.69%, 0.48%, 0.38% and 0.28%, with t -statistics of 8.36, 9.02, 10.35, and 11.83 respectively. The evidence from the full sample is also robust on a value-weighted basis. Across size groups,

for large stock the spread is negative, and firms in the high decile earn greater abnormal returns than those in the low decile. The spread of large stocks is small but still significant. Abnormal returns for small and micro stock are positive and significant for both raw and risk-adjusted portfolios for years $t+1$ to $t+3$. There is no significant variation in results when controlling for pricing factors and similar results are obtained on a value-weighted basis. For the whole-sample period there is an obvious size effect which may be due to a non-monotonic relation between FIN_FLEX and LSIZE, but within size groups no size effect exists. It appears that firms that reduce financial flexibility (show negative change) earn greater average monthly abnormal returns compared to firms that increase financial flexibility (show positive change). Firms that reduce the financial flexibility are smaller in size, and highly levered, which make them financially constrained, but still they take a risk of taking on the growth options as compared to firms that increase financial flexibility. Thus, low flexibility firms are required to earn higher returns than high flexibility firms, which explain the effect of size in earnings abnormal returns.

In the case of portfolios ranked by ST_CREDIT, results in Panel B indicate that on an equal-weighted basis the characteristics of the adjusted abnormal return hedge portfolio are significant only for year $t+3$. The large stock hedge portfolio at years $t+1$, $t+2$ and $t+3$ earns adjusted abnormal returns of -0.15%, -0.16%, and -0.17% at 1% significance level, respectively. For small stocks, the magnitude of the abnormal return is slightly higher and is more significant than it is for large stocks, but the spread is still negative meaning firms that show a decrease in short-term credit earn a lower return than firms that show increases in short-term credit. For micro stocks, only the raw return of 0.20% in year $t+1$ and the

adjusted hedge returns of 0.28% for the subsequent year $t+3$ are significant. The results are similar when controlling for pricing factors and when obtained on a value-weighted basis.

The results for LT_INV ranked portfolios in Panel C indicate that for the market sample on an equal-weighted basis the mean monthly adjusted abnormal returns for years $t+1$ to $t+3$ range from 0.34% to 0.20%, and are significant at the 1% level. Market raw hedge returns is 0.63% in year $t+1$. For the large stock portfolio the raw return in $t+1$ year is 18% and the characteristics adjusted abnormal returns are significant only for years $t+1$ and $t+2$. The magnitude of abnormal returns doubles for small stock portfolios compared to large stock portfolios. For micro stocks, the average monthly adjusted hedge returns for the subsequent three years are 0.50%, 0.43%, and 0.33%, with t -statistics of 9.07, 10.15, and 9.32 respectively.

In the case of CVT_DEBT ranked portfolios, reported in Panel D, the only significant incremental returns observed are in year $t+1$ for the market sample. Across the size samples a significant raw return is observed in $t+1$ and adjusted abnormal returns in years $t+2$ and $t+3$. The results for large stocks are more pronounced on a value-weighted basis. However, for small and micro firms, negative and insignificant abnormal returns are found, except for characteristics adjusted abnormal returns in year $t+3$.

Finally, PSK_USE ranked portfolios, reported in Panel E, indicate that on an equal-weighted basis the hedge portfolio average monthly characteristics adjusted abnormal returns are significant only across the size samples. The return spread is negative for large stocks but significant at years $t+1$, $t+2$, and $t+3$ at -0.09%, -0.14%, and -0.14% with t -stats of -1.96, -4.45, -5.93, respectively. For small stocks, the hedge returns are positive for the

Table 2.8

Average monthly abnormal returns for latent factors decile portfolios 1, 2 and 3 years after portfolio formation (whole sample period)

Panel A: <i>FIN_FLEX</i> ranked portfolios																
	Equal weighted				t-values (Equal weighted)				Value weighted				t-values (Value weighted)			
	<i>raw_ew</i> <i>t+1</i>	<i>adj_ew</i> <i>t+1</i>	<i>adj_ew</i> <i>t+2</i>	<i>adj_ew</i> <i>t+3</i>	<i>t(raw_ew)</i> <i>t+1</i>	<i>t(adj_ew)</i> <i>t+1</i>	<i>t(adj_ew)</i> <i>t+2</i>	<i>t(adj_ew)</i> <i>t+3</i>	<i>raw_vw</i> <i>t+1</i>	<i>adj_vw</i> <i>t+1</i>	<i>adj_vw</i> <i>t+2</i>	<i>adj_vw</i> <i>t+3</i>	<i>t(raw_vw)</i> <i>t+1</i>	<i>t(adj_vw)</i> <i>t+1</i>	<i>t(adj_vw)</i> <i>t+2</i>	<i>t(adj_vw)</i> <i>t+3</i>
<i>Market Hedge (L-H)</i>	0.69	0.47	0.38	0.28	8.36	9.02	10.35	11.83	0.42	0.32	0.25	0.18	5.34	6.67	7.35	8.38
Four Factor α	0.69	0.48	0.40	0.29	8.13	8.77	10.47	12.04	0.43	0.33	0.27	0.19	5.28	6.53	7.63	8.67
<i>Large Firms Hedge (L-H)</i>	-0.32	-0.10	-0.02	-0.05	-3.28	-1.77	-0.72	-1.96	-0.33	-0.11	-0.03	-0.06	-3.35	-1.94	-0.88	-2.09
Four Factor α	-0.29	-0.10	-0.03	-0.06	-2.93	-1.67	-0.80	-1.99	-0.29	-0.10	-0.03	-0.06	-2.97	-1.82	-0.94	-2.15
<i>Small Firms Hedge (L-H)</i>	0.17	0.11	0.10	0.07	1.67	1.75	2.12	1.96	0.16	0.11	0.10	0.07	1.57	1.68	2.12	1.95
Four Factor α	0.15	0.13	0.12	0.08	1.44	1.91	2.40	2.19	0.14	0.13	0.12	0.08	1.34	1.83	2.41	2.19
<i>Micro Firms Hedge (L-H)</i>	0.98	0.67	0.51	0.38	10.30	9.35	10.69	11.51	0.83	0.58	0.42	0.31	8.71	8.30	8.81	9.33
Four Factor α	0.98	0.67	0.52	0.39	9.78	9.05	10.71	11.77	0.83	0.59	0.44	0.33	8.28	8.06	8.99	9.69
Panel B: <i>ST_CREDIT</i> ranked portfolios																
<i>Market Hedge (L-H)</i>	0.14	0.01	0.04	0.06	2.42	0.29	1.26	3.12	0.05	0.05	0.05	0.08	0.75	1.42	1.80	3.55
Four Factor α	0.15	0.03	0.05	0.07	2.47	0.58	1.60	3.44	0.06	0.05	0.05	0.08	0.79	1.37	1.58	3.27
<i>Large Firms Hedge (L-H)</i>	-0.14	-0.15	-0.16	-0.17	-2.27	-3.97	-6.57	-7.46	-0.11	-0.14	-0.15	-0.16	-1.81	-3.79	-6.41	-7.36
Four Factor α	-0.14	-0.15	-0.16	-0.17	-2.32	-3.79	-6.16	-7.11	-0.11	-0.13	-0.15	-0.16	-1.89	-3.65	-6.02	-7.05
<i>Small Firms Hedge (L-H)</i>	-0.15	-0.20	-0.22	-0.17	-1.99	-3.00	-4.65	-5.29	-0.16	-0.20	-0.22	-0.18	-2.12	-3.10	-4.66	-5.43
Four Factor α	-0.14	-0.18	-0.20	-0.16	-1.73	-2.59	-4.04	-4.71	-0.15	-0.19	-0.20	-0.17	-1.87	-2.72	-4.06	-4.86
<i>Micro Firms Hedge (L-H)</i>	0.20	0.03	0.06	0.09	2.98	0.50	1.67	3.60	0.14	0.01	0.02	0.05	2.21	0.10	0.66	2.24
Four Factor α	0.21	0.04	0.07	0.10	3.06	0.70	1.87	4.00	0.16	0.02	0.03	0.06	2.30	0.30	0.93	2.67
Panel C: <i>LT_INV</i> ranked portfolios																
<i>Market Hedge (L-H)</i>	0.63	0.34	0.27	0.20	11.35	8.75	8.20	7.44	0.50	0.27	0.20	0.15	9.03	7.53	6.67	6.07
Four Factor α	0.61	0.34	0.27	0.20	10.72	8.44	7.93	7.22	0.47	0.26	0.20	0.15	8.37	7.11	6.27	5.76
<i>Large Firms Hedge (L-H)</i>	0.18	0.11	0.06	0.00	3.05	2.79	1.93	0.17	0.18	0.11	0.06	0.01	3.01	2.95	2.11	0.34
Four Factor α	0.16	0.10	0.05	0.00	2.52	2.43	1.50	-0.13	0.15	0.10	0.05	0.00	2.48	2.57	1.67	0.03
<i>Small Firms Hedge (L-H)</i>	0.41	0.21	0.05	0.00	5.36	3.54	1.22	-0.03	0.41	0.21	0.06	0.00	5.38	3.61	1.30	0.03
Four Factor α	0.39	0.19	0.04	-0.01	4.99	3.16	0.97	-0.28	0.39	0.20	0.05	-0.01	5.03	3.23	1.05	-0.23
<i>Micro Firms Hedge (L-H)</i>	0.86	0.50	0.43	0.33	12.80	9.45	10.58	9.72	0.78	0.47	0.40	0.31	11.48	8.75	9.98	9.20
Four Factor α	0.85	0.50	0.43	0.33	12.24	9.07	10.15	9.32	0.77	0.46	0.39	0.31	10.92	8.27	9.31	8.65

Table 2.8 (Continued)

Panel D: <i>CVT_DEBT</i> ranked portfolios																
<i>Market Hedge (L-H)</i>	0.03	-0.02	0.01	0.04	0.84	-0.67	0.69	2.78	0.08	-0.01	0.04	0.04	2.74	-0.23	2.39	3.29
Four Factor α	0.03	-0.01	0.02	0.04	0.96	-0.28	0.93	2.96	0.09	0.00	0.04	0.05	2.81	0.08	2.55	3.42
<i>Large Firms Hedge (L-H)</i>	0.19	0.06	0.13	0.12	3.79	1.40	4.14	5.03	0.19	0.06	0.13	0.13	3.83	1.51	4.43	5.80
Four Factor α	0.17	0.05	0.12	0.12	3.32	1.16	3.77	4.76	0.17	0.05	0.12	0.13	3.34	1.22	3.97	5.43
<i>Small Firms Hedge (L-H)</i>	0.05	-0.07	-0.06	-0.03	0.72	-1.19	-1.72	-1.07	0.04	-0.08	-0.06	-0.03	0.59	-1.37	-1.79	-1.07
Four Factor α	0.09	-0.03	-0.03	-0.01	1.29	-0.54	-0.80	-0.50	0.09	-0.04	-0.03	-0.01	1.18	-0.71	-0.86	-0.50
<i>Micro Firms Hedge (L-H)</i>	-0.06	-0.02	0.01	0.06	-1.18	-0.56	0.28	3.21	-0.01	-0.01	0.04	0.06	-0.26	-0.18	1.45	3.20
Four Factor α	-0.06	-0.01	0.01	0.07	-1.18	-0.32	0.35	3.29	-0.02	0.00	0.04	0.07	-0.35	-0.02	1.39	3.24
Panel E: <i>PSK_USE</i> ranked portfolios																
<i>Market Hedge (L-H)</i>	-0.02	0.01	-0.03	-0.02	-0.50	0.26	-1.04	-1.22	-0.02	0.01	-0.01	-0.01	-0.42	0.49	-0.29	-0.37
Four Factor α	-0.03	0.00	-0.03	-0.03	-0.66	0.00	-1.12	-1.35	-0.02	0.01	-0.01	-0.01	-0.40	0.42	-0.33	-0.55
<i>Large Firms Hedge (L-H)</i>	-0.25	-0.09	-0.14	-0.14	-3.94	-1.96	-4.45	-5.93	-0.23	-0.08	-0.14	-0.14	-3.65	-1.82	-4.34	-6.01
Four Factor α	-0.22	-0.06	-0.12	-0.13	-3.30	-1.35	-3.77	-5.38	-0.20	-0.06	-0.12	-0.13	-3.02	-1.22	-3.67	-5.46
<i>Small Firms Hedge (L-H)</i>	0.09	0.14	0.12	0.13	1.20	2.43	2.87	3.42	0.09	0.14	0.12	0.13	1.20	2.44	2.90	3.47
Four Factor α	0.08	0.12	0.09	0.10	0.96	2.02	2.09	2.61	0.08	0.12	0.09	0.10	0.96	2.03	2.12	2.67
<i>Micro Firms Hedge (L-H)</i>	-0.07	-0.03	-0.07	-0.07	-1.15	-0.72	-1.94	-2.39	-0.08	-0.04	-0.05	-0.04	-1.58	-1.00	-1.43	-1.69
Four Factor α	-0.09	-0.05	-0.07	-0.07	-1.50	-1.05	-1.92	-2.26	-0.09	-0.05	-0.04	-0.04	-1.73	-1.12	-1.27	-1.58

For the period of July 1986 to December 2010, monthly decile portfolios are formed by ranking on the last year latent factor. There is a one-year lag between the fiscal year end and the portfolio formation month. The time-series average of monthly equal (value) weighted abnormal returns (the abnormal returns are the difference between the stock returns and size, book-to-market, and momentum matched benchmark portfolio returns) are reported for each decile and hedge portfolio with respective t -statistics. Hedge portfolio is created by going long in the lowest ranked latent factor portfolio and short in the highest ranked latent factor portfolio. Comparatively, the alpha estimates (t -statistics) are obtained for CAPM, Fama and French's (1993) three-factor model, and Carhart's (1997) four-factor model for both equal- and value-weighted hedge portfolios but for clarity only results for the four-factor model are reported. The bold numbers indicate significance at less than the 10% level.

subsequent three years at 0.14%, 0.12%, and 0.13% with t -statistics of 2.43, 2.87, and 3.48, respectively. For micro stocks, the adjusted hedge returns for the subsequent three years are -0.03%, -0.07%, and -0.07% with t -statistics of -0.72, -1.94, and -2.39, respectively. The results are similar when controlling for pricing factors and when results are obtained on a value-weighted basis.

2.4.3.3 Annual buy-and-hold returns

Figure 2.7 (a-j) shows the abnormal annual buy-and-hold returns and the spread of low and high latent growth portfolios from 1986 to 2008 both on an equal(value)-weighted basis. Also obtained are the (unreported) abnormal returns for both equal- and value-weighted portfolios for three capitalisation levels and for a whole sample portfolio that excludes micro stocks. In a trading strategy, which is based on buying the low latent factors decile and selling short the high latent factors decile on an equal-weighted basis, FIN_FLEX is profitable in 20 years, ST_CREDIT is profitable in 17 years, LT_INV is profitable in 19 years, CVT_DEBT is profitable in nine years, and PSK_USE is profitable in nine out of the 23 years in the sample. Mean annual abnormal returns on the CVT_DEBT and PSK_USE ranked extreme deciles are quite similar, so their hedge (low-minus-high) almost converges to zero. This is probably due to the similar level of positive profitability and firm size across the two extreme deciles.

Figure 2.8 (a & b) shows that the equal(value)-weighted annual buy-and-hold returns spreads for FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE for the whole sample are 17%, 2.3%, 10.8%, -2.0%, and -0.3% (13%, -5%, 8%, -0.4%, and 0.6%) respectively. For the large stock only the spread for both equal(value)-weighted portfolios is 7.5%, -0.2%, 4.3%, 0.9%, and -1.2% (7.7%, -0.06%, 4.3%, 0.8%, and -1.2%)

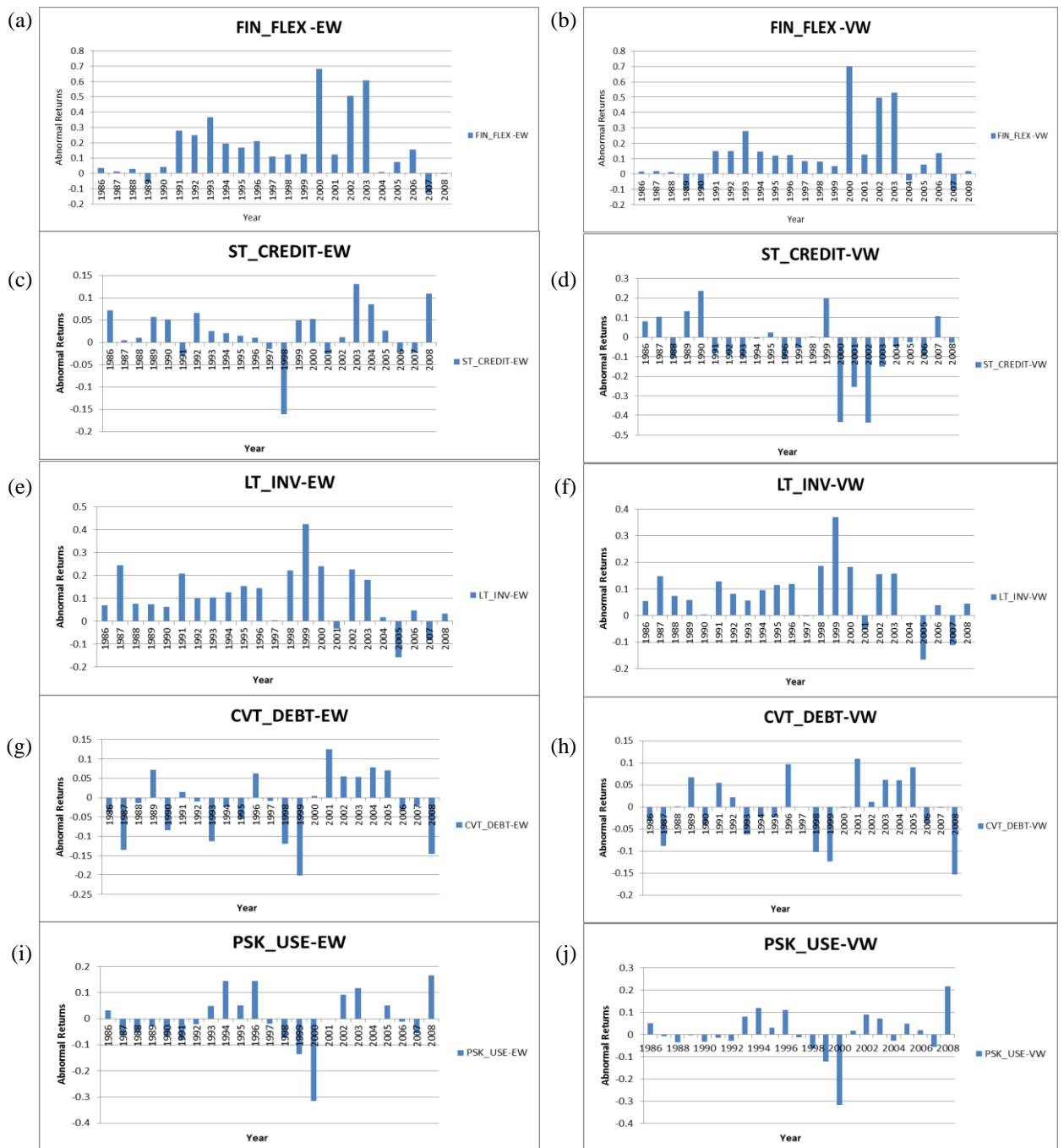


Figure 2.7 (a-j): Time-series average annual buy-and-hold returns spread

The figures show the annual buy-and-hold raw returns spread between low and high deciles based on decile ranks of latent growth factors (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE) in the previous year for the period 1986 to 2008. There is a one-year lag between the fiscal year end and raw returns. The capitalisation levels are constructed following NYSE capitalisation break points as defined in Fama and French (2008).

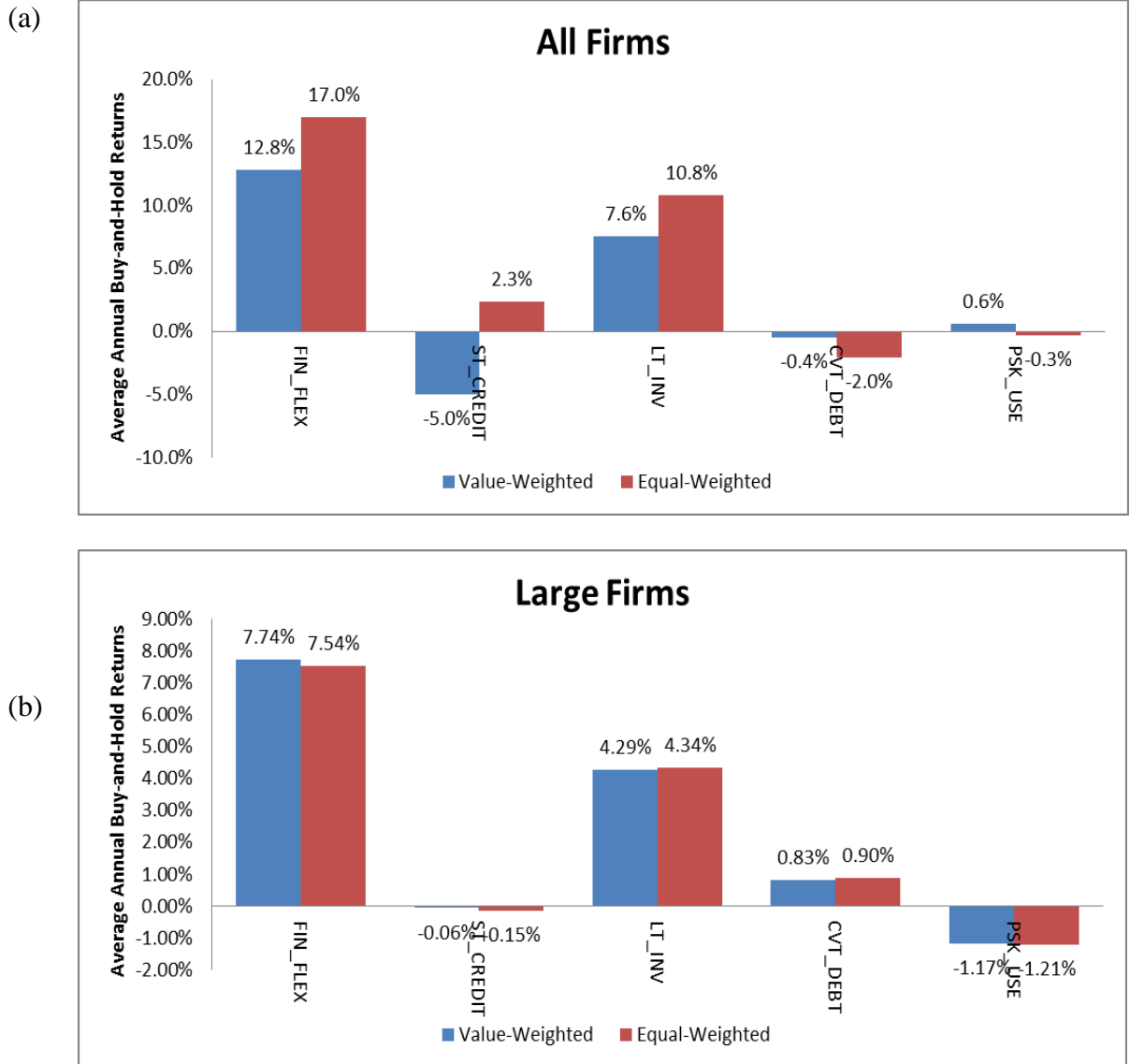


Figure 2.8 (a & b): Annual buy-and-hold returns

The figures show the average annual buy-and-hold raw returns spread between low and high deciles based on decile ranks of latent growth factors in the previous year for a period of 1986 to 2008 across the sample and large firms. There is a one-year lag between the fiscal year end and raw returns.

respectively. The above buy-and-hold returns based upon each latent factor suggest that these factors could be important in explaining the market performance of stocks; a better test is to put these factors into a controlled model to control their unique contribution to market returns (See., section 2.4.2).

2.5 Conclusion

Factor analysis provides a better measure of common corporate financing decisions among firms by analysing the shifts and interactions among a comprehensive set of balance sheet items. This study adds to the large body of literature that examines the interactions of investment and financing decisions in a setting with market imperfections³⁸. These new measures are designed to overcome missing variable biases that are expected due to in-built balance sheet data duplication and redundancies. These issues further exaggerate the problem of investor limited attention (Hirshleifer, et al., 2004).

A single or pre-determined group of accounting items cannot be representative of certain types of business decision. This study argues that when a firm implements a corporate financing decision, this is reflected in various and dynamic shifts in accounting items. For example, double-entry bookkeeping and distinct accounting practices ensure information of a certain dimension gets buried in different accounts on both sides of the balance sheet. Furthermore, one accounting item could be a reflection of many different business activities. It is thus critical to identify common interactive movements when evaluating accounting information.

The study shows that the information contained in the complete balance sheet (including supplementary items) can be optimally decomposed into five latent factors. The

³⁸ See for example, Myers (1974), Jensen and Meckling (1976), Myers (1977), Myers and Majluf (1984), Dotan and Ravid (1985), Dammon and Senbet (1988), Parrino and Weisbach (1999), Barclay and Smith (1995b), Mehran et al., (1999), Rajan and Zingales (1995), Mello and Parsons (1992), Mello et al. (1995), Fries et al. (1997), Ericsson (2000), Mauer and Ott (2000), Goldstein et al. (2001), Parrino et al. (2005), Baker and Wurgler (2002), Ju et al. (2005) Welch (2004), Morellec (2004), Leary and Roberts (2005), Mauer and Sarkar (2005), Flannery and Rangan (2006), Brennan and Schwartz (1984), Mauer and Triantis (1994), Leland (1998), Moyen (2007), Titman and Tsyplakov (2007), La Porta et al. (1997), La Porta et al. (1998), Graham et al. (2001), and Tufano (2001), among others.

latent factors mimic basic corporate financing decision types that are fundamental to business value. Specifically, it is found that these five factors contain about 70% of all accounting ratio changes. These extracted factors are better forecasters of future stock returns than are other well-established factors known in the literature such as firm size, book-to-market, momentum, and accounting growth measures (including total asset growth).

The analysis, focusing on the Q theory of investment, suggests that latent factors can predict the firm value for up to three years and thus add value to firm. The valuations (Q) are positive irrespective of the negative or positive shifts in the latent factors. In addition, the study is able to relate the latent factors with future cash flow and earnings, the relationship being particularly strong with cash flow rather than with earnings.

Overall, the new factors are certain firm decisions that reflect their financial status and expected investment opportunities. The estimation of these decisions requires an understanding of the correlational structure of accounting data that eliminate measurement errors. Consequently, these factors overcome the issues arising from investors limited attention, to some extent, if not completely. These factors are aligned with Q theory in relation to firm fundamentals (firm value and earnings) and stock returns. This study introduces is an integrated measure of both financial and investment flexibility (FIN_FLEX) unknown in prior research. Finally, in contrast to Gow et al. (2010) and consistent with Petersen (2009) in multiple test methods, this study concludes that two-way standard error clustering is not necessary both in accounting and finance analysis owing to results that are similar to one-way time-fixed standard error clustering regressions.

CHAPTER THREE

ESSAY TWO

FINANCIAL FLEXIBILITY: A NEW RISK FACTOR?

This chapter investigates the sources of return predictive power of the changes in financial flexibility (latent factor) in both risk-based and mispricing-based empirical models. Section 3.1 presents a brief overview of the study and in Section 3.2 the data and methodologies are described. Section 3.3 presents and discusses the descriptive statistics, correlation matrix, and the relationship of financial flexibility with prominent factors. Section 3.4 discusses the empirical findings of asset pricing models used in this study. Section 3.5 concludes Chapter 3. The chapter's references are presented in the bibliography section of this thesis.

Financial flexibility: A new risk factor?

Abstract

This essay finds that the changes in financial flexibility based on a zero-cost return portfolio (FLEX) bear significant positive premium of 0.69% that is driven by covariance (risk). Neither prominent pricing characteristics (e.g., firm size, book-to-market ratio, and buy-and-hold returns) nor associated pricing factors (e.g., SMB, HML, and UMD) can capture the risk ingrained in the FLEX. On the contrary, FLEX makes the HML redundant by capturing its inherent risk. FLEX is an unanticipated factor constructed using unobserved information. This factor has an average t-statistic of 5.56, which is well above the latest yardstick of 3.0. Conclusively, FLEX is a new priced risk factor.

JEL classification: G11; G12; G31; G32

Key words: Financial flexibility, cross-section of stock returns, asset pricing, anomalies, risk factors, mispricing, risk.

Acknowledgements: I am grateful to the participants at the 2014 SEF Research Symposium at Victoria University, Wellington, the School of Economics and Finance Seminar at Massey University, Palmerston North, New Zealand, and Fourth SIRCA Young Research Workshop in Sydney for useful comments.

3.1 Introduction

Financial flexibility – firm’s capacity to respond in a timely and value-maximizing way to unexpected shocks to the firm’s cash flows or investment opportunity set (Denis, 2011) – is considered by the US CFOs as the most important determinant of capital structure (Graham and Harvey, 2001) and the critical missing link for an empirically viable capital structure theory (DeAngelo and DeAngelo, 2007). Financial flexibility considerations have a significant impact on financial policies, which include cash management, capital structure, and payout policies (Denis, 2011). For example, firms attain flexibility via the management of cash holding, dividend initiations, stock repurchases, and unused debt capacity (DeAngelo and DeAngelo, 2007; Denis, 2011; Denis and McKeon, 2012; Marchica and Mura, 2010). Bonaimé et al. (2013) find risk management and payout policies contributing to financial flexibility. Apart from firm-specific implications, there is limited evidence that relates financial flexibility with stock return predictability. The first essay in this thesis finds that a shift in financial flexibility (FIN_FLEX)³⁹ is a robust negative determinant of stock returns⁴⁰. FIN_FLEX is able to earn substantial abnormal returns annually ranging from 17% (13%) at the equal (value) weighted basis. However, this essay examines whether information relating to the FIN_FLEX is priced or not.

Ranking on FIN_FLEX characteristics, this study constructs a zero-cost return mimicking portfolio (financial flexibility factor ‘FLEX’). This factor is tested as to whether

³⁹ Employs correlational structure of accounting growth information content, to obtain the unobserved measure of shifts in financial flexibility, following Subrahmanyam’s (2010) suggestions in return predictability tests.

⁴⁰ Walkshäusl and Lobe (2011) argue that there is still enough gap to optimise the ability to explain the cross-section of average stock returns by adding factors, using new variable definitions or a new construction methodology.

it aligns to the asset pricing or mispricing hypothesis⁴¹. This study complements the last 20 years' literature that documents numerous accounting information based return anomalies⁴². The prominent asset pricing models such as the classic capital asset pricing model (CAPM) (Lintner, 1965; Sharpe, 1964) and the three-factor risk model (FF3) (Fama and French, 1993, 1996) are unable to explain these return anomalies. The asset pricing literature is divided into two schools of thought. One school of thought considers prices as random (Kendall and Hill, 1953), and patterns in prices as a result of chance (Roberts, 1959). Supporting the chance model, Fama (1970) finds capital markets are efficient (rational) and presents the three-tier efficient market hypothesis (EMH). In contrast, in the other school of thought Kahneman and Tversky (1979), Thaler (1993), Shleifer (2000), and Kahneman and Tversky (2000) show humans are irrational in making decisions owing to cognitive biases. These biases create inefficiencies that surface as patterns or trends in financial markets. And there is hardly anything efficient or rational in determining the price of a financial instrument (Prechter Jr, 2001). Consequently, both rational and irrational camps provide their suited justifications of any data pattern, and reveal the severe contradictions between the two schools of thoughts.

One key argument this research makes is that investors appreciate the changes in the firm's financial flexibility owing to its greater value for CFOs. Thus, investors are expected to react rationally to the arrival of such information and readily incorporate it into prices. Supporting this argument, the evidence suggests that FLEX is a priced factor, and it bears a

⁴¹ Investors underreact to characteristics-related information and misprice the stocks (Daniel and Titman, 1997; Daniel et al. 2001).

⁴² See, for example, Sloan (1996), Richardson and Sloan (2003), Fairfield et al. (2003), Hirshleifer et al. (2004), Titman et al. (2004), Daniel and Titman (2006), Bradshaw et al. (2006), Lyandres, Sun and Zhang (2008), Fama and French (2008), Pontiff and Woodgate (2008), Xing (2008), Cooper et al. (2008), Chen et al. (2011).

positive and significant premium⁴³. This premium is found independent of existing asset pricing factor models and does not relate to its own or any other characteristics. On the contrary, financial flexibility may potentially be related to, and be able to explain, the risk inherent in the book-to-market ratio. Conclusively, FLEX is a potential candidate for a (risk) factor model.

The FLEX's candidature as a new risk factor is consistent with the arguments extended by Pukthuanthong and Roll (2014). The equities in the asset portfolios closely resemble the real assets. The FIN_FLEX characteristic is not known *a priori* (being latent), and FLEX's risk premium is driven by covariances. Roughly based on the descriptive statistics reported in Table 2.2, the FLEX can be closely matched with the portfolios of low minus high leverage firms. It can also be comparable to the hedge portfolio of high (low) book-to-market ratio firms, and to the low (high) R&D intensity firms. In addition, the average monthly FLEX premium of 0.69% with *t*-statistics of 5.56 is considerably large enough to pass a new benchmark of significance criteria of the *t*-statistics greater than 3.0 (Harvey et al., 2013). The concept of financial flexibility is quite old, but has been limited to corporate finance application and not to asset pricing. Consequently, we may expect a decline in FLEX premium soon after publication due to more informed tradings or because of statistical biases if any (McLean and Pontiff, 2014).

⁴³ This essay labels the measure of change in financial flexibility as FIN_FLEX, financial flexibility factor (return portfolio) as FLEX and its covariance as b_{FLEX} .

The financial flexibility (FIN_FLEX) measure under examination is the first latent growth factor obtained via factor analysis in essay one⁴⁴. It comprises of components such as simultaneous shifts in common equity-total, common equity-liquidation, cash and equivalents, and common equity-tangible scaled by lagged total assets. The components are forward looking (Ang and Smedema, 2011), preserve flexibility, and enhance firm value (Denis, 2011; Gamba and Triantis, 2008). FIN_FLEX accounts for the largest part of common variance (about 40%) in the complete balance sheet and is orthogonal to other factors obtained via factor analysis. The component's composition indicates that it is an integrated unobserved measure of both investment and financial flexibility.

This study observes that cash holding is highly causal to equity financing. It is because firms on average issue equity more often than debt (Fama and French, 2005; Frank and Goyal, 2003). Furthermore, firms issue equity to save the unused debt capacity for future investment distortions that may occur due to shares' undervaluation (DeAngelo and DeAngelo, 2007). The shifts in common equity-liquidation capture the concerns of the equity holders that favour payout policies to discipline the managers and to save their liquidation capital stake (Jensen, 1986; Officer, 2011). Supporting this argument, Gambia and Triantis (2008) find that the value of financial flexibility does not simply depend on the cost of external financing but also on the liquidation capital. The FIN_FLEX measure is thus constituted of all relevant components, and is well-aligned with the definition extended by Marchica and Mura (2010) that there is no well-defined measure of flexibility in the

⁴⁴ Tsay (2005) observes that the number of factors is more well-defined in statistical factor models than the way the three factors are obtained using the Fama-French (1993) approach. According to Fabozzi and Markowitz (2011) the latent (unobserved) factors are preferred because the observed factors may be measured with errors or have been already anticipated by investors. Factor analysis can explain complex phenomena through a small number of basic factors.

literature. This is an unobservable factor that depends largely on managers' perceptions of future growth options.

In the purview of the FIN_FLEX measure, this study argues that firms with high flexibility are less constrained, have higher growth opportunities, are larger in size and possess greater cash holdings compared to firms with low flexibility.⁴⁵ The high flexibility firms are usually profitable and have a higher investment-to-asset ratio, due to more tangible assets, and higher liquidation value. These firms employ tangible assets to avail of the positive NPV investment opportunities more often than the firms that are of low flexibility. In consonance with Berk et al.'s (1999) model, this reduces the systematic risk of the firm's cash flows and leads to lower than expected stock return among high flexibility firms. In contrast, low flexibility firms with low liquidation value pose a high risk to equity-holders, making equity capital costly. Accordingly, low flexibility firms should earn high returns to compensate for the higher cost paid for external finance. Apart from theoretical arguments, the empirical literature that distinguishes between risk-based and non-risk explanation for investment and financing growth anomalies are mixed⁴⁶.

Similar to the first essay, this essay finds an important variation in the distribution of the FIN_FLEX and a significant negative relation to subsequent monthly excess return. The relation is robust in terms of the inclusion of prominent characteristics like firm size, book-to-market ratio, and momentum. In particular, regression controlling for gross

⁴⁵ Here and after high flexibility would mean (the most positive change), and low flexibility would mean (the most negative change).

⁴⁶ See Anderson and Garcia-Feijóo (2006), Xing (2008), Cooper et al. (2008), Gray and Johnson (2011), Titman et al. (2004), Lam and Wei (2011) among others.

profitability (GPA) – defined in Novy-Marx (2013) – does not erode the significance or return predictability of FIN_FLEX. Instead, GPA becomes an insignificant future excess return predictor. In unreported results, this study controls for other measures of profitability – firm earnings and free cash flows scaled by total assets – but finds FIN_FLEX persistently a robust return predictor.

Unlike the observed investment growth anomalies, the FIN_FLEX and its factor covariance loading (b_{FLEX}) are uncorrelated as per their Pearson (Spearman) correlation of -0.2% (-4.3%). Similarly, the Fama-MacBeth (1973) univariate cross-sectional regression of FIN_FLEX on the corresponding loading b_{FLEX} finds no significant correlation as the coefficient of b_{FLEX} is insignificantly negative with t -statistics of -0.65. Furthermore, in Fama-MacBeth's (1973) cross-sectional regression that includes both characteristics (FIN_FLEX) and loading (b_{FLEX}), the regression coefficient of loading is significantly different from zero at the 10% level. The characteristics and even the factor model controls (Fama-French three factor model and momentum) are unable to subsume the explanatory power of b_{FLEX} .

The main asset pricing test in this research follows the two-stage cross-sectional regression approach extended by Cochrane (2001) for determining whether the FLEX factor is priced.⁴⁷ Under the risk-based doctrine, the study aims to find a positive and significant premium on the financial flexibility factor (FLEX). The FLEX portfolio is constructed following Core et al.'s (2008) and Kim and Qi's (2010) methodology for

⁴⁷ See, Black et al. (1972), Fama and MacBeth (1973), Grauer and Janmaat, (2009, 2010), Brennan et al. (2004), Campbell and Vuolteenaho (2004), Petkova (2006), Huang and Wang (2009), Core et al. (2008), Kim and Qi (2010) and Ogneva (2012) on the use of the two-stage cross-sectional regression framework.

accrual quality factor, though it differs on some nuances. In particular, this essay runs two-stage cross-sectional regressions for 25 LSIZE-LBTM portfolios; 100 portfolios ranked on FIN_FLEX latent growth factor; and for triple sorted 64 LSIZE-LBTM-FIN_FLEX portfolios. The results show that there is a significant positive FLEX premium for the above three types of portfolio tests. The evidence is in favour of the study's motivation that FLEX is a priced factor. The results are robust to the inclusion of the prominent factors such as MKT_RF, SMB, HML, UMD and LIQ. Instead, FLEX makes their slope coefficients insignificant.

In further asset pricing tests, this study in Daniel and Titman's (DT, 1997) tests concludes that in the tug-of-war between covariances and characteristics, covariance drives the return predictability associated with the FIN_FLEX characteristic. The DT test outcome can be explained via a negative correlation between FLEX factor loading (b_{FLEX}) and its characteristic (FIN_FLEX). Then, a positive correlation between factor loading and subsequent returns resulting in an outcome where the factor loading is driving the returns.

The final objective the research aims to ascertain is whether the FLEX factor is non-redundant for explaining average returns. The non-redundancy would mean that the existing pricing factors or models are not sufficient enough to capture the risk inherent with the FLEX factor. This essay runs univariate and multivariate time-series regressions of the FLEX factor on the Fama and French (1993) three-factor model (FF3), Carhart's (1997) four-factor model (FF3+UMD) or FF4, and on the Pastor and Stambaugh (2003) five-factor model (FF3+UMD+LIQ) or FF5. The results indicate that only two factors, MKT_RF (market factor) and HML (value factor), are related to the FLEX factor. However, a

positive significant intercept remains; that is the proportion of variation in FLEX unrelated to MKT_RF and HML. The positive intercept suggests that the FLEX factor cannot be replaced by these factors. On the other hand, FLEX alone is capable of capturing the risk inherent with the HML. Similarly, Fama and French (2013) show that RMW – the hedge return portfolios of robust and weak profitability stocks – and CMA – hedge return portfolio of low and high investment stocks – make the HML redundant in explaining average returns. In contrast, this study shows that the underlying characteristics of RMW and CMA get subsumed when predicting the future excess stock returns in Table 2.7. These findings reiterate the argument that financial flexibility factor is a new priced risk factor.

In order to support the risk-based explanation, in a further test this study provides evidence that future stock return predictability based on FIN_FLEX is independent of the information environment. The information environment is an important determining factor of investor misvaluation (Hirshleifer, 2001; Zhang, 2006). The misvaluation occurs when the information is being processed inefficiently or sub-optimally due to limited attentions or cognitive biases. This research shows that when controlling for various measures for information uncertainty, the financial flexibility consistently earns high returns on low FIN_FLEX quintile and low returns on high FIN_FLEX quintile. Thus, the evidence suggests that investors do not misvalue the financial flexibility related information. Furthermore, the alpha value of hedge portfolios (FLEX5-FLEX1) across different factor models (i.e., FF3, FF3 + UMD, and FF4 + UMD + LIQ) in the high information uncertainty (U4) portfolio is always low compared to the low information uncertainty (U1) portfolio. Accordingly, financial flexibility's (FIN_FLEX) return predictability is due to risk.

The rest of the essay is organised as follows. Section 2 describes the data, latent growth factor construction, zero-cost factor construction, empirical methods, and some preliminary results. Section 3 discusses the main results based on several asset pricing models. Section 4 concludes the study.

3.2 Data and Methodologies

3.2.1 Sample and data construction

This study uses a universe of non-financial and non-utilities, US stocks listed on NYSE, AMEX, and NASDAQ. The dataset consists of the CRSP monthly stock return files (1983-2011) and the Compustat annual industrial files for balance sheet data (1984-2009). The accounting data requires complete information for total assets for each firm. The balance sheet items are converted into the growth variables defined as:

$$X_{i,t-1} = (B_{i,t-1} - B_{i,t-2}) / \text{Total Asset}_{t-2}. \quad (1)$$

Where, $B_{i,t-1}$ stands for the balance sheet entry i at time $t-1$ and $B_{i,t-2}$ is the lagged balance sheet entry for the same firm. The individual growth item is further winsorized at 1% and 99% to normalise the data, since the balance sheet data are generally positively skewed. Subsequently, the data is standardised to match the scales of the diverse items for simplicity in making relations between the different variables of interest. This data treatment only focuses on balance sheet data ($X_{i,t-1}$) not on stock returns. The asset growth information ($X_{i,t-1}$) is then decomposed into five well-defined and orthogonal latent factors using a factor analysis procedure.

Among other firm characteristics is the log of the market value of equity (LSIZE) at June of year $t-1$. The log book-to-market ratio (LBTM) is measured as the ratio of the book value obtained at the fiscal year end that is June $t-1$ to the market value of equity at the December of year $t-1$. The first latent growth factor financial flexibility (FIN_FLEX) and other prominent firm characteristics are matched to the 12 monthly stock returns that start within 12 months after the end of the fiscal year. This covers any lag between the announcements and disseminations of the financial information.

3.2.2 Measurement and characteristics of FIN_FLEX

Graham and Harvey's (2001) survey paper indicates that financial flexibility is the first-order determinant of the capital structure and might be the missing link in capital structure theories (DeAngelo and DeAngelo, 2007). Financial flexibility is either obtained from internal capital or from external capital. External capital sources are debt financing and equity financing. The order, level and a mix of both external financing sources are subject to various considerations. The choice of financing is a function of the tax shields, cost, firm size, firm development stage, liquidation capital, distress costs and many other forms of financing constraints. These frictions make the financial flexibility interesting. Firms with higher flexibility are found with lower financing frictions than that of lower flexibility. The financial constraints hinder managers in availing of the first best investment opportunity (Almeida, et al., 2011). However, greater flexibility such as the accumulation of cash comes with a dark side, where managers exploit the excess cash in hand and invest in negative NPV projects (Jensen, 1986). The initiations of dividends and share repurchase actually serve as constraints on this dark side of excess liquidity (Officer, 2011). Firms can

maintain liquidity via holding the cash, cash from cash flows, and external equity. Consequently, firms in dynamic settings have a mix of these internal and external sources of liquidity.

The order of financing is a fundamental concept of the pecking order theory and the level of external debt or leverage is the prime characteristic of the trade-off theory of capital structure. The pecking order theory fails to explain why firms use equity as financing sources more often than debt; whereas, the trade-off theories fail to provide answers to questions like why firms time the market and increase leverage when stock prices rise, why there are delays in rebalancing the leverage, and why profitable firms forego interest tax shields by maintaining low debt (Fama and French, 2005). Besides this, firms regularly issue equity (Fama and French, 2005) that is more than the debt issues on average (Frank and Goyal, 2003). Equity issuance saves the debt capacity to avoid future investment distortions and creates value for firms (DeAngelo and DeAngelo, 2007). Thus, investment and financing decisions should be forward looking and not just dependent on the current state (Gamba and Triantis, 2008).

Denis (2011) argues that cash holding (a major source of flexibility conservation), is the critical link between liquidity management, capital structure and equity payout policies and is important enough to avail of first best investment opportunities. The equity holders under agency conflict are more worried about their stake, the liquidation capital. They see payout policies as a sign of a healthy firm and greater liquidation capital value. Then, we may argue that financial flexibility is an outcome of interactions among these three financial policies, suggesting that there exist great, regular and linear causal relationships

between cash holdings, common equity, and liquidation capital. Supporting these arguments, Gamba and Triantis (2008) find that investment and financial flexibility are close substitutes and financial flexibility is not only dependent on the cost of external capital, but also liquidation capital.

An interaction of these accounting attributes of firms' investment and financing decisions can be identified via bringing changes in relevant annual balance sheet items. The changes between two consecutive years will reflect that the investment and financing decisions are forward looking, not just marginal. These changes would reflect the managers' assessment of firms' financial health in order to avail of the first best investment option. The financial health, managers' perceptions, and agency problems are catalysts in determining the cash management, capital structure, and payout policies that help preserve flexibility.

The existing literature defines financial flexibility as the ability of a firm to respond in a timely and value-maximising manner to unexpected changes (negative or positive shocks) in firms' cash flows or investment opportunity sets (Ang and Smedema, 2011; Denis, 2011). The response to these unexpected shocks is again dependent on the management's assessment, which is unobserved in nature, making it difficult to observe the direct measure of financial flexibility. In fact, there is no well-defined measure of flexibility in the literature (Marchica and Mura, 2010).

This study uses the financial flexibility measure (FIN_FLEX) that is obtained using a factor analysis model employing a correlation structure of accounting information

content. The measure is based on the simultaneous shifts in annual balance sheet accounting items and thus is forward looking. It integrates both investment and financial flexibility related accounting information such as simultaneous shifts in common equity-total (Item#A60), common equity-liquidation (Item#A238), cash and equivalents (Item#A1), and common equity-tangible (Item#A11). The respective loads of these components 0.89, 0.88, 0.86, and 0.84 are time-series regularity. In 25 annual cross-sections of the US common stocks, FIN_FLEX is the first factor in the factor analysis model output, explaining almost 40% of the variation in the complete balance sheet changes observed. The measure is largely consistent with earlier literature on financial flexibility that manager's assessments are recorded as transactions in the accounting statements, and by bringing changes in the accounting items, management's perceptions can be observed. Using the factor analysis as a technique to capture the correlational structure of this accounting data, this study examines whether firms issue equity more often than debt, whether equity is more aligned with cash or not, and whether equity holders worry about the liquidation capital or not. This study further shows that financial flexibility is the first-order consideration for CFOs as it appears as the first factor in factor analysis output every year.

Table 3.1 shows the characteristics of financial flexibility by relating to prominent characteristics such as firm size and the book-to-market ratio. The table presents the mean and median values across the FIN_FLEX ranked quintiles. The ranking uses the full sample breakpoints that include stock from NYSE, AMEX, and NASDAQ. The ranking based on

Table 3.1
Characteristics of financial flexibility portfolios

Portfolios	FIN_FLEX	LSIZE	LBTM
Panel A: Mean			
Low	-0.473	4.452	0.512
2	-0.222	5.147	0.598
3	-0.141	5.439	0.568
4	-0.040	5.315	0.471
High	1.004	5.131	0.339
Panel B: Median			
Low	-0.395	4.276	0.446
2	-0.228	4.954	0.533
3	-0.157	5.342	0.503
4	-0.059	5.248	0.404
High	0.396	5.063	0.280

The table reports the prominent characteristics across the financial flexibility (FIN_FLEX) ranked quintiles for US common stocks (NYSE, AMEX, NASDAQ) over the period 1985-2009. Panel A reports mean values and panel B median values calculated across the years. The prominent characteristics include a log of firm size (LSIZE) and a log of the book-to-market ratio (LBTM).

the NYSE breakpoints also shows similar results (unreported). The mean and median values show that FIN_FLEX is unrelated to firm size, but is near monotonically related to book-to-market ratio. The strength and causality of the relationship between FIN_FLEX and book-to-market ratio is tested in further tests.

Table 3.2 presents the Fama-MacBeth (1973) cross-sectional regressions of subsequent monthly returns in excess of the risk free rate on the financial flexibility measure with prominent characteristics as controls. The control variables are: log of the market value of equity (LSIZE), log of the book-to-market ratio (LBTM), gross profitability (GPA), buy-and-hold returns for months -6 to -2 (BHRET6), and buy-and-hold

Table 3.2**Fama-MacBeth cross-sectional regressions**

$$R_{i,t+1} - R_{ft} = \alpha + \beta_1 \text{FIN_FLEX}_{i,t-1} + \beta_2 \text{Controls}_{i,t-1} + \varepsilon_{i,t+1}$$

	Model 1	Model 2	Model 3	Model 4
Intercept	0.0102*** (2.68)	0.0195*** (3.54)	0.0180*** (3.03)	0.0187*** (2.94)
FIN_FLEX	-0.0031*** (-4.22)	-0.0025*** (-3.91)	-0.0033*** (-3.89)	-0.0027*** (-2.75)
LSIZE		-0.0024*** (-3.90)	-0.0022*** (-3.85)	-0.0021*** (-3.57)
LBTM		0.0052** (2.20)	0.0026 (1.17)	0.0031 (1.30)
GPA				0.0031 (1.61)
BHRET6			-0.0035 (-1.18)	-0.0038 (-1.27)
BHRET36			-0.0022*** (-3.34)	-0.0021*** (-3.27)
Adj R-Sq	0.003***	0.016***	0.025***	0.029***

The table reports the full period (July, 1986 to December, 2010). The individual monthly subsequent stock return in excess of risk-free rate are regressed using Fama and MacBeth's (1973) regressions on the financial flexibility latent growth measure and the control characteristics. The controls include logarithm of firm size (LSIZE) that is measured stock price times shares outstanding. The logarithm of the book-to-market ratio is measured as the ratio of the market value of equity in June $t-1$ fiscal year end and the book equity estimated at the December of fiscal year $t-1$. The gross profitability (GPA) measured is the difference of revenues minus cost of goods sold scaled by total assets (Novy-Marx, 2013). The buy-and-hold return for months -6 to -1 (BHRET6) and buy-and-hold return for the month (-36 to -1). *** for 1%; ** for 5%; and * for 10% level of significance based on t -statistics.

returns for months -36 to -2 (BHRET36). The regression evidence shows that FIN_FLEX has significantly determined the subsequent excess returns even when controls are included. In subsequent sections, the method to convert the latent growth characteristic (FIN_FLEX) into return factor (FLEX) using the zero-investment factor mimicking portfolio strategy is described. The return factors can be helpful to identify the real force (risk or mispricing) of the return predictive power of the underlying characteristics. In order to conclude in favour of risk, FLEX is expected to have a significant and positive premium that would not be

explained by the existing asset pricing models. To further strengthen the argument this study asks whether there is a substantial correlation between characteristic (FIN_FLEX) and factor loading (b_{FLEX}). If the answer is no, then the study will conclude that financial flexibility based factor loading (covariance or risk) derives the average stock returns and not mispricing.

3.2.3 Financial flexibility return factor (FLEX)

In this section the study constructs the new financial flexibility latent growth based return factor, which may or may not proxy for risk. In order to be consistent with the other known accounting factors (for example accrual quality), it follows Francis et al. (2005) and Core et al. (2008) to construct the new return factor. The characteristic FIN_FLEX is matched with 12 month returns starting from the 12th month after the end of the fiscal year (for example, if the fiscal year ends in June $t-1$, then the returns are collected from July of year $t+1$ to June of year $t+2$).

The portfolio returns for this characteristic are measured in three ways. First, the full sample firm observations are sorted and ranked into five quintiles based on FIN_FLEX each month. Then, for each quintile the monthly equal-weighted average returns are calculated. The new factor (FLEX) buys bottom two and sells top two FIN_FLEX quintiles. The hedge portfolio for the long and short portfolios may mimic risk inherent in the financial flexibility. This study reports results for this factor portfolio as the main testing factor portfolio.

In the second way of constructing zero-cost investment return portfolios, the study controls for firm size. First firms are sorted into 100 percentile groups every month, and then each size group is divided into FIN_FLEX quintiles. The average equal-weighted returns are calculated for firm size and FIN_FLEX ranks. Then, hedge portfolios are constructed for each size rank by going long in bottom two FIN_FLEX ranked quintiles and short in top two FIN_FLEX ranked quintiles.

Then in the third case, this essay controls for both firm size and book-to-market ratio, in each month. It independently sorts the stocks into tercile on size, book-to-market, and FIN_FLEX. Their interaction gives us 27 portfolios. The average equal-weighted returns are allocated to each portfolio. Then, the hedge portfolios are constructed by going long in the bottom nine FIN_FLEX portfolios and short in the top nine FIN_FLEX portfolios. The last two methods of constructing the return factor help us to mitigate multicollinearity concerns, if any, owing to the significant correlations of FLEX factor with MKT_RF and HML. However, there is no impact on the final conclusion of the asset pricing test due to the variations in construction of the return mimicking factors.

3.2.4 Asset pricing methodologies

3.2.4.1 Two stage cross-sectional regressions (2SCSR)

This research uses the standard two-stage cross-sectional regressions (2SCSR) methodology, a prominent framework to explore a risk-based explanation for return anomalies. This methodology has been developed by Black et al. (1972) and Fama and MacBeth (1973) for CAPM testing along with others (Grauer and Janmaat, 2009, 2010). Brennan et al., (2004) and Petkova (2006) have used 2SCSR in the case of conditional

CAPM, Campbell and Vuolteenaho (2004) for the two-beta model, and Huang and Wang (2009) for an investment-based asset pricing model. Recently, the 2SCSR has been used for testing the accrual quality factor, and alternative three-factor model beyond the US market for risk-based explanations (Core, et al., 2008; Kim and Qi, 2010; Ogneva, 2012; Walkshäusl and Lobe, 2011).

Cochrane (2001) explains the procedure of two-stage cross-sectional regressions. The cross-sections of average returns are run on the betas, in order to distinguish whether a factor is a return or not. Intuitively, the risk premium should be directly related to the betas, where betas proxies risk exposures; thus, higher returns are expected for higher beta of assets to factors, which bear high risk premiums. Under 2SCSR methodology first time-series regressions of the following form are run, to get the betas on average future returns.

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FLEX}FLEX_t + \varepsilon_{it} \quad (2)$$

Where, $(R_{it} - R_{ft})$ are excess future returns for a portfolio of firms; (MKT_RF_t) , (SMB_t) , and (HML_t) are Fama and French (1993) risk factors that are representing the market, firm size, and book-to-market ratio premiums. The Fama and French (1993) risk factors are used by extensive asset pricing literature (Aboody et al., 2005; Francis, et al., 2005; Petkova, 2006; Stambaugh, 2003). The FLEX represents the new return factors that are constructed by going long on financial flexibility characteristics ranked in the bottom two quintiles of returns and shorting the top two quintiles.

Then, under the second step we estimate the factor risk premium γ from cross-sectional regression, with betas on the right-hand side as independent variables and γ as regression coefficients, and v_i are the cross-sectional regression residuals (pricing errors) as shown in equation (3).

$$R_{it} - R_{ft} = \gamma_0 + \gamma_1 b_{i,MKT_RF} + \gamma_2 b_{i,SMB} + \gamma_3 b_{i,HML} + \gamma_4 b_{i,FLEX} + v_{it} \quad (3)$$

Apart from Fama and French's (1993) three risk factors, this research adds a momentum (UMD) factor as given in Ken French's data library and Pastor and Stambaugh's (2003) liquidity factor (LIQ). In this case, the coefficient magnitude on γ_4 determines whether there is a statistically significant premium on the FLEX factor loading or not. If there is a significant premium then the factor is categorised as a risk factor (Core, et al., 2008; Gray and Johnson, 2011); otherwise it would be the characteristics that derive the returns.

In particular, for asset pricing tests based on using a two-stage cross-sectional regressions (2SCSR) approach, three types of portfolios are used as test assets. These are 25 LSIZE-LBTM portfolios; 100 FIN_FLEX portfolios, and 64 LSIZE-LBTM-FIN_FLEX portfolios, where the portfolio returns are value-weighted monthly returns. Kim and Qi (2010) note that the use of portfolios as test assets helps induce variations in returns and alleviates the errors-in-variable bias.

3.2.4.2 Characteristics versus covariance test

Daniel and Titman (1997) argue that characteristics dominate the covariances. For example, the book-to-market ratio characteristic remains a significant determinant for subsequent returns even after controlling for the covariances. The authors prove this by a characteristic versus covariance test that make it possible to distinguish whether characteristics or factor loadings predict returns. DT finds that as expected returns do not positively correlate with the factor loadings after controlling for characteristics; factor loadings do not predict returns. Thus, it is characteristics that predict the returns and this effect is due to investor mispricing of stocks. However, there is wide criticism over DT's approach being biased against covariances. For example, Zhang (2005) and Lin and Zhang (2013) argue that the measurement errors in covariances (betas) are responsible for their demise against the characteristics. The covariances are generally estimated using rolling regressions of 36 or 60 months that induce the time-lag between the covariances and portfolio formation and lead to the reduction of covariances' ability to predict returns. Moreover, Lin and Zhang (2013) argue that the reversal of sorting order in the DT test from first characteristics and then covariances to first covariances and then characteristics, would weaken the characteristics' substantially and relatively enhance the predictive power of covariances.

3.2.4.3 Fama and MacBeth's (1973) cross-sectional regressions

To further investigate the properties of both characteristics and corresponding factor loading and to strengthen arguments in favour of either the risk-based or characteristics-based explanations, this study runs Fama and MacBeth's (1973) cross-sectional

regressions. First of all, it tests both characteristic and factor loading separately to examine whether they significantly predict the subsequent excess returns or not. Second, including both of them together in one regression, helps in observing whether characteristics subsume the factor loading return predictive power (supporting the mispricing doctrine) or factor loading survives the test and remains significantly positive (supporting the risk doctrine). Finally, keeping control variables in place that are the FF3 risk factors and past six month return characteristic BHRET6; the factor loading is estimated through time-series regressions over 60 months or requiring at least 24 months, using a four-factor model that is Fama and French's (1993) three-factor model augmented with the financial flexibility factor (FLEX) (see equation 2).

3.2.4.4 Financial flexibility and information uncertainty

The interpretations about a certain factor's candidature as a risk factor or merely a mispriced characteristic may vary based on the nature of the information environment. The risk doctrine asks for no effect of risk environment on return generating pattern. On the contrary, the mispricing doctrine documents that the variation in return generation is expected with a variation in the quality of the environment. Precisely, the cost of capital increases with the degradation of the information quality (Easley and O'hara, 2004). Consistently, Armstrong et al. (2011) find a positive relationship between firms' cost of capital and information asymmetry. The degradation of information quality also magnifies the effects of cognitive biases that result in misvaluations (Hirshleifer, 2001).

In the case of this study, the greatest reduction (increase) in financial flexibility is coupled with higher (lower) expected returns. If FLEX echoes the risk, this particular trend should persist irrespective of the changes in the information environment. On the other hand, the FLEX premium would amplify (reduce) in an (a) opaque (transparent) information quality environment, if FLEX echoes the mispricing. Current literature echoes mispricing more than risk when information is uncertain. Time information is highly uncertain the longer underreaction to analyst forecast revision and the momentum effect is observed (Zhang, 2006), investor sentiments effect the assets that are difficult to arbitrage and if they are valued subjectively (Baker and Wurgler, 2006). Because, the information uncertainty poses substantial impediments to arbitrage and corrective trades (Hirshleifer, 2001).

To test the information uncertainty effects on FIN_FLEX, this research presents the mean returns and the factor model regression alphas on the portfolios ranked on various information uncertainty measures and the FIN_FLEX. It borrows four information uncertainties from Zhang (2006) and Durate and Young (2009), literature that include inverse of market value of equity (1/MV), inverse of firm age (1/AGE), the probability of informed trading (PIN), and the firm cash flow volatility (CFVOL).

3.3 Descriptive statistics, correlation matrix, and relationship of financial flexibility with prominent factors

3.3.1 Descriptive statistics

Table 3.3 shows the descriptive statistics and the correlations of the prominent Fama-French three factors (MKT_RF, SMB, and HML), the momentum factor (UMD), and

the new latent growth factors (FLEX) along with their characteristics and factor loadings. This essay in Panel A finds that the average monthly premium (mean) of FLEX 0.69% with *t*-statistics of 5.56 is significantly different from zero. This *t*-statistic is considerably higher than the new bar of 3.0 statistics proposed by Harvey et al. (2013). However, the average monthly premium on MKT_RF, SMB, HML, UMD, and LIQ are 0.52%, 0.12%, 0.29%, 0.56%, and 0.51% respectively and, except SMB and HML, other factors are significantly different from zero. This significant premium on FLEX can be attributed to the unobserved nature of the underlying characteristics of the factor⁴⁸.

Panel B of Table 3.3 presents the correlation among the return factors; the Spearman correlations are given in upper diagonal and Pearson in lower diagonal. The FLEX is only significantly correlated with MKT_RF and HML with Pearson (Spearman) correlation coefficients of -0.203 (-0.232) and 0.561 (0.573) respectively. Similarly, among Fama and French's (1993) three factors, there exists a significant Pearson (Spearman) correlation of -0.335 (-0.171) among the SMB and HML factors.

Further, Panel C of Table 3.3 presents the descriptive statistics on the factor loadings (covariances) and the characteristics. The covariances are estimated over the past 60 months for each firm using the four-factor model that is Fama and French's (1993) factor model augmented with FLEX factor requiring a minimum of 24 months data. The descriptive statistics indicate that the standard deviation of FLEX is about 33%, 84%, and 150% larger than the HML, SMB, and MKT_RF respectively. The trend shows the unsteady nature of the FLEX over time.

⁴⁸ According to Fabozzi and Markowitz (2011) the latent (unobserved) factors are preferred to the observed factors because observed can be measured with error and they might have been already anticipated.

Table 3.3**Summary statistics and correlations for factors, factor loadings, and underlying variable**

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FLEX}FLEX_t + \varepsilon_{it}$$

Variable	Mean	t-stat	Std. Dev	Min	Max	
<i>Panel A: Descriptive statistics on factors</i>						
MKT_RF	0.52	1.890	0.047	-0.232	0.125	
SMB	0.12	0.600	0.034	-0.164	0.220	
HML	0.29	1.560	0.032	-0.126	0.138	
UMD	0.56	1.910	0.050	-0.347	0.184	
LIQ	0.51	2.210	0.039	-0.105	0.212	
FLEX	0.69	5.560	0.021	-0.059	0.085	
	MKT_RF	SMB	HML	UMD	LIQ	FLEX
<i>Panel B: Pearson and Spearman correlation matrix between factors</i>						
MKT_RF	1	0.221	-0.321	-0.110	-0.004	
		0.000	<.0001	0.060	0.951	
SMB	0.220	1	-0.171	-0.011	-0.050	
			0.003	0.851	0.392	
HML	-0.299	-0.33598	1	-0.116	-0.034	
				0.048	0.559	
UMD	-0.178	0.056	-0.162	1	0.045	
					0.444	
LIQ	-0.035	-0.048	-0.072	0.028	1	
FLEX	-0.203	-0.020	0.561	-0.081	0.024	
	0.000	0.729	<.0001	0.166	0.686	
Variable	Mean	Std. Dev	Min	Max		
<i>Panel C: Descriptive statistics on factor loadings and characteristics</i>						
b _{MKT}	1.068	0.833	-12.752	12.831		
b _{SMB}	0.875	1.129	-14.748	20.308		
b _{HML}	0.008	1.568	-41.276	24.661		
b _{FLEX}	0.277	2.082	-58.616	61.877		
FIN_FLEX	0.026	0.868	-3.103	9.077		
LSIZE	5.097	2.117	-1.648	13.170		
LBTM	0.498	0.347	0.000	6.318		

Table 3.3 (Continued)

	b_{MKT}	b_{SMB}	b_{HML}	b_{FLEX}	FIN_FLEX	LSIZE	LBTM
<i>Panel D: Pearson and Spearman correlation matrix between factor loadings and characteristics</i>							
b_{MKT}	1	0.028	0.095	0.133	0.021	0.133	-0.060
		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
b_{SMB}	-0.003	1	0.078	0.005	-0.013	-0.229	0.035
	0.015		<.0001	0.000	<.0001	<.0001	<.0001
b_{HML}	0.096	0.057	1	-0.420	-0.030	0.053	0.146
	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001
b_{FLEX}	0.185	0.039	-0.469	1	-0.043	-0.187	0.063
	<.0001	<.0001	<.0001		<.0001	<.0001	<.0001
FIN_FLEX	0.023	0.021	-0.049	-0.002	1	0.112	-0.185
	<.0001	<.0001	<.0001	0.169		<.0001	<.0001
LSIZE	0.099	-0.213	0.051	-0.169	0.002	1	-0.342
	<.0001	<.0001	<.0001	<.0001	0.044		<.0001
LBTM	-0.046	-0.0003	0.109	0.068	-0.147	-0.344	1
	<.0001	0.806	<.0001	<.0001	<.0001	<.0001	
<hr/>							
	LSIZE on b_{SMB}		LBTM on b_{HML}		FIN_FLEX on b_{FLEX}		
<i>Panel E: Univariate Fama-MacBeth regression of characteristics on factor loadings</i>							
	-0.423***		0.026***		-0.001		
	(-45.02)		(14.11)		(-0.65)		

This table in Panel A reports the descriptive statistics of the prominent pricing factors and the FLEX factor. The pricing factors include: three Fama and French (1993) factors, the MKT_RF the premium on the market in excess of risk-free rate; SMB the premium on the size portfolio that is long on small size stock portfolios and long on big stock portfolios; and HML is the premium on the High minus Low book-to-market ratio stock portfolios. Carhart's (1997) momentum factor (UMD); and Pastor and Stambaugh's (2003) liquidity risk factor (LIQ). Panel B reports the Pearson (Spearman) correlation matrix with Lower (Upper) diagonal. Panel C presents the descriptive statistics of the factor loading on the Fama-French three factors given above and factor loading on FLEX factors and characteristics (firm size, book-to-market ratio, and financial flexibility). The loadings are estimated over a 60 month rolling window with minimum requirement of 24 months using the equation in the table header. Panel D provides the Pearson (Spearman) correlation matrix for factor loadings and characteristics. Panel E shows results for the univariate Fama and MacBeth (1973) cross-sectional regressions of characteristics (LSIZE, LBTM, and FIN_FLEX) on their respective factor loading. Panel E reports the coefficient estimate and *t*-statistics. *** for 1%; ** for 5%; and * for 10% level of significance.

Panel D of Table 3.3 presents the correlation matrix on factor loading and associated characteristics. The FIN_FLEX has significant correlations with competing characteristics such as firm size and the book-to-market ratio. The firm size and the book-to-market ratio have significant correlations with their corresponding factor loadings. In contrast, FIN_FLEX has an insignificant and negative correlation with its corresponding factor loading. The negative correlation means that for high financial flexibility there will be a low factor loading, which is consistent with the literature that investment growth, or in this study the latent growth (FIN_FLEX), and subsequent stock return are negatively related. The negative robust FIN_FLEX-return relationship is shown in Table 3.2. This negative relation means high financial flexibility firms will earn lower returns, which in turn implies that factor loading (b_{FLEX}), and returns are positively related⁴⁹. This positive relation between factor loading and the returns would justify any positive premium and leads to the conclusion that it is factor loading that drives the returns supporting the risk-based explanations.

In order to support the above arguments derived from Panel E of Table 3.3, this essay runs the univariate Fama and MacBeth (1973) regressions of characteristics (LSIZE, LBTM, and FIN_FLEX) on the corresponding factor loading. The LSIZE and LBTM are significantly correlated with right signs with their corresponding factor loadings. However, FIN_FLEX has a negative and insignificant relation to corresponding factor loading with coefficient and t -statistics of -0.001 and -0.65.

⁴⁹ The positive relation is confirmed in unreported results.

3.3.2 FLEX: A non-redundant factor

In this subsection, the study shows that the financial flexibility anomaly is independent and is able to explain the return variance without the influences of any known pricing factors. In fact, Table 3.4 in Panel A runs time-series regressions of the FLEX factor on the prominent asset pricing factors such as MKT_RF, SMB, HML, UMD, and LIQ. The intercept is expected to be statistically zero if known risk factors or risk models capture the risk inherent with the FLEX factor; and check the magnitude of adjusted- R^2 to be high. If both conditions are not true, then it confirms the orthogonal nature of the FLEX factor.

Among seven rows of Panel A of Table 3.4, the first row shows the results for the capital asset pricing model (CAPM). The second, third, fourth, and fifth rows indicate the results for the univariate time-series regressions on SMB, HML, UMD, and LIQ separately. The sixth row provides results for Fama and French's (1993) three-factor model. The final row displays the results for FF5 that include the parameter coefficients along with t-values to show the level of significance.

The CAPM model results show that the intercept coefficient is significant, indicating that the magnitude of variations in FLEX, which is unrelated to MKT_RF, carries significant mean. In the rest of the models, FLEX is only related to the HML. However, the intercept in all models is positive and carries significant mean. The adjusted- R^2 is only large enough in the models that include HML. Consequently, it may be argued

that the variation of the FLEX factor cannot be explained by other prominent factors or asset pricing models and behaves as an independent factor.

Table 3.4

Time-series regressions

Pane A: Time-series regression of financial flexibility anomaly (FLEX) on pricing factor models

Variable	Intercept	MKT_RF	SMB	HML	UMD	LIQ	Adj R-Sq
Coeff.	0.007***	-0.093***					0.0381
t-stat	(6.02)	(-3.55)					
Coeff.	0.007***		-0.013				0.0004
t-stat	(5.56)		(-0.35)				
Coeff.	0.006***			0.374***			0.3125
t-stat	(5.62)			(11.58)			
Coeff.	0.007***				-0.035		0.0032
t-stat	(5.68)				(-1.39)		
Coeff.	0.007***					0.013	0.0381
t-stat	(5.45)					(0.40)	
Coeff.	0.006***	-0.030	0.126***	0.405***			0.3439
t-stat	(5.6)	(-1.3)	(3.91)	(11.74)			
Coeff.	0.006***	-0.029	0.129***	0.410***	-0.003	0.041	0.3451
t-stat	(5.27)	(-1.22)	(4.02)	(11.56)	(-0.13)	(1.59)	

Panel B: Time-series regressions of pricing factors on FLEX

		Constant	Slope on FLEX	Adj R-Sq
MKT_RF	<i>Coeff.</i>	0.008***	-0.447***	0.0414
	<i>t-stat</i>	(2.93)	(-3.55)	
SMB	<i>Coeff.</i>	0.001	-0.032	0.0004
	<i>t-stat</i>	(0.68)	(-0.35)	
HML	<i>Coeff.</i>	-0.003*	0.843***	0.3149
	<i>t-stat</i>	(-1.78)	(11.58)	
UMD	<i>Coeff.</i>	0.007**	-0.190	0.0066
	<i>t-stat</i>	(2.24)	(-1.39)	
LIQ	<i>Coeff.</i>	0.005**	0.044	0.0006
	<i>t-stat</i>	(1.97)	(0.40)	

Panel A presents the coefficient estimates and related *t*-statistics for time-series regressions of the FLEX factor on prominent pricing factors such as market premium (MKT_RF), size premium (SMB), value-growth premium (HML), momentum premium (UMD), and liquidity premium (LIQ). The first five rows show the univariate regression coefficients and *t*-statistics and the last two rows show multivariate results for three-factor and five-factor model regressions. In Panel B, similar time-series regressions are run in which Fama-French's three factors are regressed separately on the FLEX factor. For each regression the adjusted R-squared is reported. *** for 1%; ** for 5%; and * for 10% level of significance.

In further tests given in Panel B of Table 3.4, the sides are switched in the time-series regressions and univariate regressions are run of prominent pricing factors on the FLEX factor. Similar to results in Panel A, this study finds evidence that the FLEX factor is only related with MKT_RF and HML. However, the large coefficient of FLEX 0.843 with *t*-statistics of 11.58 and marginally significant intercept of HML regression suggests that FLEX may be able to replace the HML.

3.4 Empirical Results

3.4.1 Is financial flexibility priced in the cross-sectional regressions?

Table 3.5 examines whether FLEX is an important determinant of the excess value-weighted returns on several test asset portfolios using 2SCSR proposed in Cochrane (2001), as the main test approach. The test assets are the 25 LSIZE-LBTM, 64 LSIZE-LBTM-FLEX, and 100 FLEX portfolios. In the second stage results, the time-series means of coefficients and Fama-MacBeth (1973) *t*-statistics for 25 LSIZE-LBTM, 64 LSIZE-LBTM-FLEX, and 100 FLEX portfolios are presented in Panel A, B, and C respectively. The variations in monthly excess value-weighted returns are expected to be explained by the FLEX factor in the presence of the factors of CAPM, FF3, FF4 and the FF5 models. The details of the test are discussed under section 3.2.4.

Table 3.5**Two-pass cross-sectional regressions**

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FLEX}FLEX_t + \varepsilon_{it}$$

$$R_{it} - R_{ft} = \gamma_0 + \gamma_1 b_{i,MKT_RF} + \gamma_2 b_{i,SMB} + \gamma_3 b_{i,HML} + \gamma_4 b_{i,FLEX} + v_{it}$$

Panel A: 25 LSIZE-LBTM sorted portfolios

Model	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8
Intercept	2.586*** (6.47)	2.907*** (6.90)	2.878*** (6.93)	2.895*** (7.04)	1.768*** (3.68)	2.334*** (5.92)	2.557*** (6.11)	2.572*** (6.14)
b _{MKT}	-0.908*** (-3.11)	-1.407*** (-4.44)	-1.528*** (-4.38)	-1.541*** (-4.49)	-0.319 (-0.83)	-0.862*** (-3.00)	-1.186*** (-3.47)	-1.188*** (-3.48)
b _{SMB}		0.112 (0.94)	-0.010 (-0.09)	-0.013 (-0.11)		0.031 (0.29)	-0.026 (-0.24)	-0.035 (-0.32)
b _{HML}		0.261** (2.33)	0.329*** (3.00)	0.333*** (2.94)		0.234** (2.00)	0.297** (2.50)	0.306** (2.54)
b _{UMD}			-0.743* (-1.76)	-0.755* (-1.82)			-0.572 (-1.43)	-0.587 (-1.47)
b _{LIQ}				-0.054 (-0.16)				-0.152 (-0.44)
b _{FLEX}					0.360*** (3.77)	0.309*** (3.08)	0.260*** (2.63)	0.264*** (2.66)
Adj R ²	0.15	0.45	0.48	0.49	0.29	0.50	0.52	0.53

Panel B: 100 FIN_FLEX sorted portfolios

Model	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8
Intercept	1.182*** (3.50)	0.833*** (2.63)	0.700** (2.23)	0.695** (2.23)	0.839*** (2.59)	0.853*** (2.68)	0.762** (2.41)	0.766** (2.45)
b _{MKT}	-0.333 (-1.34)	-0.053 (-0.23)	0.114 (0.50)	0.117 (0.52)	-0.040 (-0.17)	-0.048 (-0.21)	0.062 (0.27)	0.059 (0.26)
b _{SMB}		0.095 (0.70)	0.098 (0.73)	0.094 (0.69)		-0.018 (-0.14)	0.002 (0.02)	0.004 (0.03)
b _{HML}		0.285** (2.28)	0.286** (2.28)	0.281** (2.23)		0.145 (1.17)	0.168 (1.35)	0.169 (1.36)
b _{UMD}			0.219 (1.02)	0.217 (1.01)			0.149 (0.71)	0.149 (0.71)
b _{LIQ}				-0.003 (-0.01)				-0.028 (-0.12)
b _{FLEX}					0.210*** (2.64)	0.218*** (2.68)	0.200** (2.50)	0.201** (2.50)
Adj R ²	0.09	0.13	0.13	0.14	0.12	0.14	0.14	0.15

Table 3.5 (Continued)*Panel C: 64 LSIZE-LBTM-FIN_FLEX sorted portfolios*

Model	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8
Intercept	1.90*** (4.53)	2.229*** (5.34)	2.262*** (5.31)	2.226*** (5.37)	0.886* (1.93)	1.528*** (4.32)	1.550*** (4.09)	1.512*** (4.09)
b _{MKT}	-0.870* (-1.73)	-1.519*** (-3.13)	-1.804*** (-3.29)	-1.78*** (-3.29)	-0.227 (-0.41)	-0.802* (-1.84)	-0.860* (-1.75)	-0.838* (-1.72)
b _{SMB}		0.323 (1.29)	0.080 (0.35)	0.097 (0.42)		0.024 (0.10)	-0.004 (-0.02)	0.014 (0.06)
b _{HML}		0.253 (1.09)	0.410* (1.85)	0.393* (1.74)		0.238 (1.02)	0.261 (1.16)	0.243 (1.05)
b _{UMD}			-1.521** (-2.19)	-1.441** (-2.14)			-0.158 (-0.25)	-0.073 (-0.12)
b _{LIQ}				0.304 (0.49)				0.404 (0.64)
b _{FLEX}					0.666*** (4.43)	0.675*** (4.90)	0.671*** (4.91)	0.675*** (4.91)
Adj R2	0.10	0.24	0.25	0.27	0.19	0.28	0.28	0.31

This table reports the results of two-stage cross-sectional regressions for asset pricing tests on three test assets. Panel A provides results for 25 LSIZE-LBTM portfolios for the full sample. The 25 LSIZE-LBTM portfolios are constructed at the start of each month, by sorting and ranking into both LSIZE and LBTM quintiles independently. The intersection of both sets of controls gives 25 portfolios. Panel B provides results for 100 FIN_FLEX portfolios for the full sample. The 100 portfolios for FIN_FLEX are constructed at the start of each month, by sorting and ranking into 100 FIN_FLEX percentiles independently. Panel C provides results for 64 LSIZE-LBTM-FIN_FLEX portfolios. The 64 portfolios are constructed at the start of each month, by sorting and ranking firms into 4 LIZE, 4 BTM, and 4 FIN_FLEX quartiles and their intersections as 64 portfolios separately. The portfolio returns are monthly excess value-weighted returns. The first stage estimates factor betas from a time-series regression for each firm using the corresponding model. In the second stage, estimated betas are used as explanatory variables in a cross-sectional regression for each month. For each model's intercept and explanatory variables, time-series means and Fama-MacBeth (1973) *t*-statistics of the coefficients of factor loadings are reported. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

The tests are designed to determine whether FLEX is priced and bears a positive significant premium in the presence of the controls discussed above. Panel A, B, and C of Table 3.5 present the results and convey the message that FLEX bears positive and significant premium and thus is a priced factor. Strikingly, the prominent factor model

controls do not survive the addition of the FLEX factor. For example, results from the first four models of the three panels show that only MKT_RF and HML bear significant premiums. However, in a more conservative test portfolio given in panel C, when these four models are augmented with FLEX, none of the five (MKT_RF, SMB, HML, UMD, and LIQ) factor premiums remains significant. Only FLEX survives the test and behaves as a priced factor, which indicates that the risk inherent in the control factors is captured by the FLEX factor.

This essay interprets the above evidence as being that FLEX is a priced risk factor. However, in view of existing literature (Prombutr et al., 2012), there is the chance that the FLEX premium could be the result of mispricing of FIN_FLEX. Because FIN_FLEX is estimated as the simultaneous growth in two consecutive previous fiscal year ends, that is 12 months earlier than the stock returns, it may lead to investor underreaction to the arrival of new financial flexibility information content. To resolve and unravel this problem, Daniel and Titman (1997) propose the characteristic vs. covariance test, which is discussed in section 3.2.4.1 and empirical evidence is provided in the next sub-section.

3.4.2 Do covariances dominate characteristics?

This essay in Table 3.3 argues that there is no significant positive correlation between the FIN_FLEX and corresponding factor loading. Thus, there remains no obligation to distinguish characteristics from covariances. However, such a test shall strengthen the arguments of the risk-based explanations.

3.4.2.1 Portfolio evidence based on Daniel and Titman's (1997) test

The test results are given in Table 3.6; test construction uses US common stock listed on the NYSE, AMEX, and NASDAQ for the period of July, 1986 to December, 2010 excluding the regulated utilities firm and financial firm. This essay uses the triple sort methodology. First firms are sorted each month into three LSIZE terciles and three FIN_FLEX terciles independently. Then for the nine intersections of LSIZE-FIN_FLEX, firms are further sorted into quintiles based on b_{FLEX} creating 45 LSIZE-FIN_FLEX- b_{FLEX} portfolios. The factor loading for each firm is obtained by running the four-factor model, that is FF3 augmented with FLEX factor, over a rolling window of the previous 60 months of the portfolio formation date or requiring at least 24 months' data. The value-weighted returns are obtained for 45 portfolios each month and their averages in excess of the risk-free rate are reported in Panel A of Table 3.6. The last row of the panel shows the average excess returns when controlling for FIN_FLEX and the last column shows the average excess return when controlling for FLEX loading. The relationship between b_{FLEX} and returns should be monotonic positive (shown in the last row), if the covariances dominate the characteristics. Interestingly, the relation is non-monotonic and neither positive nor negative, but a nonlinear that shows decreasing order first and increasing later making a U-shape. However on average returns on High b_{FLEX} are higher than Low b_{FLEX} with a positive spread of 0.054. And, out of nine High b_{FLEX} portfolios four have larger returns than Low b_{FLEX} portfolios.

Table 3.6

Characteristics versus covariances tests

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FLEX}FLEX_t + \varepsilon_{it}$$

Panel A: Mean excess returns of 45 LSIZE-FIN_FLEX- bFLEX portfolios

Characteristics			b _{FLEX} Quintiles: Mean excess returns					
LSIZE	FIN_FLEX	LSIZE/FIN_FLEX	1	2	3	4	5	Average
0	0	S/H	1.611	1.816	1.419	1.831	2.547	1.845
0	1	S/M	1.787	1.377	1.041	1.369	1.828	1.481
0	2	S/L	1.257	1.086	0.991	1.351	0.892	1.115
1	0	M/H	1.002	0.834	0.635	0.816	0.984	0.854
1	1	M/M	1.055	0.613	0.645	0.870	0.834	0.804
1	2	M/L	0.738	0.675	0.615	0.862	0.324	0.643
2	0	B/H	1.102	0.945	0.675	0.586	0.729	0.807
2	1	B/M	0.372	0.676	0.706	0.505	0.850	0.622
2	2	B/L	0.250	0.423	0.243	0.845	0.672	0.486
Average			1.019	0.938	0.774	1.004	1.073	
			5.976	6.488	6.958	7.051	4.716	

Panel B: FF3+FLEX alphas of 45 LSIZE-FIN_FLEX- bFLEX portfolios

Characteristics			b _{FLEX} Quintiles: FF3+FLEX Intercept					t(Intercept)				
LSIZE	FIN_FLEX	LSIZE/FIN_FLEX	1	2	3	4	5	1	2	3	4	5
0	0	S/H	-0.014	0.247	-0.052	0.172	0.485	-0.05	0.93	-0.19	0.61	1.21
0	1	S/M	0.696	0.467	0.199	0.205	0.350	2.71	2.14	0.94	0.80	0.97
0	2	S/L	0.297	0.161	0.077	0.634	-0.218	1.07	0.67	0.29	1.92	-0.62
1	0	M/H	-0.309	-0.091	-0.464	-0.459	-0.569	-1.30	-0.53	-2.53	-2.25	-1.91
1	1	M/M	0.182	-0.194	-0.074	0.070	-0.096	0.85	-1.24	-0.46	0.42	-0.48
1	2	M/L	0.096	0.140	0.106	0.298	-0.307	0.45	0.90	0.66	1.81	-1.31
2	0	B/H	0.272	0.171	0.045	-0.146	-0.215	1.10	0.82	0.22	-0.81	-0.98
2	1	B/M	-0.090	0.283	0.238	-0.074	0.061	-0.39	1.66	1.70	-0.55	0.37
2	2	B/L	0.014	0.067	-0.088	0.352	0.149	0.05	0.33	-0.49	2.18	0.62
Average			0.202	0.190	0.165	0.104	-0.206					

Table 3.6 (Continued)

Panel C: Characteristics-balanced portfolios

Characteristics			Characteristics-balanced portfolios: Mean returns and Time-series regression												
LSIZE	FIN_FLEX	LSIZE/FIN_FLEX	Avg Ret	t(Avg Ret)	Intercept	Mkt_RF	SMB	HML	FLEX	tIntercept	tMkt_RF	tSMB	tHML	tFLEX	AdjR2
0	0	S/H	0.476	1.89	0.212	0.079	0.361	-0.360	0.357	0.88	1.47	4.98	-3.92	2.76	0.21
0	1	S/M	0.016	0.07	-0.305	0.088	0.136	-0.251	0.439	-1.31	1.71	1.95	-2.83	3.53	0.09
0	2	S/L	-0.050	-0.20	-0.021	0.019	0.218	-0.364	0.031	-0.08	0.35	2.91	-3.83	0.23	0.14
1	0	M/H	-0.017	-0.07	-0.314	0.178	0.261	-0.289	0.324	-1.38	3.53	3.82	-3.33	2.66	0.19
1	1	M/M	0.018	0.11	-0.007	0.055	0.191	-0.014	-0.044	-0.04	1.48	3.77	-0.21	-0.49	0.07
1	2	M/L	-0.113	-0.69	-0.123	0.023	0.235	-0.104	-0.019	-0.74	0.63	4.73	-1.64	-0.21	0.13
2	0	B/H	-0.366	-1.76	-0.402	0.026	-0.039	-0.022	0.047	-1.78	0.52	-0.58	-0.26	0.39	-0.01
2	1	B/M	0.154	0.79	-0.103	0.014	-0.091	0.007	0.362	-0.51	0.31	-1.50	0.10	3.34	0.06
2	2	B/L	0.422	1.86	0.210	-0.035	-0.289	-0.038	0.397	0.91	-0.68	-4.16	-0.43	3.21	0.11
Average			0.060	0.500	-0.095	0.050	0.109	-0.159	0.210	-0.80	1.89	3.06	-3.52	3.31	0.14

Panel C: Alternate characteristics-balanced portfolios

Characteristics			Alternate Characteristics-balanced portfolios: Mean returns and Time-series regression												
LSIZE	FIN_FLEX	LSIZE/FIN_FLEX	Avg Ret	t(Avg Ret)	Intercept	Mkt_RF	SMB	HML	FLEX	tIntercept	tMkt_RF	tSMB	tHML	tFLEX	AdjR2
0	0	S/H	0.936	2.50	0.499	0.126	0.593	-0.494	0.558	1.41	1.61	5.55	-3.65	2.93	0.23
0	1	S/M	0.041	0.12	-0.347	0.098	0.159	-0.330	0.548	-1.01	1.28	1.54	-2.52	2.97	0.06
0	2	S/L	-0.364	-1.02	-0.515	0.063	0.224	-0.554	0.315	-1.42	0.78	2.05	-4.00	1.62	0.11
1	0	M/H	-0.018	-0.05	-0.260	0.188	0.379	-0.350	0.239	-0.79	2.59	3.84	-2.80	1.36	0.16
1	1	M/M	-0.221	-0.87	-0.279	0.063	0.299	-0.047	-0.017	-1.05	1.07	3.74	-0.47	-0.12	0.07
1	2	M/L	-0.414	-1.63	-0.403	0.032	0.396	-0.176	-0.063	-1.60	0.58	5.21	-1.83	-0.47	0.15
2	0	B/H	-0.373	-1.23	-0.486	0.033	-0.141	-0.020	0.171	-1.48	0.46	-1.43	-0.16	0.97	0.00
2	1	B/M	0.478	1.66	0.150	-0.059	-0.089	0.023	0.507	0.50	-0.88	-0.98	0.20	3.15	0.06
2	2	B/L	0.422	1.40	0.136	0.023	-0.173	-0.058	0.438	0.42	0.33	-1.80	-0.47	2.55	0.03
Average			0.054	0.32	-0.167	0.063	0.183	-0.223	0.299	-1.01	1.73	3.68	-3.54	3.39	0.16

This table reports the results for characteristics versus covariances test on 45 LSIZE-FIN_FLEX- b_{FLEX} sorted portfolios. The data used is for US common stocks listed on NYSE, AMEX, and NASDAQ excluding the regulated and financial firms for the period 1986 to 2010. The b_{FLEX} is estimated over a 60 month rolling window with a minimum requirement of 24 months using time-series regressions. First, the stocks are sorted into terciles separately on LSIZE and FIN_FLEX; their intersection gives 9 LSIZE-FIN_FLEX portfolios. Then each 9 portfolios are ranked into quintiles making 45 portfolios containing value-weighted returns. Panel A reports the mean excess value-weighted returns for 45 LSIZE-FIN_FLEX- b_{FLEX} portfolios. Panel B reports the intercept coefficients and t -statistics of time-series regression for 45 LSIZE-FIN_FLEX- b_{FLEX} portfolios using a four-factor model that is Fama-French's three-factor model augmented with FLEX factor. Panel C reports the average returns and their t -statistics for the 9 characteristics balanced (CB) portfolios. The CB portfolios are determined by going long top two b_{FLEX} quintiles and shorting bottom two b_{FLEX} quintiles within each of 9 LSIZE-FIN_FLEX groups. Panel C also reports the four factor time-series regression coefficients and t -statistics along with adjusted R-squared. The bottom of the panel reports the mean return and regression results for a single portfolio formed by equally weighting the 9 CB portfolios. Similarly, alternative 9 CB portfolios are constructed, which go long top one b_{FLEX} quintiles and shorting bottom one b_{FLEX} quintiles within each of 9 LSIZE-FIN_FLEX groups in Panel D. Statistics similar to Panel C for Panel D are obtained.

If FLEX is a priced factor, then it should capture all variations in returns and the regression intercepts should be jointly zero and show no pattern with b_{FLEX} . The intercepts are reported in Panel B of Table 3.6. The results largely favour the risk doctrine as 35 out of 45 portfolio intercepts are insignificant and not different from zero. And out of 10 significant intercepts, seven are positive and three are negative. Among nine Low b_{FLEX} intercepts six are positive with only one significantly different from zero and not all nine Low b_{FLEX} intercepts are higher than High b_{FLEX} intercepts.

Further, under characteristics-balanced (CB) portfolios, this study constructs two types of nine zero-cost CB portfolios. For the first CB portfolio, it longs two top FLEX loadings and shorts two bottom FLEX loadings; and for the second CB it longs top one FLEX loading and shorts bottom one FLEX loading. It further produces the mean returns and t -statistics for these nine CB portfolios and runs time-series regressions of these two CB portfolio returns on a four-factor model that is FF3 augmented with the FLEX factor. The results are reported under Panel C and D of Table 3.6. The average returns on the spread of the long and short portfolio are positive in five out of nine portfolios in Panel C and on four out of nine portfolios in Panel D. In each panel, out of nine portfolios two portfolios bear significant positive returns. The average of these average returns of nine portfolios in each panel is again positive though insignificant. The insignificant average could be because of the weak negative relationship between FIN_FLEX and its factor loading.

Next, the intercepts from the time-series regressions are expected to be jointly zero, if FLEX is a priced risk factor and negative otherwise. The insignificant intercepts in

Panels C and D reflect that the four-factor model is adequate in explaining the CB portfolio returns. The coefficients on average are positive except that of HML, confirming that the strategy of a characteristics-balanced spread portfolio on FLEX factor loading can result in positive average returns. Accordingly, evidence suggests that the FLEX factor loading is capable of deriving returns and is consistent with the risk-based explanations.

3.4.2.2 Cross-sectional regressions evidence

Table 3.7 reports the results of the Fama and MacBeth (1973) cross-sectional regression of subsequent excess returns on the characteristics and the factor loadings of the financial flexibility while controlling for FF3 factor model and the past six month return characteristic BHRET6. The risk-based explanation demands that the characteristics control should not subsume the premium or coefficient on the factor loading; and mispricing-based explanations demand that the characteristics should be able to subsume the explanatory power of the loading in cross-sectional regressions. Evidence suggests that in regressions that include both characteristics (FIN_FLEX) and loading (b_{FLEX}), the coefficient of loading 0.075 is significantly different from zero at the 10% level with t -statistics of 1.84. The control of characteristics (FIN_FLEX) and the control of Fama-French's three-factor model and momentum are unable to subsume the explanatory power of b_{FLEX} .

3.4.2.3 Information uncertainty and financial flexibility

Table 3.8 using the information environment assesses whether the return predictability of the financial flexibility (FIN_FLEX) information content is due to the risk or mispricing. If conjecture of the risk is true, then the low FIN_FLEX would keep earning

Table 3.7**Fama-MacBeth cross-sectional regressions**

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FLEX}FLEX_t + \varepsilon_{it}$$

Panel A: Fama-MacBeth cross-sectional regressions (Testing FLEX)

Model	M-1	M-2	M-3	M-4	M-5	M-6
Intercept	1.113*** (3.06)	1.088*** (2.91)	1.052*** (2.85)	0.778*** (3.34)	0.708*** (3.07)	0.664*** (2.89)
b_{MKT}				0.192 (1.20)	0.206 (1.29)	0.127 (0.92)
b_{SMB}				0.093 (0.96)	0.103 (1.07)	0.078 (0.83)
b_{HML}				0.097 (0.84)	0.077 (0.67)	0.057 (0.55)
b_{FLEX}	0.072* (1.75)		0.075* (1.84)	0.092* (1.84)	0.084* (1.69)	0.085* (1.78)
FIN_FLEX		-0.612*** (-4.44)	-0.592*** (-4.36)		-0.499*** (-4.47)	-0.490*** (-4.62)
BHRET6						-0.249 (-0.91)
Adj-R ²	0.004	0.002	0.01	0.03	0.03	0.03

In this table, results are reported for the Fama-MacBeth (1973) cross-sectional regressions. Firm excess returns are regressed each month on FLEX factor and control variables. The control variables are the Fama-French (1993) three risk factors. MKTRF is the excess return on the market portfolio. SMB is the return to size factor-mimicking portfolio. HML is the return to book-to-market factor mimicking portfolio. FLEX is the Long (Short) portfolio of Low (High) financial flexibility stock portfolios. The factor loading is estimated through time-series regressions over a 60 month rolling window with at least 24 months requirement of the form given in the table header. For each model's intercept and explanatory variables, time-series means and Fama-MacBeth (1973) t -statistics of the coefficients of factor loadings are reported. *, **, and *** represent the model's significance at 10%, 5%, and 1% respectively.

higher future returns than the high FIN_FLEX firms, irrelevant of the information environment. In the literature, the information environment is an important determining factor of investor misvaluation (e.g., Hirshleifer 2001; Zhang 2006). The misvaluation occurs when the information is being processed inefficiently or sub-optimally due to the limited attentions or cognitive biases.

This section uses four proxies of information uncertainty, for example, the inverse of the market size ($1/MV$), inverse of firm age ($1/AGE$), probability of informed trading (PIN), and the cash flow volatility (CFVOL). MV is the product of the fiscal year-end closing price and the shares outstanding. AGE is the number of years since the firm first appeared in the CRSP database. The PIN is obtained as in Durate and Young (2009). CFVOL is the standard deviation of the five-year operating cash flow scaled by the total assets, which requires at least three years data.

Panel A of Table 3.8 shows that controlling for various information uncertainty proxies, the financial flexibility consistently earns high returns on low FIN_FLEX quintile and low returns on high FIN_FLEX quintile. The hedge FIN_FLEX portfolio (FLEX5-FLEX1) average returns decrease from insignificant -0.28% for lowest information (U1) quartile to significant -2.09% ($t=8.79$) for highest information uncertainty (U4) quartile. Similarly, the information uncertainty hedge (U4-U1) shows a monotonic decrease of mean hedge returns of 1.99% ($t=3.1$) for FLEX1 and 0.18% ($t=0.27$) for FLEX5. Consequently, this evidence suggests that investors do not misvalue the financial flexibility related information. Moreover, under Panels B and C of Table 3.8, the alpha value of hedge portfolios (FLEX5-FLEX1) in high information uncertainty (U4) portfolio are always low across different models (i.e., FF3, FF4, and FF5) as compared to the low information uncertainty (U1) portfolio. Thus, the evidence indicates no role for information uncertainty in FIN_FLEX's return predictability, which supports the risk-based explanation of the FLEX factor.

Table 3.8**Information uncertainty and financial flexibility**

$$R_{it} - R_{ft} = \alpha + \beta_{i,MKT_RF}MKT_RF_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,UMD}UMD_t + \beta_{i,FLEX}LIQ_t + \varepsilon_{it}$$

Panel A: Mean returns for FLEX and 1/MV sorted portfolios

	1/MV	FLEX1	FLEX2	FLEX3	FLEX4	FLEX5	FLEX5-FLEX1
U1 (low)	1.09	1.10	0.87	0.82	0.81	-0.276 (-1.00)	
U2	1.18	1.02	1.05	1.08	0.34	-0.843 (3.37)	
U3	1.69	1.35	1.16	1.21	0.60	-1.091 (-4.66)	
U4 (high)	3.08	1.93	1.55	1.50	0.99	-2.087 (-8.79)	
U4-U1	1.993 (3.18)	0.825 (1.73)	0.677 (1.49)	0.688 (1.44)	0.182 (0.27)	-1.811 (-4.97)	

Panel B: Factor model intercepts for FLEX and 1/MV sorted portfolios

1/MV	FF3 Alpha					FF4 Alpha		FF5 Alpha
	FLEX1	FLEX2	FLEX3	FLEX4	FLEX5	FLEX1-FLEX5	FLEX1-FLEX5	FLEX1-FLEX5
U1 (low)	0.12	0.23	0.06	-0.01	0.06	0.28	0.30	0.35
t-value	0.89	2.55	0.79	-0.11	0.38	1.39	1.46	1.68
U2	-0.02	-0.15	-0.07	0.00	-0.63	-0.27	-0.23	-0.20
t-value	-0.11	-1.21	-0.54	0.04	-3.59	-1.25	-1.04	-0.90
U3	0.55	0.15	0.01	0.14	-0.41	-0.62	-0.60	-0.58
t-value	1.99	0.91	0.05	1.05	-2.13	-2.77	-2.65	-2.55
U4 (high)	2.07	0.90	0.50	0.55	0.07	-1.66	-1.64	-1.67
t-value	5.77	3.79	2.49	2.75	0.25	-7.09	-6.91	-6.98
U4-U1	1.95	0.67	0.44	0.56	0.01	-1.94	-1.94	-2.02

Panel C: Mean return and factor model alphas for other information uncertainty proxies.

1/AGE	Mean Ret	FF3 Alpha		FF4 Alpha	FF5 Alpha
	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1
U1 (low)	-0.118	0.41	0.40	0.46	
t-value	-0.37	1.60	1.52	1.75	
U2	-0.445	0.03	0.10	0.16	
t-value	-1.11	0.09	0.26	0.44	
U3	-0.919	-0.40	-0.30	-0.23	
t-value	-2.2	-1.11	-0.82	-0.63	
U4 (high)	-0.997	-0.59	-0.60	-0.54	
t-value	-1.55	-0.91	-0.91	-0.81	
U4-U1	-0.878 (-1.32)	-1.00	-0.99	-1.00	

PIN	Mean Ret	FF3 Alpha		FF4 Alpha	FF5 Alpha
	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1
U1 (low)	-0.398	0.05	0.35	0.30	
t-value	-1.12	0.14	1.12	0.95	
U2	-0.766	-0.36	-0.34	-0.36	
t-value	-2.22	-1.06	-0.97	-1.03	
U3	-0.559	0.09	0.09	0.15	
t-value	-1.29	0.21	0.22	0.35	
U4 (high)	-0.756	-0.04	-0.31	-0.45	
t-value	-1.61	-0.10	-0.66	-0.97	
U4-U1	-0.356 (-0.61)	-0.09	-0.66	-0.75	

Table 3.8 (Continued)

CFVOL	Mean Ret	FF3 Alpha	FF4 Alpha	FF5 Alpha
	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1	FLEX5-FLEX1
U1 (low)	-0.23	0.30	0.25	0.33
t-value	-0.67	1.02	0.85	1.12
U2	-0.44	0.13	0.12	0.13
t-value	-1.19	0.43	0.37	0.41
U3	-0.15	0.35	0.41	0.52
t-value	-0.43	1.05	1.23	1.53
U4 (high)	-0.56	0.005	0.11	0.22
t-value	-1.33	0.01	0.29	0.59
U4-U1	-0.323 (-0.60)	-0.29	-0.14	-0.11

This table reports the mean returns and the factor model regressions alphas on the portfolios sorted by the measures of information uncertainty and financial flexibility (FIN_FLEX). First the stocks are sorted in five FIN_FLEX quintiles in each month and later each FIN_FLEX quintile is sorted into four quartiles based on the information uncertainty measures. Four information uncertainty proxies have been used, which include: 1/MV, 1/AGE, PIN, and CFVOL. 1/MV is the inverse of the market value, which is the product of the closing price at fiscal year-end and the number of shares outstanding. 1/AGE is the inverse of the firm's age, which is the number of years since the firm first appeared in CRSP data. PIN is the probability of informed trading from Durate and Young (2009). CFVOL represents cash flow volatility, which is the standard deviation of cash flow from operations scaled by the average total assets in the past 60 months (requiring at least 36 months' data). In panel A, the table reports the mean returns for 20 1/MV and FIN_FLEX portfolios, and four hedge portfolios (FLEX5-FLEX1). Panel B runs factor model time-series regressions of form given in the table header for the Fama-French (FF3) three-factor model, four-factor model (FF3+UMD), and the five-factor model (FF4+LIQ). Panel C details mean returns and the factor model alphas for the rest of the proxies of information uncertainty, but only hedge portfolio results (FLEX5-FLEX1) are shown. The returns are value-weighted of each firm within portfolios that are shown in the table along with the *t*-statistics.

3.5 Conclusions

This study investigates the question of whether the financial flexibility anomaly (estimated as a return spread between low and high financial flexibility characteristics based stock portfolios) carries a significant positive premium. Is such a premium characteristics driven or risk driven? And, if it is a risk factor does it survive when regressed on the prominent pricing factor models that are CAPM, Fama and French's (1993) three-factor model, Carhart's (1997) four-factor model, and Pastor and Stambaugh's (2003) five-factor model?

This study suggests that financial flexibility is a priced risk factor. This evidence contradicts the literature on observed investment growth based anomalies that mainly is in support of mispricing or characteristics based explanations. This research finds a significant positive premium on financial flexibility factor (FLEX) and it is the only factor that survives in the 2SCSR framework proposed by Cochrane (2001). In such regressions, the significant premiums on the MKT_RF and HML are subsumed by the addition of the FLEX factor. Preliminary correlations and time-series regression results show that FLEX is related to MKT_RF and HML. There are also strong indications that HML can be replaced by FLEX, and not otherwise.

The arguments supporting covariances are further strengthened by the evidence that there is a negative Pearson (Spearman) correlation between FIN_FLEX and b_{FLEX} . Then, DT's characteristic versus covariance tests support covariances. Next, in Fama and MacBeth's (1973) cross-sectional regression framework, the characteristic (FIN_FLEX) was found to be unable to capture the explanatory power of the b_{FLEX} . Finally, FIN_FLEX return predictability and FLEX premium are irrelevant to the information environment.

This study argues that investors react rationally to the arrival of information relating to financial flexibility. This is because of the perception that the financial flexibility decision is the most important decision in the context of capital structure. For example, the majority of CFOs⁵⁰ will take it rationally and for the greater good of the firm. Thus, the FLEX factor may capture the manager's unobserved perceptions of future growth opportunities and in turn the risk inherent in firms' cash flows. Further research could focus

⁵⁰ See, Graham and Harvey (2001).

on the relation between the value (HML) and financial flexibility (FLEX) anomalies. From the point of view of the research, the part of the book-to-market ratio that is priced is driven by underlying financial flexibility information content.

CHAPTER FOUR

ESSAY THREE

LATENT ACCOUNTING GROWTH, FIRM SIZE, AND FIRM GROUPING DYNAMICS

This chapter investigates how firms differ in terms of basic corporate financing and investment decisions (Latent Factors) and firm size. This investigation applies a distinct method of grouping a cross-section of firms into 12 clusters. A brief overview of the study is presented in Section 4.1. Data sample, variable estimation, and research methods are given in Section 4.2. The clustering results and other empirical analysis results are discussed in Section 4.3. Section 4.4 concludes the study. The chapter's references are presented in the bibliography section of this thesis.

Latent accounting growth, firm size and firm grouping dynamics

Abstract

This study provides a classification of US common stocks based on common corporate financing decisions and firm size. The firms are clustered into 12 possibly optimal internally homogeneous and externally heterogeneous clusters. The cluster differences show prominently non-linear relationships among firm size and age; firm size and book-to-market ratio. Further, one can always find some special features (differences to other clusters) in every cluster. Finally, the study confirms that across the 12 cluster models, the clustering creation is a robust predictor of future firm value, cash flows and earnings. The result suggests that the simple linear regression for all firms is not acceptable in most applications.

JEL classification: C38, G12, G32

Key words: Cluster analysis, accounting growth, factor analysis, corporate finance policies, business prediction model

An intelligent being cannot treat every object it sees as a unique entity unlike anything else in the universe. It has to put objects in categories so that it may apply its hard-won knowledge about similar objects encountered in the past to the object at hand.

- S. Pinker

4.1 Introduction

The first essay identified a pattern of five unobserved, orthogonal, and forward-looking systematic factors⁵¹ in firms' balance sheets. These factors – financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-term investment (LT_INV), convertible debt usage (CVT_DEBT), and preferred stock usage (PSK_USE) – are obtained using the correlation structure of the balance sheet information content. These factors are common among businesses and mimic management perceptions or actions about firms' growth opportunities. This particular feature enables the new factors in forecasting stock returns and creating firm values. The stock returns comove when ranked on five factors because of differences in growth opportunities in the portfolio evidence. These empirical patterns suggest the role of the five factors in firm categorisation. Thus, this study asks how firms agree and differ in corporate financing decisions. It finds that US listed firms are clustered into 12 natural groups that are internally homogeneous and heterogeneous externally. This study further seeks to understand whether latent factors mimic manager's perceptions; this classification would show optimal differentiation in growth opportunities as well. Again, it would make sense to know how such a firm grouping differs from the existing SIC code based industry classifications. And if it is so, does it have any advantages in firm analysis and evaluation? Particularly, this essay examines the performance of clustering criteria in explaining future firm value, cash flows, and earnings.

⁵¹ These factors are growth variables and are interchangeably named as shifts or corporate finance decisions or factors.

This study employs the cluster analysis technique⁵², as a grouping method to provide sufficient differentiation within a universe of US common stocks. Cluster analysis is a nonlinear technique that is sometimes called classification pattern recognition or data reduction technique (Jobson, 1992b). The clustering criteria in this essay include five latent factors and firm size. The latent factors are change variables and may not be related to firm size. The first essay demonstrates that the new factors and firm size have a non-monotonic relationship. This is also anticipated because the input data for the factor analysis are financial balance change ratios – no size measures are included. Numerous empirical results show that size is an important variable in describing return rates (Fama and French, 1992). Thus, the addition of the firm size as a sixth firm grouping criteria makes the analysis credible. Cluster analysis classifies firms into several groups that have similar profiles within the group and heterogeneous profiles between the groups. In particular, cluster analysis finds natural groups⁵³ where the population is unknown based on some internal criterion instead of finding convenient groups (Jensen, 1971; Lusardi and Tufano, 2009). The choice of clustering criteria and number of clusters requires the use of statistics and judgment to ensure that the clustering makes sense and is correct (Lusardi and Tufano, 2009)⁵⁴.

⁵² This essay uses a K-means clustering method, which is implemented using FASTCLUS procedure in SAS. The K-means clustering algorithm selects initial cluster centroids (means of cluster centres), and starts to assign the stocks to their nearest cluster centroid. Cluster centroid is re-estimated with every new entry. This process is continued until no stock changes any cluster membership, meaning till optimal differentiation between clusters and optimal similarity within clusters is achieved.

⁵³ Jobson (1992b) states that natural clusters are anticipated to satisfy the features of external isolation and internal cohesion.

⁵⁴ The number of clusters is specified based on the values of the cubic clustering criteria and pseudo F statistics. Higher values with no sudden jumps are required for both statistics, which can be observed if one plots these values across the expected number of clusters (two to sixteen). The cluster output is also checked to finalise the decision of selecting cluster number 12 as an optimal number of clusters.

Using the technique of clustering analysis in firm classification is based upon the general theory of business operating and financing. The first essay tries to construct a multivariate measurement system to describe business activities, and the result provides five common decision measures. However, the relation of these measures to operating efficiency and profitability is not necessarily a simple linear type. For example, it may not be easily concluded that the higher the cash used for asset growth, the lower the future profitability or market return (see Tables 2.6 and 2.7) because the real mechanism of growth could be very different. It could be caused by either a greater growing opportunity or simply a product restructuring activity. To separate these two, some other measurement needs to be put together with the growth variable. Clustering is one of the techniques that can classify systematically individuals into groups based upon their overall measures.

The first essay tests the predictability of the five latent factors to the operating performance and market value. This chapter tries a very different assumption: the market value and operating performance is not a common linear relation to all firms, so is not testable by normal regression techniques. There is no such function as follows for all styles of firms (as equation 2.3 reproduced here):

$$Q_{it} = \alpha + \beta_1 LatentGrowth_{i,t-n} + \beta_2 Controls_{i,t} + \varepsilon_{it} \quad (4.1)$$

In contrast to the linear models, the quadratic-form and mixed-effect regression may be not enough either. They try to take care of some of the nonlinearity, but may not be able to justify complete heterogeneity. The clustering technique does not have such types of assumptions; it simply asks what are the common patterns chosen by the managers for

these firms. If some common practices of the profession exist, there could be some optimum pursued by the managers. Based upon this understanding, we can identify or partition firms in their respective clusters of style. Because of the nature of cluster analysis – it does not require any *a priori* hypothesis – it might be a valuable and important contribution to the methodology of business analysis and the result could be complementary to the common linear investigation.

Clustering techniques are not completely new in financial analysis. In finance literature, Elton and Gruber (1970, 1971) are among the first researchers to propose heuristic approaches for partitioning stock into homogeneous groups. Later, Goetzmann and Watcher (1995) employ factor analysis and K-means cluster analysis for classification of real estate portfolio diversification. Brown and Goetzmann (1997) classify the mutual fund manager's investment style by K-means clustering. Among others, Abarbanell et al. (2003) employ K-means clustering to bring about the separation of institutions into groups of unique investment styles; and Elton et al. (2007) employ K-means clustering to classify mutual funds. Recently, Lusardi and Tufano (2009) have employed hierarchical clustering methods to determine groups of individuals with similar financial experiences. However, this study uses this technique to classify firms into groups for the purpose of testing the difference in financial performance and firm value models. It uses five latent factors and firm size as criteria and K-means clustering algorithm as a technique to cluster the firms into groups by operating and financing style.

This research differs from traditional industry classification systems that group firms using product and services outputs. Instead, it uses K-means clustering technique and

latent corporate finance decisions to cluster firms. The widely used industry affiliation is based on the Standard Industrial Classification (SIC) codes⁵⁵. SIC codes are based on company's production processes, which are subject to change due to evolution in technology, service style, and type of business⁵⁶. There exist differences among government agencies and even among data vendors⁵⁷. Among other classifications, we have the North American Industry Classification System (NAICS) and Global Industry Classification Standards (GICS codes). GICS has been found to be significantly better at predicting stock return comovements⁵⁸. In GICS firms are partitioned based on both operational characteristics and investors' attitudes, which differentiates it from other the two classification schemes. Fama-French (1997) also introduced industry reclassification based on SIC codes that are being widely used in asset pricing research. However Bhojraj et al. (2003) find evidence that GICS is out-performing FF, NAICS and SIC industry classifications. In a similar vein, Ramnath (2002) puts forth a five-analyst-based definition for classification of firms and shows that the scheme also differs from the SIC codes significantly. Chan et al. (2007) also find disagreements among such industry classifications.

⁵⁵ Grinold et al. (1989), Connor (1995), King (1966), Lessard (1974) and Roll (1992) study the influence of industry association.

⁵⁶ Clarke (1989) studies the SIC codes for grouping stocks in homogeneous markets based on some similar firm characteristic variables. He concludes that SIC is not successful.

⁵⁷ Kahle and Walkling's (1996), research uses a 1-digit SIC code to partition economic activity into 10 groups but Government agencies use 2-digit codes. The comparison of Compustat and CRSP reveals disagreement the between two vendors, as there is disagreement for 79% of 4-digit SICs, 50% for 3-digit SICs, 36% of 2-digit SICs, and 21% of the 1-digit SICs. Their predictions are similar to Guenther and Rosman (1994) and Fan and Lang (2000) in which alternative measures of homogenous grouping of input-output (IO) found richer descriptions of firms' association than SIC industry classification.

⁵⁸ Kirhnan and Press (2003) using the methodology of Guenther and Rosman (1994), show that NAICS has the edge over SIC codes for manufacturing, transportation, and service industries.

Contrary to above traditional transitive and disjointed industry classifications, the recent literature focuses on the firm-specific industry classifications. Among such studies, Hoberg and Phillips (2010a, 2010b) construct the “text-based network industry classifications” or TNICs. Their classification is based on the textual analysis of the firm-specific business description from 10-K filings. Rauh and Sufi (2012) and Lewellen (2012) employ firm-specific data from the CapitalIQ and 10-K filings to design industry groups, respectively. These classifications are found to be superior and more accurate in explaining financial variables such as profitability, stock returns, and leverage than traditional classifications.

This research is not directly contending with existing industry classifications. However, it proposes that firm-specific corporate finance decisions should be the criteria for grouping firms instead of product market outputs. These decisions relate to the firm-specific health and structure, not to their industry⁵⁹. Firms within the same industry are dissimilar in size, debt structure, and growth opportunities suggesting nonlinearity. Rauh and Sufi (2010) demonstrate that their sample firms simultaneously use heterogeneous types, sources, and priorities of debt. Using nonlinear technique, and corporate finance decisions as clustering criteria, 12 firm groups that are well-diversified are obtained – each group contains firms relating to 10 broad SIC code based industries.

These 12 firm clusters have manifold implications. First, firms are partitioned in their natural groups with criteria clearly marked in the average statistics. For example

⁵⁹ The literature that relate firms’ decision-making with health and structure of their industry include, for example, Brander and Lewis (1986), Maksimovic (1988), Dixit (1989), Bolton and Scharfstein (1990), Maksimovic and Zechner (1991), Hopenhayn (1992), Williams (1995), Miao (2005) and Spiegel and Tookes (2013).

cluster 1 firms on average show the most negative shifts in hybrid decisions (CVT_DEBT and PSK_USE) and a positive shift in FIN_FLEX. Similarly, among other clusters, cluster 2 criterion is PSK_USE; cluster 3 shows a negative shift in all latent factors; ST_CREDIT is a criterion for cluster 4; and LSIZE is a criterion for cluster 5 which represents larger firms. Then, cluster 6 criterion is the most negative shift in LT_INV; cluster 7 criterion is the most positive shift in CVT_DEBT; cluster 8 shows maximum positive change in LT_INV; and cluster 9 indicates the most positive shift in FIN_FLEX and a negative shift in ST_CREDIT. In the last three clusters, cluster 10 constitutes firms that are smaller and show the most negative shift in FIN_FLEX; cluster 11 shows the second most negative shifts in hybrid decisions; and cluster 12 (similar to cluster 3) shows negative shifts in all five latent factors. The pattern of criteria values suggests that all size firms engage in all five latent decisions and the speed of shift – both negative and positive – depends on their financial health and growth opportunities.

Second, cluster analysis offers firm segregation on the basis of growth opportunities (Tobin's Q) and financial strategies. For instance, cluster 1 on average has higher R&D intensity, lower book-to-market ratio, lower book leverage, pays zero dividends, and belongs to an industry with lower real sales growth. All these characteristics resemble the high IST (investment-specific technology) beta portfolio (Kogan and Papanikolaou (KP), 2014). Consistent to the high IST beta portfolio, cluster 1 firms have a higher Tobin's Q. In addition, cluster 9 resembles high IST beta portfolio firms, but differs from cluster 1 in values of FIN_FLEX, ST_CREDIT, and PSK_USE. However, compared to the KP finding, the investment rates in physical capital of firms in clusters 1 and 9 are higher and firm size is not smaller. Furthermore, clusters 3 and 10 mimic the properties of low IST beta

portfolios. Collectively, these four clusters constitute 45.4% of the whole data sample. To add to the KP finding, these clusters also differ in terms of information uncertainty and financial constraints, possibly due to a difference in age and size. Specifically, firms in clusters 1 and 9 on average exhibit higher information uncertainty than clusters 3 and 10. In terms of financial constraints, clusters 1 and 3 are less constrained than clusters 9 and 10.

Evidently, the greater the shift (both negative and positive) in financial flexibility (FIN_FLEX) is the clustering criteria in the above four clusters. Cluster 1 and 9 firms have more positive shifts in financial flexibility, and firms in clusters 3 and 10 have more negative shifts. From chapter 3, this study finds that the return predictability of FIN_FLEX is driven by covariances and FIN_FLEX based zero-investment return portfolio (FLEX) – that is hedge of the most negative (positive) shift in financial flexibility – is a good candidate for a risk factor. Thus, very prudently, this essay claims greater similarities between KP's IMC portfolio and the FLEX factor. The IMC portfolio is estimated as a portfolio of stock returns of investment-good producers minus stock returns of consumption-good producers. Similar to KP, this essay argues that shifts in financial flexibility (FIN_FLEX) are an important source of systematic risk and give rise to a systematic pattern in the stock return comovement of firms with different growth opportunities. This is the reason that the FLEX factor is capable of subsuming the return predictive power inherent in HML.

Third, this research does not find any conformity between the 12 clusters, and the 10 industries defined by Fama-French's (FF) SIC codes. In all clusters the member firms represent all ten industries. The result indicates that the clustering classification is not well

aligned with industry codes, they are different in measuring criteria. It claims that clustering is a better measure because it is based upon investment and financial styles which are fundamental to firm value.

In terms of profitability, the average future monthly raw (accumulative) returns for clusters 10 and 7 are 2.58% (2.25%) and 0.16% (0.99%), respectively. Cluster 10 consists of micro size, but not young firms, and they show on average a negative change in the five latent factors. Cluster 7 firms are above median size, but younger than cluster ten firms. They show positive change in all latent factors, except PSK_USE. This finding is consistent with the evidence in chapter 2, where firms in the low latent factor decile earn higher abnormal returns than high latent factor decile firms.

Fourth, the average statistics of the 12 clusters also reveal that the maturity of a firm is associated with more shifts in the corporate financing clustering criteria irrespective of their size. Among mature firms 47% are large and 44% are small and micro firms. However, firms with a median size show a mixed corporate finance behaviour. On average median sized firms are young and they constitute about 8% of the overall sample. These findings indicate the nonlinear relationship between firm size and age. A similar non-linear relationship is detected between book-to-market ratio and firm size, and the book-to-market ratio and the firm age. The average log book-to-market ratio of median (young) firms is lower than the corresponding ratio of large size firms.

Finally, Table 4.5 shows the result for predictive regression that all six clustering criteria offer interesting patterns in forecasting future firm value, cash flows, and earnings. For example, in Panel A this study finds firms that show most positive (negative) shifts in

FIN_FLEX create (destroy) future firm value. Panel B shows that low FIN_FLEX firms have robust predictive power of cash flows as compared to high FIN_FLEX firms. Then, Panel C shows that firms with negative shifts in FIN_FLEX have more future earning forecasting capability compared to firms with positive shifts in FIN_FLEX.

This chapter is structured as follows: The next section 4.2 describes the data, variable estimation, and methods used in this study. Then, section 4.3 provides the empirical results and discussion, while the final section 4.4 presents the concluding remarks.

4.2 Data Sample, Variable Estimation and Research Methods

4.2.1 Data sample

The data sample consists of US common stocks, excluding non-financial and non-utilities (i.e., stocks with four-digit SIC code 6000-6999 and 4900-49000) listed on the NYSE, AMEX, and NASDAQ. The market data is obtained from the CRSP database during the period 1983-2011, and the accounting data from the Compustat annual industrial files during the period 1984-2009. The study requires full information for total assets for each firm. To ensure survival, it requires firms to have complete information for two consecutive years.

The individual growth item is further winsorized at 1% and 99% to normalise the data; the balance sheet data is generally positively skewed. Then, the data is standardised to match the scales of the different items for simplifying relations between the distinctive variables of interest. The winsorization and standardisation are applied to accounting items

not on stock returns. The final data sample has 876,348 firm-month observations for the sample research period.

4.2.2 Estimation of clustering criteria

The clustering criteria include financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-term investment (LT_INV), convertible debt (CVT_DEBT), preferred stock usage (PSK_USE) and the firm size (LSIZE). All variables except LSIZE are change variables. The first five criterion are extracted from the factor analysis model as described in chapter 2. Factor analysis is a tool to reduce large input variable dimensionality. The model combines the variables together based on their common variance and excludes unique variance. The FIN_FLEX is the combination of the shifts in common equity (Item # A60), common equity-liquidation (Item#A238), cash and equivalents (Item#A1), common equity-tangible (Item#A11), invested capital (Item#A37), total current assets (Item#A4), cash (Item#A162), capital surplus (Item#A210), common shares outstanding (Item#A25) balance sheet items for a year $t-2$ to $t-1$ scaled by the total assets (Item#A6) of year $t-2$.

The important components that make up the short-term credit (ST_CREDIT) factor, are current liabilities-total (Item#A5), debt in current liabilities (Item#A34), total liabilities (Item#A181), total debt (Item#A9+A34), current liabilities-others (Item#A72), notes payable (Item#A206), and accounts payable (Item#A70). The short-term credit components have high loadings in the factor, thus it is labelled as ST_CREDIT. The third factor LT_INV represents the changes in long-term investment. The key components of LT_INV

include plant, property, and equipment – net (Item#A8), plant, property, and equipment – gross (Item#A7) and the long-term debt (Item#A9).

The fourth factor is predominantly comprised of hybrid components and the long-term debt. The components include debt convertible preferred stock (Item#A39), total long-term debt convertible (Item#A97), debt convertible and subordinated (Item#A154), and long-term debt (Item#A9). The last latent factor (PSK_USE) is also comprised of hybrid components, which include preferred stock (Item#A130), preferred stock liquidation (Item#A10), preferred stock nonredeemable (Item#A209) and total long-term debt convertible (Item#A97). The sixth criterion variable used is firms' market size following Fama and French (1992)'s method. It is the natural logarithm of the market value of equity; that is, stock price times the common stocks outstanding at the June of year $t-1$.

4.2.3 Descriptive statistics

Table 4.1 summarises the descriptive statistics of six clustering criteria in Panel A and the correlation matrix in Panel B for the data sample during the period 1985-2009. Panel A shows that FIN_FLEX has an average value of 0.026 representing the last quartile firms. The last quartile firms have a positive shift in financial flexibility. The distribution of FIN_FLEX is highly right-skewed, and its standard deviation is large with 0.868. This distribution reflects a negative median value that shows a reduction in their financial flexibility for more than 50% of firms.

The average value of -0.073 for ST_CREDIT shows that firms in the full data sample on average reduce short-term credit. The average value refers to the median and above median firms. The standard deviation is the least among these criteria. The mean and median values for LT_INV – the measure of the change in long-term investment – are 0.038 and -0.125, respectively. The mean value of LT_INV is representative of the above median firms and thus is right-skewed and has a standard deviation of 0.844.

Table 4.1
Descriptive statistics

Variable	Mean	Std Dev	25%	Median	75%	Max	Min	N
<i>Panel A. Descriptive Statistics</i>								
FIN_FLEX	0.026	0.868	-0.252	-0.153	0.010	-3.103	9.077	876348
ST_CREDIT	-0.073	0.603	-0.244	-0.131	0.023	-7.374	9.392	876348
LT_INV	0.038	0.844	-0.290	-0.125	0.144	-4.826	10.266	876348
CVT_DEBT	-0.002	0.908	-0.139	-0.057	0.045	-6.382	10.079	876348
PSK_USE	-0.057	0.813	-0.059	-0.004	0.054	-9.435	9.487	876348
LSIZE	5.097	2.117	3.569	4.969	6.496	-1.648	13.170	876348
<i>Panel B. Correlation</i>								
	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE			
ST_CREDIT	0.058	1						
	<.0001							
LT_INV	-0.061	0.066	1					
	<.0001	<.0001						
CVT_DEBT	0.013	0.006	0.011	1				
	<.0001	<.0001	<.0001					
PSK_USE	-0.151	-0.012	-0.000	0.014	1			
	<.0001	<.0001	0.644	<.0001				
LSIZE	0.002	-0.026	0.060	0.012	-0.011			
	0.0443	<.0001	<.0001	<.0001	<.0001			

FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE are the latent growth variables representing the change in firms' financial flexibility, short-term credit, long-term investment, convertible debt usage, and preferred stock usage respectively. These factors are extracted via the factor analysis model and a combination of several causal accounting growth items given in the firms' balance sheet. LSIZE is the log of the market value of equity at June in fiscal year t-1. Panel A reports the descriptive statistics for 876,348 firm-month observations for the above six clustering criteria. Panel B reports the correlation values along with p-values for showing the significance level of the correlations.

The CVT_DEBT – the latent measure of change in convertible debt usage – has mean and median values of -0.002 and -0.057, respectively. The mean and median statistics suggest that collectively firms reduce convertible debt usage. The mean and median values

of PSK_USE – the latent measure of the change in preferred stock usage – are -0.057 and -0.004, respectively. The average value is representative of the 25th percentile firms and is negatively skewed with a substantial standard deviation of 0.908. The overall sample firms on average try to reduce the preferred stock usage.

Finally, the mean and median values for the log of the market value of equity (LSIZE) are 5.097 and 4.969, respectively. The mean value represents the above median firms and thus distribution of LSIZE is right-skewed, and the standard deviation of 2.177 is highest among the six clustering criteria. In Panel B of Table 4.1, no significant high correlations are observed between the clustering criteria.

4.2.4 Research methods

This study intends to classify the universe of common stocks into their homogeneous clusters based on basic corporate financing decisions and firm size as explained under sub-section 4.2. There are mainly two types of cluster analysis. First is hierarchical cluster analysis, which is used for a limited number of observations. The second type is non-hierarchical (K-means) cluster analysis that is used for a large number of observations. Cluster analysis and its technique detail are given in Gordon (1999), Hair et al. (2006), and in Everitt et al. (2011). This essay employs K-means cluster analysis technique to cluster the firms and obtain the internally homogeneous and externally heterogeneous groups (Hartigan, 1975).

4.2.4.1 K-means cluster analysis

The simple K-means cluster analysis is a pattern recognition technique that groups the objects (firms) to discover the common practice styles in a dataset. Each group is called a cluster. The technique joins firms together into clusters by reducing the rows in the dataset. Firms in a cluster are similar to one another and firms in different groups are dissimilar in terms of clustering attributes. The grouping process starts with the similarity in observational firms over the variables. Cluster analysis does not have dependent variables. The quantification of similarity is done through distance measures between the objects in one group and between groups.

Cluster analysis employed here in this research measures distance by Euclidean Distances measure that is the length of a straight line between two observations. It helps to identify relatively homogeneous clusters of firms based on selected variables. Each cluster has a Centroid, which is the mean of all observations within the cluster also referred as a cluster center. The Euclidean distances are employed to allocate observations to each cluster, and least squares estimation used for cluster centers. Following iterative relocation algorithm of K-means is used, along with the discreteness constraints, as given in Lauprete (1998).

$$\begin{aligned} \text{Minimize} \quad & J(U, v) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik}) (d_{ik})^2 \\ \text{s.t} \quad & \sum_{i=1}^c u_{ik} = 1, \forall k = 1, \dots, n \\ & u_{ik} \in [0,1], \forall i = 1, \dots, c, \forall k = 1, \dots, n \\ & (d_{ik})^2 = (x_k - v_i)' A_i (x_k - v_i), \forall i, \forall k \end{aligned}$$

where, U is a $(c \times n)$ matrix of weights u_{ik} , and v is $(p \times c)$ matrix for which i^{th} column v_i is a vector representing the cluster i 's cluster center. Thus i indicate the cluster, c indicates the number of clusters, k represents the point, n is the number of points, u_{ik} equals 1 if point k belongs to cluster i , and 0 otherwise. x_k , is a p -dimensional vector, which is k^{th} data point. While $(d_{ik})^2$ is the distance of k point to cluster i , defined in terms of a positive definite symmetric matrix A_i . K-means cluster analysis is optimization process implying following condition:

$$v_i = \frac{\sum_{k=1}^n (u_{ik})x_k}{\sum_{k=1}^n (u_{ik})}$$

where, v_i are the cluster centroids. This analysis partitions the firms into groups with similarities optimally. It should be noted that this is near to optimal, but not necessarily optimal. The distance measure that incorporates our corporate finance decisions and firm size is described as under:

$$d(i,k)^2 = (FIN_FLEX_i - FIN_FLEX_k)^2 + (ST_CREDIT_i - ST_CREDIT_k)^2 + (LT_INV_i - LT_INV_k)^2 + (CVT_DEBT_i - CVT_DEBT_k)^2 + (PSK_USE_i - PSK_USE_k)^2 + (LSIZE_i - LSIZE_k)^2$$

where, FIN_FLEX , ST_CREDIT , LT_INV , CVT_DEBT , and PSK_USE are the latent growth variables representing the change in firms' financial flexibility, short-term credit, long-term investment, convertible debt usage, and preferred stock usage respectively.

According to Bushee (1998) there are no standard objective criteria for choosing the number of clusters. It is the long standing question with no clear solution, but often a matter of trial and error, educated guess, and judgement (Hair et al., 2006). This study uses the Pseudo-F statistics – the ratio of between-cluster variances to within cluster variance – and the cubic clustering criteria (CCC) measure – it is estimated by comparing the observed R-

squared to the approximate expected R-squared using an approximate variance-stabilizing transformation – for the choice of the right number of the clusters, where a higher positive number of both statistics are better with no sudden jumps. The negative numbers are indicative of the outliers. The positive CCC values indicate that R-squared is greater than expected when data sample is uniformly distributed, and thus would be a sign of the possible presence of clusters.

The use of K-means clustering techniques has manifold advantages. First, it provides the descriptions of the individual cluster that helps to identify what is unique in each cluster. Second, the firms with similar operating and financing styles can be identified and allocated to one group. Third, the clusters of different operating and financing styles can be compared in regard to their performance which might be explained by the differences. However, in K-means clustering methods, several analyses need to be performed before reaching the conclusion of how many clusters can be obtained. Thus, this research hypothesises that firm groups exist. This should not be a problem for the business analysis, if one closely checks the F and CCC statistics until there exists heterogeneity in the data sample.

This study uses the SAS software FASTCLUS procedure to select the initial K centroid, where K is the number of clusters pre-specified, and then each observation is allocated to the nearest centroid. The group of such observations allocated to a centroid is known as a cluster. K-means cluster analysis updates the centroid every time a new observation is added. This process is repeated until no observation changes the centroid.

This algorithm was first proposed by Elton and Gruber (1971) in financial research for reducing uncertainty about Markowitz's model inputs. Later among others⁶⁰, Goetzmann and Wachter (1995) used this procedure in real estate portfolio diversification. The authors claim that this procedure addresses the problem of estimation error in the Markowitz model of mean-variance by aggregating cross-sectional asset series. Further, they suggest that this model projects the forecasts with higher precision by decreasing dimensionality. Apart from such findings, they suggest that clustering offers clear recommendations for improving diversification.

4.3 Results

4.3.1 Cluster number choice

This section reports the results of the cluster analysis application. Cluster analysis produces 12 clusters using FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE as clustering criteria. The final number of clusters is 12 based on the higher value for the pseudo F statistics of (20001.57) and the higher value of 96.869 for CCC. There should be no sudden jumps in the value of both statistics if they are plotted across the expected number of clusters. It appears from figure 4.1 that the choice of 12 clusters is suitable for the data sample. In Figure 4.1, the FIG 1 shows that the F-statistics before 12 cluster number is very volatile, but afterwards the trend is smooth and decreasing. Similarly, in Figure 4.1, the FIG 2 shows that the CCC value of cluster number 12 is higher than the lower and higher cluster numbers. The CCC values after 12 are smoother than the

⁶⁰ See, for example, Brown and Goetzmann (1997), Bushee (1998), DeLong (2001), Abarbanell et al. (2003), and Elton et al. (2007).

previous clusters, thus making it the optimal cluster number to bring maximum differentiation in the dataset.

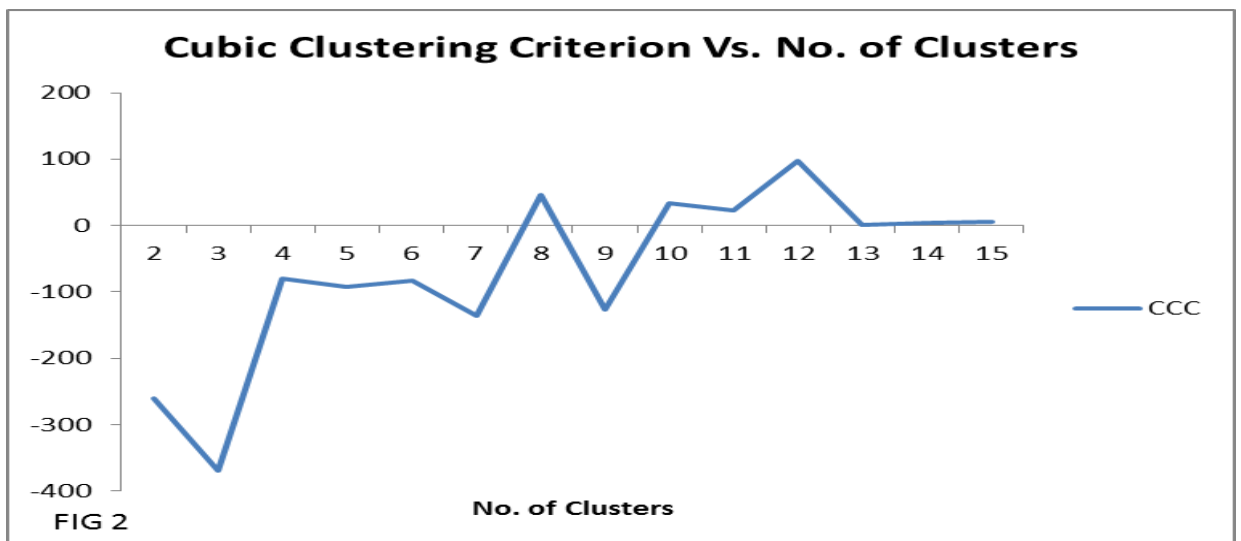
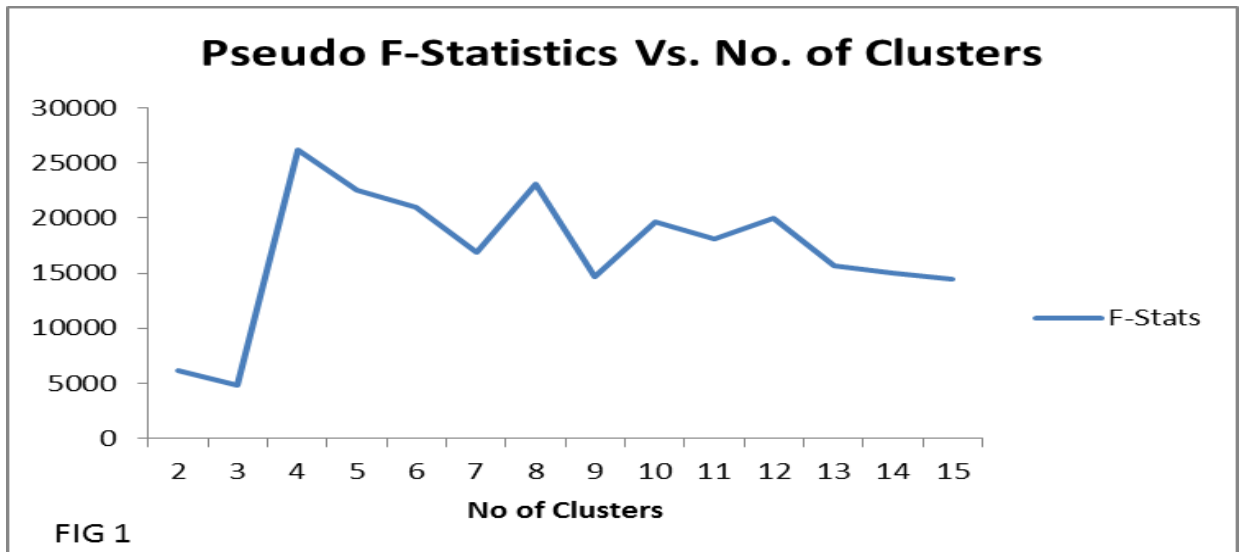


Figure 4.1: Choice of number of clusters and pseudo F-statistics and cubic clustering criteria

Figure 4.1 shows the plot of the pseudo F-statistics for the respective number of clusters that can be obtained employing the cluster analysis. Figure 2 plots cubic clustering criteria values across the number of clusters. The firms in the each cluster are grouped via six clustering criteria, including: FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE (defined in the header of tables 1 and 2).

4.3.2 Cluster summary

Panel A of Table 4.2 reports the cluster summary showing the number of clusters, firm observation frequency, the root mean square (RMS) standard deviations, the maximum distance from seed (centroid) to the observation, the radius exceeded, the nearest number, and the distance between cluster centroids. Broadly, it finds that clusters 3, 5, 6, 10 and 12 have maximum firm membership, lower RMS standard deviation, lower distance from seed to observation and lower distance between cluster centroids compared to other clusters that have higher, but similar values of the above attributes shown in Panel A. Cluster 12 has the highest membership with 23,281 firms and cluster 1 has the lowest membership with 357 firms. Panel B observes that among clustering criteria, LSIZE, FIN_FLEX, and PSK_USE have higher R-square value and ST_CREDIT with lowest R-square value. The overall R-square is 74.2%.

4.3.3 Cluster descriptive statistics

Table 4.3 reports the descriptive statistics for the K-means cluster analysis results of six clustering criteria attributes (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE) along with other accounting and market firm attributes. Among the other attributes, RET is the monthly raw future stock return; Lead12 is the next one-year accumulative monthly abnormal stock return; AGE is the logarithm of the firm's age; Q represents the Tobin's Q in the logarithm form; RDI indicates the research and development intensity; I_A is the investment-to-asset ratio representing the physical investment rate; LBTM is a measure of the logarithm of the book-to-market ratio; LEV is the proxy for the book leverage; ROA is a measure of firm profitability and ACCR_NEW

Table 4.2**Cluster summary and variables statistics**

<i>Panel A: Cluster Summary</i>						
Cluster	Frequency	RMS Std Deviation	Max Distance from Seed to Observation	Radius Exceeded	Nearest Cluster	Distance Between Cluster Centroids
1	357	1.3389	7.9526	> Radius	11	4.4194
2	429	1.2782	9.3047	> Radius	3	4.4881
3	21966	0.4253	6.2448	> Radius	10	1.8902
4	459	1.5241	10.5354	> Radius	3	4.6318
5	12543	0.5812	6.4846	> Radius	12	2.4822
6	1683	0.8724	6.2039	> Radius	3	2.4686
7	1223	1.2974	9.7178	> Radius	12	5.719
8	1486	1.2326	9.4087	> Radius	12	4.0081
9	607	1.5038	9.5953	> Radius	6	4.2006
10	11894	0.5175	5.5511	> Radius	3	1.8902
11	779	1.1297	7.7782	> Radius	1	4.4194
12	23281	0.4298	6.2509	> Radius	3	1.9788

<i>Panel B: Statistics for Variables</i>				
Variable	Total STD	Within STD	R-Square	RSQ/(1RSQ)
FIN_FLEX	0.87483	0.4439	0.7426	2.8852
ST_CREDIT	0.6140	0.5219	0.2776	0.384203
LT_INV	0.8509	0.6326	0.4475	0.80983
CVT_DEBT	0.9122	0.5735	0.6048	1.53019
PSK_USE	0.8198	0.4034	0.7579	3.129678
LSIZE	2.1245	0.8228	0.8500	5.668022
OVER-ALL	1.1463	0.5828	0.7415	2.86873

Pseudo F Statistic = 20001.57
Approximate Expected Overall R-Squared = 0.70126
Cubic Clustering Criteria =96.869

Table 4.2 reports the summary and clustering criteria variable statistics for 12 clusters obtained via the cluster analysis model. Panel A reports the frequency of firms' membership for each cluster across the 25 year (1985-2009) sample of 76,707 firm-year observations. Other information in this panel includes root means square standard deviation, maximum distance from seed to observation, radius limit exceeded, nearest cluster, and distance between cluster centroids. Panel B reports the total standard deviation, within standard deviation, and R-square related information for six clustering criteria. At the bottom of the table, we report the pseudo F statistics, approximate expected overall R-squared, and cubic clustering criteria. FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE are the latent growth variables representing the change in the firms' financial flexibility, short-term credit, long-term investment, convertible debt usage, and preferred stock usage respectively. These factors are extracted via the factor analysis model and a combination of several causal accounting growth items given in the firms' balance sheets. LSIZE is the log of the market value of equity at June in fiscal year t-1.

is a proxy of accounting accrual measure. The cash flow volatility is represented by CFVOL and SA is a financial constraint index. ZSCORE is an unlevered Z-score proxy for

bankruptcy introduced by MacKie-Mason (1990). SG_N is a proxy for real sales growth adjusted for inflation using the 2004 CPI index. ISG represents the real sales growth for the same three-digit SIC industry, and PayDiv shows proportions of firms that pay in dividends (see appendix A.2 for variable definitions).

Cluster 1 firms on average show the second most positive change in financial flexibility, and the most negative change in both convertible debt usage and preferred stock usage. Cluster 1 firms, comprising about 0.46% of the sample, are of median size, but are younger with an average age of 1.69 years. These firms employ higher growth opportunities (Tobin's Q) to intensively invest in the R&D. The higher RDI, and reduction in CVT_DEBT and PSK_USE are supported by borrowing short-term and decreasing the long-term investment (LT_INV). Among other attributes, on average these firms pay no dividends, are less profitable, have a lower log book-to-market ratio, and lower values for information uncertainty proxies. The firm size and age based financial constraint index (SA) takes the average value of -2.33, and ZSCORE takes the average value of 0.85.

Cluster 2 firms' criterion is the most positive change in the preferred stock usage (PSK_USE). In this cluster, comprising about 0.56% of the sample, firms increase LT_INV, FIN_FLEX, and show negative change in ST_CREDIT and CVT_DEBT. The average firms are small in size, but they are not young. The average age is 6.29 years. Cluster 2 firms have a higher chance of bankruptcy; the mean value of -2.03 for ZSCORE is the lowest among the 12 clusters. Cluster 2 firms are 20% levered, and only 11% of the firms pay dividends.

Cluster 3 firms on average show negative change in latent growth clustering criteria. In this cluster, comprising about 28.64% of the sample, firms on average are smaller in market size and are mature with an average firm age of 9.19 years. The average value of Tobin's Q is 0.41. The leverage level is similar to cluster 2, and 21% of the firms pay dividends. These firms with lower information uncertainty, lesser financial constraint index level, and with the second most positive ZSCORE of 3.40 indicate that there is no possibility of a firm's bankruptcy.

The main clustering criterion for cluster 4 is the most positive change in ST_CREDIT. These firms, comprising about 0.60% of the sample, are not large firms, but show a higher investment rate in fixed assets across the clusters. To capitalise on growth opportunities – with Tobin's Q of 0.71 – and to finance physical investments, the firms raise short-term credit. This increases the leverage ratio to 34%, which is 14% higher than the industry median leverage. Cluster 4 firms' real sales growth is higher than for any other cluster, and only 8% of its firms pay dividends. However, this high sales growth rate does not translate into positive profitability.

Firms in cluster 5, comprising about 16.35% of the sample, are larger and more mature than firms in other clusters. These big and profitable firms pay more in dividends – 63% of firms pay dividends – instead of exploiting growth opportunities. These firms have lower values for information uncertainty and financial constraint, and have higher values of ZSCORE, which indicates no possibility of bankruptcy.

Table 4.3 finds that most negative change in the long-term investment (LT_INV) – LT_INV's average value of -0.47 – is the criteria for firm grouping in cluster 6. In cluster 6, comprising about 2.2% of the sample, firms have an average size of 4.74 and age of 4.77 years. These firms, with considerable Tobin's Q (1.14), avoid investing in the fixed assets, instead retiring both short-term and convertible debt, and invest in R&D. Cluster 6 firms are small and young, but they are not the value firms. The LBTM ratio is very low (0.29) showing that firms that are growth firms may be due to their investment and financing behaviour. The leverage ratio is about 50% lower than the similar two-digit SIC code industry median leverage. However, in such a growth, financial flexibility, and higher Tobin's Q, the average statistics for information uncertainty (CFVOL) and financial constraints (SA) of 0.16 and -2.22 are higher respectively. One possible explanation would be an increase in PSK_USE as a financing source.

In cluster 7, comprising about 1.60% of the sample, firms show the highest positive change in the convertible debt usage (CVT_DEBT) compared to other clusters. These firms are above median firms with an average leverage of about 41%. These firms show an increase in FIN_FLEX, ST_CREDIT, and LT_INV, and reducing PSK_USE. Firms in this cluster are financially constrained, uncertain, and may gradually lean towards bankruptcy. This poor outlook is contrary to the LBTM of 0.33. This cluster on average earns lower mean RET of 0.16% and LEAD12 of 0.99% compared to the other clusters.

Cluster 8 firms, comprising about 1.94% of the sample, indicate the most positive change in long-term investments (LT_INV). The average age of firms is 6.69 years, and they pursue expansionary policies. The average size of firm is 5.13, which represents the

median size of the sample. The accounting accruals (ACCR_NEW) are highest across all the clusters. Twenty per cent of firms in this cluster pay in dividends as on average they are profitable.

Cluster 9 firms, comprising about 0.79% of the sample, on average show the most positive change in financial flexibility of 6.32 and the most negative change in short-term credit of 0.15 (see, Table 4.3). Cluster 9 differs from cluster 1 in ST_CREDIT, CVT_DEBT, PSK_USE, and AGE. These firms are young, growth firms instead of value firms. The LBTM ratio is 0.29. The real sales growth is two and half times the firms in the same two-digit SIC code industry firms. These firms have high information uncertainty and greater financial constraints, but are still not vulnerable to bankruptcy.

The average size of the firms in cluster 10 is 2.08, which is the lowest among all clusters. In this cluster, comprising about 15.5% of the sample, firms on average show the most negative change in financial flexibility and lower Tobin's Q. These firms are 'value' firms with the highest log book-to-market ratio. These firms indicate almost zero accounting accruals and the least investment in physical capital. They are more financially constrained, are not high in terms of information uncertainty and are not subject to bankruptcy. These firms are small and value firms, but definitely not young, which makes them unique. However, these firms on average earn higher future monthly raw and accumulative returns of 2.58% and 2.25% in the sample, respectively.

Table 4.3
Descriptive cluster analysis

Variable	<i>Clusters</i>											
	1	2	3	4	5	6	7	8	9	10	11	12
Ret	0.72%	1.76%	1.38%	0.19%	1.06%	0.60%	0.16%	0.26%	0.75%	2.58%	0.82%	1.12%
Lead12	1.51%	1.55%	1.51%	1.35%	1.11%	1.16%	0.99%	0.96%	1.70%	2.25%	2.20%	1.35%
FIN_FLEX	5.11	0.42	-0.15	0.97	-0.09	2.14	0.34	-0.08	6.32	-0.18	0.73	-0.09
ST_CREDIT	0.08	-0.11	-0.09	4.07	-0.11	-0.12	0.01	-0.06	-0.15	-0.10	0.00	-0.11
LT_INV	-0.38	0.23	-0.06	1.51	0.03	-0.47	0.49	3.89	-0.27	-0.17	0.14	-0.02
CVT_DEBT	-0.39	-0.06	-0.08	-0.28	-0.07	-0.16	5.56	-0.24	-0.19	-0.04	-0.34	-0.10
PSK_USE	-5.45	4.42	-0.01	0.18	-0.01	0.08	-0.06	-0.11	0.48	-0.01	-5.07	-0.01
LSIZE	5.03	3.96	3.96	3.59	8.40	4.74	5.45	5.13	4.42	2.08	4.86	5.93
AGE	1.69	6.29	9.19	5.04	14.10	4.77	7.80	6.69	2.96	8.46	3.43	10.36
Q	1.26	0.71	0.41	0.71	0.77	1.14	0.75	0.50	1.20	0.14	0.98	0.62
RDI	1.35	0.58	0.27	0.11	0.05	0.85	0.41	0.06	1.10	0.11	0.74	0.19
I_A	0.44	0.20	0.04	2.09	0.05	0.19	0.27	1.06	0.91	0.01	0.16	0.05
LBTM	0.25	0.36	0.57	0.35	0.36	0.29	0.33	0.48	0.29	0.72	0.32	0.45
PayDiv	0.00	0.11	0.21	0.08	0.63	0.07	0.13	0.20	0.04	0.11	0.05	0.37
LEV	0.03	0.20	0.19	0.34	0.22	0.05	0.41	0.39	0.03	0.23	0.13	0.18
ROA	-0.12	-0.16	-0.01	-0.01	0.12	-0.04	-0.02	0.03	-0.11	-0.06	-0.07	0.07
INCOME	-0.25	-0.22	-0.04	-0.12	0.07	-0.09	-0.09	-0.02	-0.24	-0.07	-0.14	0.02
Accr_New	0.20	0.23	0.05	0.64	0.07	0.21	0.29	0.66	0.31	0.00	0.15	0.07
CCASH	1.04	0.09	-0.01	0.07	0.03	0.59	0.19	0.04	1.07	-0.02	0.35	0.02
FCF	-0.45	-0.45	-0.09	-0.76	0.00	-0.31	-0.38	-0.68	-0.55	-0.07	-0.29	-0.04
CFVOL	0.16	0.15	0.10	0.10	0.07	0.16	0.13	0.08	0.20	0.09	0.16	0.09
SA	-2.23	-2.24	-2.41	-2.08	-3.19	-2.22	-2.68	-2.70	-2.00	-1.77	-2.34	-2.90
ZSCORE	0.85	-2.03	3.40	3.17	5.76	1.52	1.66	2.28	1.26	3.37	-0.20	5.16
ISG	1.27	3.33	2.40	3.28	3.24	2.31	2.60	3.70	2.00	2.07	1.40	3.07
SG_N	3.32	3.02	3.23	15.68	3.13	3.42	3.38	5.51	5.96	1.48	3.74	2.58
N	0.47%	0.56%	28.64%	0.60%	16.35%	2.2%	1.59%	1.94%	0.79%	15.51%	1.02%	30.35%

This table reports the descriptive statistics for the 12 clusters of firms obtained via K-means cluster analysis, a pattern recognition technique. Several descriptive characteristics, including the clustering criteria are reported. The six clustering criteria are defined in the header of the Table 2.1 and 2.2 and the other characteristics are defined in appendix A for brevity. The following are obtained: mean, standard deviation, lower quartile, median, upper quartile, maximum, minimum and the firm-month frequency for several characteristics across individual cluster. Only means are reported for brevity.

In cluster 11, comprising about 1.0% of the sample, firms are of median size, but on average are 3.43 years old. The growth opportunities and R&D intensity are supported by financial flexibility. These firms behave like growth instead of value firms. They increase the long-term investment (LT_INV), and decrease both convertible debt and preferred stock usage. These firms are less financially constrained, but on average have higher information uncertainty that categorises them as bankrupt firms.

The cluster 12 firms, comprising about 30.4% of the sample, are comparable to cluster 10 firms, to the extent that there is a negative change in all latent growth decisions (see table 4.3). These two clusters differ in age and firm size. The average firm size of cluster 12 represents above the median of all firms and the average age is the second most mature of all clusters. These firms are profitable firms with an average ROA of 0.07 and 37% of firms pay in dividends. The firm's financial health is better, as they have low information uncertainty, with lower financial constraints, and with almost no fear of bankruptcy.

Relative to other literature, this essay finds that the K-means cluster analysis shown in Table 4.3 can shed light on the issues of investment-specific technology, growth opportunities, and asset prices. The attributes of firms in clusters 1, 3, 9 and 10, to some extent, are consistent with the evidence documented by Kogan and Papanikolaou (KP) (2014). They find that firms with higher growth opportunities or with higher exposure to investment specific shocks (IST) have higher RDI, higher cash holdings, higher physical capital investment rates, and make fewer payments in dividends as compared to firms with lower IST shocks. Table 4.3 shows that clusters 1 and 9 represent high IST shock portfolio stocks, although the clusters' descriptives are partially different to KP. The firms in clusters 1 and 9

are of median size, and reduce long-term investment, and KP's firms are generally smaller and increase investment. The investment rates (I_A) in physical capital by cluster 1, and 9 firms are lower than the other clusters. On the other side, clusters 3 and 10 represent low IST shock portfolio stocks; these firms are small, but mature. It seems that these results provide a new way to identify and classify firms based on the investment-specific shocks or growth opportunities. One of the special features of the classification is the resulting systematic numerical method without any subjective inputs. Additionally, this essay finds that clustering would obscure the firm's varying investment specific behaviours; further segmentation of the two IST portfolios is due to the difference in the usage of hybrid securities which might be used to exploit the available growth opportunities.

4.3.4 Clustering versus industry patterns

This section seeks further evidence of internal cluster homogeneity and external cluster heterogeneity. The individual cluster is segmented into industry groupings using SIC codes. The implications of this exercise are threefold. First, it shows that overall clusters' homogeneous distribution holds when a cluster is further segmented into industry groupings. Second, in terms of whether the individual cluster is well-diversified across broad traditional industry groupings, the diversification evidence would imply that industry classification standards based on product market outputs are not adequate. Third, the corporate finance decisions are closely related to firm-specific health and structure instead of industry. These decisions are managers' perceptions about firms' health and structure and in turn about investment opportunities. This research believes that the existence or non-existence of

investment opportunities is the actual cause for financial transactions, choice of the financing mix, and ultimately the product market outputs.

This study argues that rigid traditional industry groupings based on product market outputs would not be able to explain issues like debt heterogeneity. In a similar way, firms belonging to the same product market would show variations in size, information uncertainty, financial constraints, and growth opportunities. Thus, the idea is that matching in terms of corporate finance decisions would be a better way to identify comparable firms. These groups when compared to the traditional classifications would provide an improved understanding of business and firm value.

Table 4.4 reports the mean values for six clustering and size variables in 10 SIC codes based industries, and the industry distribution for each cluster⁶¹. The industry groupings are based on the SIC codes that are defined in Fama-French's online data library. The 10 industry groups are: NoDur, Durbl, Manuf, Enrgy, Chems, BusEq, Telcm, Shops, Hlth, and Others (the details are given in appendix C). This essay excludes two regulated financial and utilities industries. NoDur represents non-durable consumer firms that include food, tobacco, textiles, apparel, leather, and toy industries. Durbl represents consumer durables that include cars, TVs, furniture, and household appliances. Manuf represents manufacturing that includes machinery, trucks, planes, office furniture, paper, and commercial printing industries. Energy includes oil, gas, and coal extraction and products industries. Chems includes chemicals and allied products industries. BusEq represents business equipment that includes computers,

⁶¹ Table 4.4 only presents statistics for clustering criteria, the control variables are not tabulated for brevity concerns. Table C.1.2 reports the detailed results in appendix C.

software, and electronic equipment industries. Telcm includes telephone and television transmission industries. Shops include wholesale, retail, and some services (laundries and repair shops) industries. Hlth includes healthcare, medical equipment, and drug industries. And Others include mines, construction, building materials, transportation, hotels, bus services, and entertainment industries. The other important firm attributes included in the study are monthly raw future stock returns (RET) and next one year accumulative monthly abnormal stock return (LEAD 12), book leverage (LEV), growth opportunities (Q), research and development intensity (RDI), firm profitability (ROA), firm age (AGE), size and age based financial constraints index (SA), investment-to-asset ratio (I_A), dividends paid (PayDiv), cash holdings (CASH), common equity to lagged total asset ratio (EQUITY), net debt (DEBT), an unlevered Z-score (ZSCORE), accounting accruals measure (ACCR_NEW), and (INV_AGE) (see appendix A.2 for definitions).

Table 4.4 in the last column 'N' provides the industry frequency distribution in each cluster. The table shows that there is no clear relation between industry classification, and cluster distribution. All industries are included in each cluster. From this table, it can be concluded that these two ways of classifying firms are very different and are generally independent of each other. The industry classification is based upon their product types, and the clustering is based upon their investment and finance structure; this study believes the second method is better in relation to the value of financial analysis and asset evaluation. For example, when it uses a regression model to estimate an evaluation model, one needs the assumption of linear relation and constant coefficients for explaining variables. One can claim that the firms generally closer to the homogeneous in business style will have a better chance of satisfying the normal linear regression assumption. It seems not so for an industry, because

one has to filter on homogeneity in the classification and the result does not support this either because each industry includes some firms from every cluster.

From Table 4.4, the majority of the cluster 1 firms belong to BusEq (189) and Hlth (95) industries. All industry firms are represented in this cluster. However, there is only one firm representing the Enrgy industry. This particular Enrgy industry firm brings the most negative change in the preferred stock issuance (PSK_USE), which is the clustering criterion for cluster 1. Except for the Enrgy firm, all others show a homogeneous pattern in firm attributes, except ZSCORE and RDI. In cluster 2, the Durbl industry has the least representation of eight firms, and BusEq has the highest representation of 107 firms. The distributions of clustering criteria (PSK_USE) values, FIN_FLEX, LSIZE, (and other variables not in the table such as LEV, AGE, Q, SA, and ACCR_NEW) are homogeneous across the industries. However, industries such as Enrgy and Shops have zero R&D intensity, and Telcm firms pay no dividends. The average values of ZSCORE indicate that the Enrgy, Chems, BusEq, Hlth, and Others industries show a higher possibility of going into bankruptcy (the results for ZSCORE discussed here and after are given in appendix C, Table C.1.2). Cluster 3 shows the homogenous clustering criteria distribution with a few exceptions. The Enrgy industry and Telcm industry firms show positive change in long-term investment (LT_INV). The industries such as Chems and Telcm show positive change in preferred stock issuances (PSK_USE). In addition, Hlth industry firms have a negative ZSCORE of 3.77.

Cluster 4 with a 459 firm membership makes up just 0.60% of the overall sample. This study finds the uniform distribution of ST_CREDIT – the clustering criteria for this cluster – across the industries. It does not find many differences in the distributions of the attributes.

However, the negative ZSCORE of -1.03 and -0.45 for Enrgy and Hlth industries are exceptions. Cluster 5 makes about 16.4% of the overall sample. The cluster is topped by the 2485 business equipment and the 1956 manufacturing firms. LSIZE is the featuring criterion for this cluster, and its values are homogeneous across the 10 industries. All firms on average reduce the financial flexibility, short-term credit, convertible debt, and preferred stock usage with one or two exceptions. In cluster 6, the BusEq and Hlth industries have the most number of firms represented and Enrgy the least number of firms. The mean statistics for firm attributes for 10 industries for this cluster are homogeneous. The only exception is the negative values of ZSCORE for Enrgy, Chems, and Hlth industries.

Cluster 7's clustering criterion is CVT_DEBT which is consistently homogeneous across the industries. Similarly, average attributes like FIN_FLEX, LT_INV, LSIZE, LEV, Q, AGE, SA, I_A, CASH, EQUITY, ACCR_NEW, and INV_AGE are uniform. In terms of ZSCORE mean statistics, only the Hlth industry has a negative mean value of -1.25. In cluster 8 we find that the clustering criterion is long-term investment (LT_INV) which is uniformly distributed across the industries. The distribution of the other variables is uniform across the 10 industries with few exceptions. Cluster 9 makes only 0.79% of the overall data sample. The criteria value distribution across industries on average is homogeneous. The Hlth, Chems, and Durbl industries are found to make more intensive R&D investments, and thus are prone to bankruptcy.

Table 4.4 reports that firms across the 10 industries in cluster 10 have homogeneous values for six clustering criteria. Likewise, among other firm attributes, it finds a homogeneous pattern for LEV, AGE, SA, CASH, EQUITY, and INV_AGE. Despite having

low information uncertainty, the firms on average are financially constrained across the industries. And only firms belonging to Enrgy, Telcm, and Hlth have negative ZSCOREs of -1.13, -4.43, and -1.30 respectively. In cluster 11, the attributes such as FIN_FLEX, CVT_DEBT, PSK_USE, LSIZE, LEV, Q, AGE, SA, ACCR_NEW, and INV_AGE are homogeneous. Consistent with cluster 10, the SA index shows low financial constraint levels and industries such as Enrgy, Chems, BusEq, Telcm, and Hlth show signs of bankruptcy. Cluster 12 has the highest memberships with 23,281 firms (almost 30.35% of the overall data sample). The average statistics of several attributes are similar across 10 industries, and LT_INV and RDI are exceptions. These firms are large, mature, less financially constrained, less uncertain, and have no fear of bankruptcy.

The cluster industry frequency and criteria (Table 4.4) distribution supports the statements. The clustering criteria of homogeneous in firms within each cluster, and heterogeneous between different clusters, do not resemble the industry classification⁶². Each cluster is further segmented into 10 traditional SIC codes based industry groups. The pattern suggests that corporate finance decisions are related to firms' health and structure might be more important than the type of products they produce.

⁶² See, Table C.1.3 in appendix C for mean statistics of clustering criteria and other firm attributes across 10 SIC codes based industry groups. The results clearly indicate that clustering based classification is effective in bringing the qualitative and quantitative differentiation among the firms based on operating business and firm size criterion. For example, there is positive mean for FIN_FLEX in 7.23% firms across clusters (1, 2, 4, 6, 7, 9 and 11), but full sample mean across 10 industries indicate that 40% of the firms have positive mean for FIN_FLEX. In further robustness check similar evidence is obtained in size ranked industry mean statistics for clustering criteria. It also indicates that size is a necessary criterion. Micro firms consists of 45584 (59%), small firms consists of 16080 (21%), and large firms consists of 15043 (20%) firm-year observations respectively (see, e.g., Table C.1.3).

Table 4.4**Industry characteristics' pattern across clusters**

Clust#1	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	5.92	2.08	-0.79	-0.61	-5.79	4.04	5
Durbl	4.99	-0.89	1.84	-0.72	-6.23	3.18	2
Manuf	4.65	0.12	-0.38	-0.69	-5.87	4.69	14
Enrgy	2.98	-2.63	0.26	-0.58	-6.68	2.43	1
Chems	4.73	-0.19	-1.19	0.06	-6.42	3.93	7
Buseq	5.03	0.22	-0.36	-0.30	-5.30	5.33	189
Telcm	5.06	-0.26	0.53	-0.76	-5.65	4.88	8
Shops	5.10	-0.32	0.35	-0.60	-3.84	4.55	11
Hlth	5.25	-0.09	-0.66	-0.53	-5.62	4.69	95
Others	5.47	-0.24	-0.06	-0.21	-5.89	5.22	25
Clust#2	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	0.42	-0.05	0.72	-0.79	4.95	4.10	14
Durbl	0.16	0.09	0.57	-0.46	4.57	4.15	8
Manuf	0.19	-0.16	0.00	-0.16	3.97	4.45	48
Enrgy	0.22	-0.40	0.43	-0.35	4.13	3.71	40
Chems	0.20	-0.35	0.25	0.16	3.74	4.22	12
Buseq	0.53	-0.02	-0.12	0.01	4.62	3.93	107
Telcm	0.35	-0.21	0.83	0.05	3.92	4.83	23
Shops	0.51	0.33	0.52	-0.06	4.59	3.47	36
Hlth	0.49	-0.26	-0.10	0.08	4.67	3.92	82
Others	0.52	-0.02	0.81	0.03	4.33	3.69	59
Clust#3	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	-0.12	-0.09	-0.07	-0.06	-0.01	3.98	1636
Durbl	-0.13	-0.07	-0.04	-0.06	-0.01	3.99	817
Manuf	-0.13	-0.10	-0.07	-0.06	-0.01	3.95	3192
Enrgy	-0.15	-0.19	0.10	-0.12	-0.02	3.98	782
Chems	-0.17	-0.07	-0.04	-0.09	0.01	3.97	482
Buseq	-0.16	-0.09	-0.12	-0.09	-0.01	3.95	5961
Telcm	-0.18	-0.07	0.05	-0.13	0.02	4.15	372
Shops	-0.13	-0.06	-0.04	-0.08	-0.01	3.98	2730
Hlth	-0.21	-0.09	-0.06	-0.09	-0.02	3.98	2962
Others	-0.15	-0.12	-0.01	-0.08	-0.01	3.94	3032
Clust#4	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	0.74	3.94	0.70	-0.46	0.39	3.64	30
Durbl	1.09	4.47	0.97	-0.26	-0.14	3.61	12
Manuf	0.76	4.16	1.92	-0.16	0.37	3.32	45
Enrgy	0.99	4.19	2.81	-0.47	0.02	3.35	17
Chems	0.76	3.41	1.83	-0.25	0.23	3.54	9
Buseq	1.07	4.02	1.26	-0.44	0.14	3.57	102
Telcm	1.03	3.88	1.78	-0.29	0.13	3.51	23
Shops	0.68	4.25	1.17	-0.01	-0.07	3.59	92
Hlth	1.11	4.02	1.73	-0.40	0.38	3.78	51
Others	1.30	3.98	1.82	-0.28	0.31	3.69	78

Table 4.4 (Continued)

Clust#5	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	-0.13	-0.10	-0.04	-0.03	-0.01	8.46	1203
Durbl	-0.14	-0.09	-0.03	-0.05	-0.02	8.25	372
Manuf	-0.13	-0.11	-0.03	-0.07	-0.03	8.21	1956
Enrgy	-0.19	-0.18	0.27	-0.12	-0.01	8.69	972
Chems	-0.16	-0.11	-0.03	-0.03	-0.02	8.38	657
Buseq	0.02	-0.07	-0.04	-0.11	0.00	8.45	2485
Telcm	-0.14	-0.14	0.12	0.00	-0.01	8.85	622
Shops	-0.10	-0.11	0.05	-0.10	0.00	8.34	1433
Hlth	-0.05	-0.09	0.03	-0.08	-0.02	8.55	1177
Others	-0.11	-0.12	0.09	-0.05	-0.01	8.16	1666
Clust#6	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	2.05	-0.28	-0.31	-0.19	0.12	4.45	58
Durbl	2.20	-0.24	-0.36	-0.46	0.13	4.23	35
Manuf	1.98	-0.17	-0.18	-0.29	0.13	4.57	94
Enrgy	1.93	-0.40	-0.52	-0.32	0.20	4.02	34
Chems	1.97	-0.22	-0.82	-0.01	0.22	4.33	20
Buseq	2.22	-0.05	-0.49	-0.14	0.05	4.88	629
Telcm	2.15	-0.14	-0.38	-0.19	0.13	4.97	52
Shops	2.01	-0.11	-0.13	-0.38	0.22	4.54	112
Hlth	2.13	-0.17	-0.65	-0.02	0.03	4.78	486
Others	2.13	0.00	-0.35	-0.30	0.13	4.67	163
Clust#7	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	0.20	0.27	0.61	5.37	-0.18	4.37	45
Durbl	0.18	-0.05	0.44	5.50	-0.06	4.37	29
Manuf	0.09	0.11	0.41	5.14	-0.21	5.40	104
Enrgy	0.12	-0.46	1.44	5.11	-0.32	5.41	66
Chems	0.69	-0.06	0.69	5.60	-0.27	4.95	18
Buseq	0.42	0.04	0.23	5.69	0.02	5.89	349
Telcm	0.24	0.00	0.98	5.58	0.01	5.59	43
Shops	0.33	0.13	0.46	5.54	-0.05	5.04	127
Hlth	0.48	0.01	0.40	5.84	0.00	5.73	276
Others	0.23	-0.08	0.71	5.31	-0.10	4.93	166
Clust#8	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	-0.19	0.42	3.58	-0.04	-0.11	5.06	74
Durbl	-0.30	0.41	3.32	0.01	-0.16	5.20	27
Manuf	-0.17	0.40	3.74	0.04	-0.07	4.90	134
Enrgy	-0.25	-0.42	4.14	-0.47	-0.06	5.30	323
Chems	-0.10	-0.06	4.21	-0.52	-0.25	4.97	30
Buseq	0.24	0.35	3.73	-0.31	-0.24	4.71	111
Telcm	-0.11	-0.12	4.15	0.37	-0.09	5.61	137
Shops	0.03	0.00	3.59	-0.41	-0.12	4.73	134
Hlth	0.00	0.26	3.73	-0.23	-0.16	5.00	109
Others	-0.02	-0.25	3.92	-0.33	-0.10	5.22	407
Clust#9	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	6.13	-0.74	0.02	-0.35	0.65	3.62	23
Durbl	6.24	-0.41	0.31	-0.32	0.65	3.93	11
Manuf	6.42	-0.30	-0.33	-0.34	0.73	3.74	29
Enrgy	6.97	-0.45	1.03	-0.64	0.19	5.11	8
Chems	6.47	-0.34	0.03	-0.13	1.13	3.73	10
Buseq	6.30	-0.04	-0.41	-0.13	0.38	4.79	254
Telcm	5.90	-0.49	1.68	-0.22	0.35	4.76	16
Shops	6.35	-0.13	0.33	-0.38	0.71	3.85	26
Hlth	6.34	-0.23	-0.68	0.00	0.44	4.30	164
Others	6.37	0.01	0.22	-0.63	0.64	4.19	66

Table 4.4 (Continued)

Clust#10	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	-0.16	-0.11	-0.15	-0.06	0.00	2.14	923
Durbl	-0.18	-0.10	-0.15	-0.04	-0.03	2.15	321
Manuf	-0.17	-0.09	-0.16	-0.03	-0.01	2.07	1900
Enrgy	-0.19	-0.13	-0.19	-0.03	-0.02	1.81	715
Chems	-0.16	-0.09	-0.13	0.00	-0.02	2.02	305
Buseq	-0.20	-0.10	-0.19	-0.04	-0.02	2.08	3221
Telcm	-0.17	-0.21	-0.17	-0.17	0.01	2.22	167
Shops	-0.16	-0.10	-0.15	-0.04	0.00	2.09	1597
Hlth	-0.20	-0.07	-0.15	-0.06	0.00	2.20	1207
Others	-0.18	-0.09	-0.17	-0.05	0.00	2.07	1538
Clust#11	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	0.18	-0.27	0.15	-0.43	-4.44	4.31	24
Durbl	0.62	0.07	-0.45	-0.06	-3.99	3.91	11
Manuf	0.39	0.08	0.09	-0.24	-4.56	4.80	42
Enrgy	-0.01	-0.19	0.24	-0.20	-4.23	4.53	31
Chems	0.75	-0.50	-0.21	-0.18	-5.96	4.40	9
Buseq	0.96	0.11	0.15	-0.34	-5.17	4.92	308
Telcm	0.51	-0.07	0.29	-0.32	-5.31	5.33	26
Shops	0.50	0.04	0.45	-0.46	-4.67	4.64	64
Hlth	0.83	-0.06	-0.05	-0.38	-5.61	4.94	174
Others	0.61	-0.10	0.35	-0.30	-4.70	5.01	90
Clust#12	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	N
NoDur	-0.10	-0.10	-0.06	-0.08	-0.02	5.91	1591
Durbl	-0.10	-0.09	-0.06	-0.07	-0.02	5.97	864
Manuf	-0.12	-0.11	-0.06	-0.07	-0.01	5.92	3800
Enrgy	-0.16	-0.21	0.22	-0.18	0.00	6.00	989
Chems	-0.12	-0.10	-0.09	-0.03	0.00	6.16	768
Buseq	-0.05	-0.07	-0.06	-0.12	0.00	5.89	5328
Telcm	-0.15	-0.14	0.10	-0.11	-0.02	6.01	735
Shops	-0.09	-0.09	0.02	-0.11	-0.01	5.95	2994
Hlth	-0.07	-0.11	-0.04	-0.11	-0.01	5.86	2705
Others	-0.10	-0.13	0.04	-0.09	-0.01	5.96	3507

The table reports the mean statistics of six clustering criteria and the distribution of industries across 12 clusters. Similar statistics for 19 other attributes of these clusters are shown in appendix C. The six clustering criteria include FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE. The 10 industry groups are constituted using the SIC codes given in Fama and French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) and are defined in appendix C. Financial and utilities industries are excluded. NoDur represents consumer non-durables. Durbl represents consumer durables. Manuf represents manufacturing firms. Enrgy includes oil, gas, and coal firms. Chems represents chemicals and allied products industries. BusEq represents business equipment. Telcm represents telephone and television transmission industries. Shops include wholesale, retail, and some service firms. Hlth includes health-related firms. And Others include mines, construction, building materials, transportation, hotels, business services, and entertainment industries.

4.3.5 Predictive regression analyses across clusters

Overall in the full sample regression model (all firms) this study makes the assumption that the relation between operating factors, and performance is the same for all firms. This assumption is not automatic. Theoretically, firms facing different possibilities will take a different strategy in operating and financing management. For example, when one analyses a firm's decision to take a high leverage, it might be correct when a company's operating business risk is relatively low; however, it could easily go in the opposite direction when a firm's operating situation is deteriorated. A similar argument could be made when a firm increases its long-term investment – it could be either right or wrong in different situations.

The clustering result has provided a grouping of firms with similar operating and financing situations (more homogeneous within each cluster), and the simple constant relation assumption might be much better established. This study anticipates a better regression result within each cluster than the result obtained for all firms. It believes this might be a big problem in most of the current regression models reported in literature. The clustering analysis would have a great potential to improve the quantitative business analysis in future. In such an effort, this study finds that relations among latent operating business decisions and firm earnings, value, and cash flows in predictive regressions across the clusters depict true sign of relations as compared to full sample based linear regressions, which indicate the importance of both cluster analysis and the clustering criterions.

Table 4.5 runs predictive regressions using standard clustered error OLS regressions. The dependent variables $X_{i, t+1}$ – in the model given below – include Tobin's Q_{t+1} , CF_{t+1} , and $EARN_{t+1}$. The Q_{t+1} is the measure of the firm's future value; CF_{t+1} is the measure of the firm's future cash flows in the year $t+1$, and $EARN_{t+1}$ is the firm's future earnings measure. The

independent variables include a set of clustering criteria, and a set of control variables. The clustering criteria include FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE. The control variables include ASSETG, LBTM, NegY, YB, NegAccr, PosAccr, NoDiv, Div_B, Lagret, AGE, CVOL, and SG. The clustering criteria (CC) variables and control variables (CV) are defined in appendix A. In the interests of brevity, Table 4.5 only reports coefficients and *t*-statistics of CC for the individual cluster model along with the model adjusted R-squared value.

$$X_{i,t+1} = \alpha + \beta_i CC_{i,t+1} + \beta_j CV_{i,t+1} + \varepsilon_{i,t+1} \quad (4.2)$$

4.3.5.1 Predicting the future firm value

The results in Panel A of Table 4.5 show that a change in financial flexibility (FIN_FLEX) positively determines future firm value in nine out of 12 clusters. In clusters 3, 10 and 11, comprising about 45.2% of the sample, firms show a negative relationship with the future firm value. Except for cluster 2, all other cluster coefficients for FIN_FLEX are statistically significant. The negative relationship of FIN_FLEX– Q_{t+1} is a puzzle as it is contrary to the literature that suggests positive relationships (Gamba and Triantis, 2008). One explanation could be a higher level of information uncertainty and financial constraints. For example, the cluster 11 firms are more uncertain, financially constrained and are already in the bankruptcy zone.

The second criteria ST_CREDIT has a negative relationship with future firm value in eight out of 12 clusters. The firms in clusters 4, 5 and 8, comprising about 19% of the sample, have a positive ST_CREDIT relationship with the Q_{t+1} . All coefficients of ST_CREDIT are statistically significant except cluster 4 and 11. Cluster 4 shows the most positive change and

cluster 11 shows the marginal positive change in short-term credit. In the case of LT_INV, except for cluster 1 and 5 firms, all other cluster firms have a negative relationship with the Q_{t+1} . The cluster 1 and 5 are two extremes – cluster 1 firms are young and cluster 5 firms are mature – if age is a consideration. Besides, cluster 1 firms are median sized firms and cluster 5 firms are bigger compared to firms in other clusters. The LT_INV regression coefficient is insignificant for cluster 9 firms as this cluster shows a large increase in FIN_FLEX.

The regression coefficients for CVT_DEBT show a mixed result, where the clusters 1, 2, 3, 7, 8 and 9 show a positive relationship with Q_{t+1} and the other six clusters have a negative association. The cluster number 1, 3, 4, 8 and 9 is statistically insignificant in determining the future firm value. The results for PSK_USE are also mixed; in clusters 1, 3, 6, 7, 9, and 10 PSK_USE has a positive association to Q_{t+1} and the remaining clusters have a negative relationship. The clusters 3, 6 and 8 are statistically insignificant in explaining Q_{t+1} . Cluster 2 firms show the most positive change in preferred stock issuance and have a negative relationship to Q_{t+1} . The cluster 1 firms, showing the most negative change in preferred stock issuance have a positive relationship to firm value. Overall, the relationship between PSK_USE - Q_{t+1} is not consistent, some are significantly positive and others are significantly negative. Finally, the LSIZE- Q_{t+1} relationship is significantly positive in seven out of 12 clusters. Clusters 6 and 12 have a significantly negative LSIZE- Q_{t+1} relationship. The cluster numbers 1, 4 and 11 coefficients are statistically insignificant in explaining Q_{t+1} . The same coefficient assumption for all firms is clearly violated here. Across the 12 models, the lowest adjusted R-squared of 0.23 is for cluster 10 and the highest of 0.45 is for cluster 9. The adjusted R-squared is higher than that obtained in the overall sample in Table 2.4 (most are less than 0.05).

Table 4.5

Predicting future fundamentals across the clusters

Panel A: Predicting Future Firm Value (Q_{t+1})										
Cluster		Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	Adj R-Sq	
1	Coeff.	6.905	0.245	-0.457	0.153	0.019	0.412	0.004	0.3736	
	t-value	9.31	4.28	-6.98	2.40	0.22	6.82	0.06		
2	Coeff.	3.110	0.132	-0.144	-0.362	0.244	-0.099	0.109	0.4084	
	t-value	9.88	1.39	-2.58	-7.08	5.04	-2.64	3.27		
3	Coeff.	2.392	-0.129	-0.123	-0.079	0.011	0.040	0.090	0.2398	
	t-value	55.68	-4.39	-8.09	-4.78	0.97	1.71	10.17		
4	Coeff.	2.954	0.101	0.003	-0.175	-0.016	-0.130	-0.035	0.4445	
	t-value	12.69	2.64	0.11	-6.94	-0.48	-3.38	-1.22		
5	Coeff.	1.977	2.370	0.143	0.027	-0.129	-0.462	0.234	0.3889	
	t-value	39.67	84.11	6.86	1.92	-13.74	-17.54	45.37		
6	Coeff.	3.770	0.343	-0.477	-0.225	-0.346	0.052	-0.121	0.2814	
	t-value	9.67	5.83	-8.12	-3.69	-7.01	0.57	-2.96		
7	Coeff.	1.607	0.118	0.079	-0.051	0.084	0.307	0.152	0.2849	
	t-value	9.67	2.86	2.22	-2.25	5.26	7.12	9.70		
8	Coeff.	1.655	0.229	0.021	-0.014	0.010	-0.016	0.036	0.276	
	t-value	32.41	16.43	2.18	-2.22	1.16	-1.53	5.91		
9	Coeff.	-0.479	0.438	-0.460	-0.120	0.174	0.275	0.377	0.4466	
	t-value	-0.45	3.35	-4.00	-1.42	1.12	2.62	3.38		
10	Coeff.	1.776	-0.198	-0.059	-0.030	-0.095	0.074	0.061	0.2264	
	t-value	63.61	-7.59	-4.57	-1.98	-9.34	3.73	9.14		
11	Coeff.	3.113	-0.402	-0.046	-0.805	-0.395	-0.677	-0.079	0.3914	
	t-value	5.14	-1.95	-0.28	-6.00	-3.55	-8.90	-1.26		
12	Coeff.	3.464	0.546	-0.187	-0.146	-0.054	-0.080	-0.066	0.2957	
	t-value	82.90	25.57	-12.93	-11.97	-6.39	-4.48	-10.58		

Panel B: Predicting Future Firm Cash Flows (CF_{t+1})										
Cluster		Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	Adj R-Sq	
1	Coeff.	-0.471	-0.0002	0.028	0.017	0.010	-0.016	0.066	0.4395	
	t-value	-7.50	-0.04	5.13	3.22	1.38	-3.23	11.41		
2	Coeff.	-0.064	0.067	0.017	0.049	0.006	-0.033	0.028	0.6154	
	t-value	-1.46	5.08	2.21	6.89	0.96	-6.41	5.94		
3	Coeff.	0.114	0.138	0.042	0.033	0.003	-0.036	0.004	0.5454	
	t-value	20.15	35.76	20.75	15.33	1.82	-11.72	3.11		
4	Coeff.	0.073	0.009	0.006	0.043	-0.015	-0.016	-0.026	0.2101	
	t-value	1.57	1.20	1.22	8.43	-2.23	-2.14	-4.64		
5	Coeff.	0.384	0.050	-0.035	0.003	-0.005	-0.029	-0.010	0.5807	
	t-value	87.02	19.86	-18.87	2.42	-6.41	-12.53	-22.01		
6	Coeff.	-0.148	0.043	-0.012	0.015	-0.043	0.024	0.046	0.3369	
	t-value	-3.53	6.84	-1.89	2.32	-8.10	2.43	10.48		
7	Coeff.	-0.001	-0.005	-0.020	0.005	0.008	-0.057	0.028	0.4149	
	t-value	-0.06	-0.91	-4.37	1.81	3.86	-10.03	13.56		
8	Coeff.	0.245	0.068	0.015	0.040	0.008	0.002	0.006	0.4794	
	t-value	16.58	16.92	5.19	21.45	3.33	0.54	3.61		
9	Coeff.	-0.199	0.010	0.044	0.023	0.008	-0.051	0.045	0.321	
	t-value	-2.61	1.09	5.34	3.84	0.73	-6.74	5.56		
10	Coeff.	0.053	0.049	0.042	0.003	0.017	-0.005	0.000	0.8103	
	t-value	7.38	7.26	12.73	0.88	6.44	-1.06	-0.28		
11	Coeff.	-0.223	0.170	-0.001	0.132	0.013	-0.049	0.038	0.6283	
	t-value	-3.11	6.98	-0.03	8.29	1.00	-5.50	5.11		
12	Coeff.	0.086	0.067	0.009	0.036	0.004	-0.007	0.024	0.5933	
	t-value	17.98	27.20	5.51	25.82	4.33	-3.46	33.75		

Table 4.5 (Continued)**Panel C: Predicting Future Firm Earnings ($EARN_{t+1}$)**

Cluster		Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LSIZE	Adj R-Sq
1	<i>Coeff.</i>	-0.285	0.000	0.068	-0.002	0.009	-0.009	0.067	0.5544
	<i>t-value</i>	-5.58	0.07	15.05	-0.39	1.42	-2.13	14.28	
2	<i>Coeff.</i>	-0.118	0.076	0.046	0.055	0.008	-0.016	0.021	0.4231
	<i>t-value</i>	-3.21	6.93	7.13	9.17	1.45	-3.73	5.36	
3	<i>Coeff.</i>	0.047	0.136	0.021	0.010	0.007	-0.006	0.008	0.2952
	<i>t-value</i>	10.72	45.26	13.40	5.67	5.47	-2.72	8.77	
4	<i>Coeff.</i>	-0.002	-0.006	0.009	0.016	-0.007	-0.003	0.020	0.2895
	<i>t-value</i>	-0.05	-0.97	2.33	4.15	-1.29	-0.58	4.77	
5	<i>Coeff.</i>	0.303	0.076	0.023	0.017	-0.005	0.012	-0.007	0.2731
	<i>t-value</i>	106.24	47.06	18.93	21.15	-8.71	7.88	-25.02	
6	<i>Coeff.</i>	-0.206	0.030	0.062	0.044	0.024	0.045	0.046	0.4316
	<i>t-value</i>	-5.74	5.52	11.45	7.94	5.37	5.38	12.11	
7	<i>Coeff.</i>	-0.112	-0.008	0.002	-0.003	0.011	-0.036	0.025	0.3672
	<i>t-value</i>	-7.11	-1.95	0.69	-1.41	7.42	-8.89	16.90	
8	<i>Coeff.</i>	0.075	-0.008	0.003	-0.001	0.004	-0.001	0.007	0.2842
	<i>t-value</i>	10.61	-4.03	2.40	-1.32	3.77	-0.93	8.60	
9	<i>Coeff.</i>	-0.256	0.022	0.057	0.010	0.022	-0.036	0.041	0.4002
	<i>t-value</i>	-3.89	2.68	7.99	1.94	2.27	-5.58	5.97	
10	<i>Coeff.</i>	0.049	0.087	0.007	-0.001	0.012	0.004	0.003	0.1781
	<i>t-value</i>	11.56	22.04	3.84	-0.62	7.98	1.31	2.93	
11	<i>Coeff.</i>	0.015	0.139	-0.020	0.083	0.024	-0.018	0.021	0.3744
	<i>t-value</i>	0.22	6.17	-1.10	5.62	1.96	-2.15	3.05	
12	<i>Coeff.</i>	0.063	0.094	0.023	0.025	-0.001	-0.001	0.021	0.3073
	<i>t-value</i>	16.15	47.18	17.28	21.89	-1.35	-0.84	35.86	

This table reports the results of the standard clustered error OLS regressions of future firm value (Q_{t+1}), future firm cash flows (CF_{t+1}), and future firm earnings ($EARN_{t+1}$) in year $t+1$. The independent variables are six clustering criteria variables (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE). The control variables include ASSETG, LBTM, NegY, YB, NegAccr, PosAccr, NoDiv, Div_B, Lagret, AGE, CVOL, and SG. The clustering criteria and control variables are measured in year t and defined in appendix 1. The table presents the results of predictive regressions of dependent variables across 12 clusters and show only the coefficients and t -statistics of the clustering criteria. The control variable results are not included in the interests of brevity. The predictive regressions for Q_{t+1} , CF_{t+1} , and $EARN_{t+1}$ are given in Panels A, B, and C respectively.

4.3.5.2 Predicting future cash flows

Panel B in Table 4.5 summarises the results for the estimating equation (4.2) with future cash flows as the dependent variable. FIN_FLEX in eight out of 12 clusters has a significant positive relationship with CF_{t+1} . In clusters 1, 4, 7 and 9 there is no significant relation observed. The coefficients for ST_CREDIT – CF_{t+1} relation are significantly positive in seven clusters. On the other hand, two clusters have a significant negative relationship with

future cash flows. The $LT_INV - CF_{t+1}$ association is positive in all clusters and among them 10 are significant. The results suggest that an increase in long-term investment will generally be translated into positive future cash flows.

For the $CVT_DEBT - CF_{t+1}$ relation, Table shows four significantly positive and two significantly negative coefficients. Then in the case of $PSK_USE - CF_{t+1}$, nine out of 12 clusters have a significantly negative relationship. Finally, in $LSIZE - CF_{t+1}$ relation, nine out of 12 clusters are significantly positive and three clusters are negatively related (two of them are significant). Again the R-squares in these cash flow models are greater than 0.21.

4.3.5.3 Predicting future earnings

Table 4.5 in Panel C summarises the results for estimating equation (1) to predict future earnings from operations one year ahead across the 12 clusters. The change in financial flexibility (FIN_FLEX) is found to predict $EARN_{t+1}$ significantly in 10 out of 12 clusters. One interesting outcome is that the firms showing negative changes in financial flexibility have a strong positive relationship – with large coefficients and t -statistics – with future firm earnings. The change in short-term credit (ST_CREDIT) predicts the $EARN_{t+1}$ positively across all clusters except cluster number 11. Cluster 11 is comprised of firms that show a statistically zero increase in short-term credit, in median size and in early growth stage. The findings suggest that on aggregate the short-term credit increase decisions would increase future earnings in most of the clusters.

Then, LT_INV a latent measure of changes in long-term investment predicts robustly the future firm earnings. They have a positive relationship in seven out of 12 clusters. Interestingly, the cluster firms with the most positive change in long-term investment show an insignificant LT_INV – EARN_{t+1} relationship. Next, the CVT_DEBT – EARN_{t+1} association is robust in eight out of 12 cluster regression models. The CVT_DEBT regression coefficients for cluster models 1, 2, 4 and 12 are statistically insignificant. The only robust negative CVT_DEBT – EARN_{t+1} association exists between cluster 5 firms that are larger in size, pay higher in dividends, and that reduce convertible debt marginally. The PSK_USE – EARN_{t+1} relationship is prominently negative except in cluster models 5, 6 and 10. Among 12 cluster models, only the PSK_USE – EARN_{t+1} relationship is insignificant across cluster models 4, 8, and 10. Any increase or decrease in the preferred stock issuance would translate into reduced earnings in future.

Finally, the LSIZE predicts strong positive future earnings. The only cluster with a significant negative association is cluster 5. The cluster 5 firms being larger in size with fewer growth opportunities in hand, use free cash flows to pay dividends and do not take positive NPV projects to make positive future earnings. The cluster model 10 has the least adjusted R-squared of 0.18 and cluster 1 has the highest adjusted R-squared of 0.55.

4.4 Conclusion

How do firms differ in their corporate finance decisions? In particular how can they be analysed and evaluated efficiently? The question is not well addressed using usual linear models because, the relations of corporate finance decisions on the one hand are not common

among all firms, and on the other hand, they vary in priority and composition in different real business situations. This means firms practice heterogeneous business styles. And, if that is so, then the answer requires a method of finding homogeneous firm groups. This can be done by grouping firms into their similar business styles before taking on any financial analysis.

To answer the question, this essay discriminates the firms into their internally homogeneous and externally heterogeneous groups based on common corporate financing decisions and in particular, firm size. The common corporate finance decisions are the five factors that are extracted via application of the factor analysis on the balance sheet change rates (see chapter 2). These include financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-term investment (LT_INV), convertible debt usage (CVT_DEBT), and preferred stock usage (PSK_USE).

This study uses K-means clustering technique to cluster firms. As a result it obtains 12 possibly optimal and natural clusters of firms. All clusters differ in their unique clustering criteria, and a detailed description of a long list of financial attributes is provided.

This study finds that at one point in time, firms in two clusters (clusters 1 and 9) with the most positive changes in financial flexibility, have higher Tobin's Q, pay fewer dividends and have higher R&D investment intensity. Similarly, there is another set of two clusters (clusters 3 and 10) with the most negative change in financial flexibility, and their attributes are exactly at opposite extremes relative to the first group. On the one hand, the firms in the group are found to differ in their behaviour in dealing with changes in hybrid securities like preferred stock and convertible debt. On the other hand, clusters 1 and 9 mimic high IST beta

firms, but differ in firm size, investment rate and firm age. The observation is relative to the recent literature on growth opportunities, technology shocks and asset prices (Kogan and Papanikolaou, 2014). This study also finds that clusters 3 and 10 mimic low IST beta firms. KP's IMC portfolio is similar to the FLEX factor (from chapter 3), but differs on a few nuances. Both are the important source of risk and bear a risk premium that explains or gives rise to the value factor due to the heterogeneous growth opportunities of stocks in both high and low IST (FIN_FLEX) stocks.

The clustering classification is not consistent with prior classifications based on the SIC codes. Particularly, the 10 industries – as classified based on the SIC codes defined in the Fama and French online library – are represented in all 12 clusters. However, most firms are classified in the cluster dominated by business equipment and health industry firms. This evidence entails that these clustering based firm groups are the best representative of their corporate and financial decision-making rather than their industry or product market outputs.

Apart from the six clustering criteria, this study has included firm age as another important attribute that should be used to classify the firms. It finds firm age, and firm size are not linear; and firm age, and log book-to-market ratio are not linear either. The firms based on the clustering criteria are clustered in such a way that not all large firms are mature firms, and not all young firms are small firms. Due to this particular phenomenon, this study finds very young firms are often associated with very low log book-to-market ratios, although they are not large in size.

Finally, this essay tests the predictive power of the clustering criteria to predict future firm value, cash flows, and earnings, while placing extensive controls. The results suggest that the clustering attributes are good predictors of these variables. However, the signs of relationship may vary cluster to cluster; an indication of different relational coefficients in all-firms' regressions, which indicates that the simple linear all-firms' regression is not valid. Thus clustering first followed by simple linear regression within each cluster is a better method for general business analysis and firm valuation.

This study believes that the trial of clustering classification is very successful. This is an important contribution to the general business analysis methodology. Business operations and financial investment decisions are much more complicated. The simple linear relation for all firms is not acceptable. The idea of first classifying firms into different styles may possibly influence much previous published research and improve our understanding of the relative theories and practice in the future.

CHAPTER FIVE

CONCLUSION

This chapter concludes the thesis by providing a summary of the major findings and implications for each of the three essays in Section 5.1. Section 5.2 suggests potential areas for future research.

5.1 Major Findings and Implications

5.1.1 Essay One: Latent accounting growth, corporate financing decisions, and return predictability

The first essay examines the balance sheet change rates of US common stocks from 1985 to 2009 to explore whether firms display systematic shifts in their corporate finance decisions. These decisions are embedded in various balance sheet accounts across time. Consequently they are not directly observed, and it remains unclear how many fundamental (common) factors are sufficient to represent the time-varying operating business activities.

This study accomplishes the task of obscuring the hidden operating business decisions using a factor analysis model, which is a linear data reduction technique. This model reduces the input of 37 balance sheet variables to the five unobserved (latent) factors that closely fit the theoretical (or logical) category of decision types for financial business management. The finding of the five factors makes the first major contribution to the accounting literature. These factors include: financial flexibility (FIN_FLEX), short-term credit (ST_CREDIT), long-term capital investment (LT_INV), convertible debt usage (CVT_DEBT) and preferred

stock usage (PSK_USE). These categories are typical topics in corporate finance textbooks. This close resemblance suggests that these factors are truly underlying mechanisms.

This study also establishes that these orthogonal factors provide the true decomposition of the total asset growth (ASSETG) effect. This is supported by evidence that the latent factors can capture 83% of the variations of ASSETG, and only two factors (FIN_FLEX and LT_INV) are sufficient to capture the return predictability contained in ASSETG. The new factors are also able to capture 64% of the variations of investment-to-assets (I_A), and 48% of the net operating assets (NOA). They offer additional information to the competing prominent investment, growth, and accrual measures in the related literature.

This essay finds that the new factors have significantly negative relation to subsequent stock returns. In terms of economic significance, a 1% point decrease in FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT factors increases future excess annual returns by about 3.2%, 2.4%, 2.9%, and 1%, respectively. Similar evidence is found in the portfolio setting. These factors are able to earn substantial incremental returns over and above the prominent return determinants and asset pricing factors. The abnormal returns can be obtained at least up to three years ahead. In terms of annual buy-and-hold returns, the first three factors, namely FIN_FLEX, ST_CREDIT, and LT_INV are able to earn 17%, 2.3%, and 10.8% returns over the whole sample at the equal-weighted basis.

This study also shows that when firms engage in operating business activities, the prime objective is the increase in firm value. Correspondingly, this essay seeks to find the impact of the unobserved new operating decision measures on firms' value. The robust-

cluster standard error regressions and portfolios of high (low) Tobin's Q in this research tests the implications of the Q theory of investment to the five unobserved factor decisions. The supporting evidence suggests that except for CVT_DEBT, the other latent factors are consistent with the Q theory of investment. The latent factors can predict firm value (Tobin's Q) for up to three years, and thus are value-creating activities. This finding implies that managers are able to make the right operating business judgments about firms' current status and future investment opportunities. Apart from firm value forecasting tests, this essay provides evidence that new factors are also related to future cash flows and earnings. In predictive cluster robust standard error regressions, the evidence suggests that the new factors can forecast both future cash flows and future earnings. This forecasting ability is particularly strong with cash flow rather than earnings.

5.1.2 Essay Two: Financial flexibility: A new risk factor?

The second essay examines the financial flexibility (FIN_FLEX) characteristic to answer the question of whether information relating to this characteristic is priced or not. For such a test, the FIN_FLEX mimicking return portfolio (FLEX) (i.e., return spread between low and high financial flexibility) carries a significant positive premium. The average monthly premium on the FLEX factor is 0.69% with *t*-statistics of 5.56. The test statistic value is much larger than the new conservative cut-off of 3.0 suggested by Harvey, et al. (2013). The premium also persists in the two-stage cross-section regressions (2SCSR), Fama-MacBeth (1973) regressions, and characteristics versus covariances test of Daniel and Titman (1997).

These findings suggest that FLEX is a priced factor and contradicts the mispricing or characteristics based explanations. The FLEX is able to survive the controls of its own characteristics (FIN_FLEX) and control of the five prominent risk factors: MKT_RF, SMB, HML, UMD, and LIQ, is thus non-redundant. The average return unexplained by FLEX's exposure to these five risk factors is about 0.60% per month with *t*-statistics of 5.27. In contrast, the non-redundancy evidence shows that the FLEX is able to capture the risk inherent in the HML. The average HML return is subsumed by the exposures of HML to the FLEX factor, where intercepts become negative -0.30% with *t*-statistics of -1.78, which is significant at the 10% level. In related literature, Fama and French (2013) show a similar finding where profitability based hedge portfolio RMW and physical investment based hedge portfolio CMA makes HML redundant. On the contrary, single FLEX does this job well. Moreover, in further supportive evidence, this essay shows that FIN_FLEX is able to subsume the average return predictability associated with both gross profitability and physical investment characteristics. Thus, this study assumes that FLEX would also be non-redundant to both CMA and RMW.

The possible explanation for the pricing effect of FLEX either lies in the unobserved nature of the underlying characteristics (FIN_FLEX) or the basic construction of the factor. The FLEX is not only hedge of low (high) FIN_FLEX portfolios, however it is the hedge of low (high) Tobin's Q, and high (low) leverage. In fact, the firms with a high FIN_FLEX portfolio tend to have a higher Tobin's Q, higher R&D intensity, lower book-to-market ratio, almost zero leverage, pay nearly no dividends, and are above median size relative to low FIN_FLEX firms. This essay in further robustness checks, shows that the FIN_FLEX based return predictability is independent of the firm's information environment. Overall, this

research suggests that investors react rationally to the arrival of changes in financial flexibility information, thus FLEX captures the managers' perceptions of the future investment opportunities and in turn the risk inherent in firms' future cash flows.

5.1.3 Essay Three: Latent accounting growth, firm size, and firm grouping dynamics

The final essay of this thesis emphasises the role of the common corporate financing decisions identified in the first essay to segment US listed firms. Firms are partitioned in their natural groups that are unknown *a priori*. Compared to existing firm grouping methods, for example SIC, NAICS, and GICS, this is a distinct exercise because prior research uses product-market outputs to categorise the firms into the industries. In contrast, this essay employs the common business decisions and firm size as a clustering criteria. This study argues that corporate finance decisions reflect the firms' financial health and managers' perceptions instead of the industry owing to the nonlinear and firm-specific nature of the criteria.

In particular, this essay addresses the question of: how do firms differ in their corporate finance decisions? The evidence based on the application of K-means clustering suggests that there exist 12 clusters of firms reflecting the various business management styles. The clustering criterion is homogeneous within the cluster and heterogeneous across the clusters. Thus, the main criterion for each of the 12 clusters is different. The findings further suggest that the choice of making positive shifts or negative shifts among the five corporate finance decisions mainly depends on the firm's information uncertainty, age, financial constraints, growth opportunities, and leverage levels. This clustering exercise also

adds to the literature on portfolio diversification. The 12 clusters are well-diversified across the 10 SIC code based industries. The clusters with higher average returns are the ones that show the most negative shift in the five corporate finance decisions.

This essay examining the individual clusters suggests that this classification method successfully identifies the nonlinear relationship between firm size and age; and the size and book-to-market ratio. The firm age appears to be the missing link that should also be made part of future firm groupings. Moreover, this research complements the literature on growth opportunities, investment-specific technology and asset prices (Kogan and Pananikolaou, 2014). The evidence suggests that the latent factors are underlying mechanisms for firms to exploit growth opportunities. This study finds that using these mechanisms, the high (low) IST beta portfolios given in KP (2014) can naturally be partitioned into four clusters. Clusters 1 and 9 are consistent with the high IST beta portfolio, and clusters 3 and 10 are consistent with the low IST beta stock portfolio. However, the average attributes of these clusters in this research differ from high (low) IST beta portfolios in firm size, age, and investment rates in physical capital.

The important contribution of this research is the examination of five operating business decisions in forecasting performance measures. This essay finds that the signs of the five corporate finance decisions in forecasting future firm value, earnings, and cash flows vary from cluster to cluster. This effect may not be directly observed from the full sample, possibly due to the influence of firm size. Overall, the K-means clustering technique is able to incorporate the dynamic changes in the clustering criteria.

5.2 Future Areas of Research

The first essay in this thesis provides the foundation for a new line of research to commence in the areas of asset pricing, financial accounting, and corporate finance. This new perspective on data analysis and identifying the unobserved measures of operating business activities from the balance sheets would improve business analysis. This style of analysis can be extended to other accounting and market data to unearth interesting patterns. To this end, powerful pattern recognition tools such as factor analysis, principal components, cluster analysis and discriminant analysis have been developed and have been applied successfully to other disciplines. There is a need for more attention in the finance profession to apply these numerical methods in financial analysis.

First of all, the existence of the five latent common corporate finance decisions requires their authentication in out-of-sample tests. The sample of developed G-7 economies would be ideal for an out-of-sample test as these countries provide rich accounting information. If these decisions do exist, then it would be interesting to observe their order of importance, and their impact on the accounting and market values of a firm. A similar analysis could be done to identify the industry-specific common factors. This exercise would reflect whether the overall sample effect is influenced by one particular industry, or is consistent across industries. In addition, the researchers could also extract the common corporate finance decisions specific to the size grouping. Then the firm size specific latent factors could be further tested for their return predictability and fundamental value creation findings.

The second essay of this thesis shows that the return forecasting capability of FIN_FLEX (shift in financial flexibility) is driven by the covariances and not characteristics.

These tests could be repeated to explore the sources of return predictability for the other four factors (ST_CREDIT, LT_INV, CVT_DEBT, and PSK_USE). This exercise would reflect whether factor decisions mimic risk or mispricing. Alternatively, we would understand which information content is rationally incorporated into prices. Future research may also need to explain the relationship between FLEX and macro factors; that is, whether it mimics some state variables or not. Further, if FLEX is a risk factor, then how expected returns on it would vary in relation to changes in the investors' fear factor (Volatility Index). In addition, research needs to be undertaken to answer questions like: does FLEX persist as a non-redundant factor when regressed on RMW (the hedge return of the robust and weak profitability based diversified portfolios of stocks) and CMA (the hedge return of the low and high investment based diversified portfolios of stocks)? Similarly, does FLEX truly captures HML factor? And, is FLEX premium sensitive to factor sorting variability such as independent, joint, and sequential sorts? These are only a few of several possibilities for future research to be undertaken just using the first factor (FIN_FLEX).

The final essay implies that both academics and investors need to understand the diversification effects of the clustering using corporate finance decisions as clustering criteria. New investment strategies involving a combination of 12 clusters could be devised to earn incremental and substantial long-run returns. Future studies could examine why clustering on common corporate financing decisions ends up with a cluster with varying growth opportunities and innovation intensity, and what role firm age may play in improving firm grouping dynamics. Consequently, this may shed further light on the joint determination of investment behaviour and risk premia.

REFERENCES

- Abarbanell, J. S., and Bushee, B. J. (1997). Fundamental Analysis, Future Earnings, and Stock Prices. *Journal of Accounting Research*, 35(1), 1-24.
- Abarbanell, Jeffery S., Bushee, Brian J., and Raedy, J. S. (2003). Institutional Investor Preferences and Price Pressure: The Case of Corporate Spin-Offs. *The Journal of Business*, 76(2), 233-261.
- Aboody, D., Hughes, J., and Liu, J. (2005). Earnings quality, insider trading, and cost of capital. *Journal of Accounting Research*, 43(5), 651-673.
- Allayannis, G., and Weston, J. P. (2001). The use of foreign currency derivatives and firm market value. *Review of Financial Studies*, 14(1), 243-276.
- Almeida, H., Campello, M., and Weisbach, M. S. (2011). Corporate financial and investment policies when future financing is not frictionless. *Journal of Corporate Finance*, 17(3), 675-693.
- Anderson, C. W., and Garcia-Feijóo, L. (2006). Empirical Evidence on Capital Investment, Growth Options, and Security Returns. *Journal of Finance*, 61(1), 171-194.
- Ang, J., and Smedema, A. (2011). Financial flexibility: Do firms prepare for recession? *Journal of Corporate Finance*, 17(3), 774-787.
- Armstrong, C. S., Core, J. E., Taylor, D. J., and Verrecchia, R. E. (2011). When does information asymmetry affect the cost of capital? *Journal of Accounting Research*, 49(1), 1-40.
- Baker, M., Stein, J. C., and Wurgler, J. (2003). When does the market matter? Stock prices and the investment of equity-dependent firms. *The Quarterly Journal of Economics*, 118(3), 969-1005.
- Baker, M., and Wurgler, J. (2002). Market timing and capital structure. *The Journal of Finance*, 57(1), 1-32.
- Baker, M., and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4), 1645-1680.
- Barclay, M. J., and Smith, C. W. (1995a). The Maturity Structure of Corporate Debt. *The Journal of Finance*, 50(2), 609-631.
- Barclay, M. J., and Smith, C. W. (1995b). The priority structure of corporate liabilities. *The Journal of Finance*, 50(3), 899-917.

- Bateson, G. (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*: University of Chicago Press.
- Baxamusa, M. (2011). How well do market timing, catering, and classical theories explain corporate decision? *Journal of Financial Research*, 34(2), 217-239.
- Bernard, V., and Thomas, J. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13, 305-340.
- Berk, J. B., Green, R. C., and Naik, V. (1999). Optimal Investment, Growth Options, and Security Returns. *Journal of Finance*, 54(5), 1553-1607.
- Bhojraj, S., Lee, C. M. C., and Oler, D. K. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research*, 41(5), 745-774.
- Black, F., Jensen, M. C., and Scholes, M. S. (1972). The Capital Asset Pricing Model: Some Empirical Tests. *Michael C. Jensen, Studies in the theory of capital markets, Praeger Publishers Inc., 1972.*
- Bodie, Z., and Taggart, R. A. (1978). Future investment opportunities and the value of the call provision on a bond. *The Journal of Finance*, 33(4), 1187-1200
- Bolton, P., and Scharfstein, D. S. (1990). A theory of predation based on agency problems in financial contracting. *The American Economic Review*, 93-106.
- Bonaimé, A. A., Hankins, K. W., and Harford, J. (2013). Financial flexibility, risk management, and payout choice. *Review of Financial Studies*, hht045.
- Bradshaw, M. T., Richardson, S. A., and Sloan, R. G. (2006). The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics*, 42(1-2), 53-85.
- Brander, J. A., and Lewis, T. R. (1986). Oligopoly and financial structure: The limited liability effect. *The American Economic Review*, 956-970.
- Brennan, M. J., and Schwartz, E. S. (1984). Optimal financial policy and firm valuation. *The Journal of Finance*, 39(3), 593-607.
- Brennan, M. J., Wang, A. W., and Xia, Y. (2004). Estimation and test of a simple model of intertemporal capital asset pricing. *Journal of Finance*, 59(4), 1743-1776.
- Brown, S. J., and Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics*, 43(3), 373-399. doi: 10.1016/s0304-405x(96)00898-7

- Bushee, B. J. (1998). The Influence of Institutional Investors on Myopic R&D Investment Behavior. *The Accounting Review*, 73(3), 305-333.
- Butler, A. W., Cornaggia, J., Grullon, G., and Weston, J. P. (2011). Corporate financing decisions, managerial market timing, and real investment. *Journal of Financial Economics*, 101(3), 666-683.
- Campbell, J. Y., and Vuolteenaho, T. (2004). Bad Beta, Good Beta. *The American Economic Review*, 94(5), 1249-1275.
- Cao, S. (2011). *The total asset growth anomaly: is it incremental to the net operating asset growth anomaly?* University of Illinois working paper.
- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1), 57-82.
- Chan, L. K. C., Lakonishok, J., and Swaminathan, B. (2007). Industry Classifications and Return Comovement. *Financial Analysts Journal*, 63(6), 56-70.
- Chen, L., Novy-Marx, R., and Zhang, L. (2011). An alternative three-factor model. *Available at SSRN 1418117*.
- Chung, K. H. (1993). Asset characteristics and corporate debt policy: an empirical test. *Journal of Business Finance & Accounting*, 20(1), 83-98.
- Clarke, R. N. (1989). SICs as Delineators of Economic Markets. *The Journal of Business*, 62(1), 17-31.
- Cochrane, J. H. (2001). *Asset pricing*: Princeton University Press, Princeton, NJ (2001).
- Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4), 1047-1108.
- Connor, G. (1995). The Three Types of Factor Models: A Comparison of Their Explanatory Power. *Financial Analysts Journal*, 51(3), 42-46.
- Cooper, M. J., Gulen, H., and Schill, M. J. (2008). Asset Growth and the Cross-Section of Stock Returns. [Article]. *Journal of Finance*, 63(4), 1609-1651.
- Core, J. E., Guay, W. R., and Verdi, R. (2008). Is accruals quality a priced risk factor? *Journal of Accounting and Economics*, 46(1), 2-22.
- Dammon, R. M., and Senbet, L. W. (1988). The effect of taxes and depreciation on corporate investment and financial leverage. *The Journal of Finance*, 43(2), 357-373.
- Daniel, K., and Titman, S. (1997). Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance*, 52(1), 1-33.

- Daniel, K., and Titman, S. (2006). Market Reactions to Tangible and Intangible Information. *Journal of Finance*, 61(4), 1605-1643. doi: 10.1111/j.1540-6261.2006.00884.x
- Daniel, K. D., Hirshleifer, D., and Subrahmanyam, A. (2001). Overconfidence, Arbitrage, and Equilibrium Asset Pricing. *The Journal of Finance*, 56(3), 921-965.
- DeAngelo, H., and DeAngelo, L. (2007). Capital structure, payout policy, and financial flexibility. *Marshall School of Business Working Paper No. FBE*, 02-06.
- Dechow, P. M., Richardson, S. A., and Sloan, R. G. (2008). The persistence and pricing of the cash component of earnings. *Journal of Accounting Research*, 46(3), 537-566.
- DeFond, M. L., and Jiambalvo, J. (1994). Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics*, 17(1), 145-176.
- DeLong, G. L. (2001). Stockholder gains from focusing versus diversifying bank mergers. *Journal of Financial Economics*, 59(2), 221-252.
- Denis, D. J. (2011). Financial flexibility and corporate liquidity. *Journal of Corporate Finance*, 17(3), 667-674.
- Denis, D. J., and McKeon, S. B. (2012). Debt financing and financial flexibility evidence from proactive leverage increases. *Review of Financial Studies*, 25(6), 1897-1929.
- Diamond, D. W. (1991). Debt maturity structure and liquidity risk. *The Quarterly Journal of Economics*, 106(3), 709-737.
- Dixit, A. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy*, 620-638.
- Dotan, A., and Ravid, S. A. (1985). On the interaction of real and financial decisions of the firm under uncertainty. *The Journal of Finance*, 40(2), 501-517.
- Doukas, J. A., and Pantzalis, C. (2003). Geographic diversification and agency costs of debt of multinational firms. *Journal of Corporate Finance*, 9(1), 59-92.
- Duarte, J., and Young, L. (2009). Why is PIN priced? *Journal of Financial Economics*, 91(2), 119-138.
- Easley, D., and O'hara, M. (2004). Information and the cost of capital. *The Journal of Finance*, 59(4), 1553-1583.
- Easterwood, J. C., and Kadapakkam, P. R. (1994). Agency conflicts, issue costs, and debt maturity. *Quarterly Journal of Business and Economics*, 69-80.
- Elton, E. J., and Gruber, M. J. (1970). Homogeneous Groups and the Testing of Economic Hypotheses. *The Journal of Financial and Quantitative Analysis*, 4(5), 581-602.

- Elton, E. J., and Gruber, M. J. (1971). Improved Forecasting Through the Design of Homogeneous Groups. *The Journal of Business*, 44(4), 432-450.
- Elton, E. J., Gruber, M. J., and Green, T. C. (2007). The Impact of Mutual Fund Family Membership on Investor Risk. [Article]. *Journal of Financial & Quantitative Analysis*, 42(2), 257-277.
- Ericsson, J. (2000). Asset substitution, debt pricing, optimal leverage and maturity. *Finance* 21, 39-70.
- Everitt, B., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster analysis* (Vol. 5th): West Sussex, UK: John Wiley & Sons.
- Fabozzi, F. J., and Markowitz, H. M. (2011). *The Theory and Practice of Investment Management: Asset Allocation, Valuation, Portfolio Construction, and Strategies* (Vol. 198): Wiley.
- Fairfield, P. M., Whisenant, J. S., and Yohn, T. L. (2003). Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing. *The Accounting Review*, 78(1), 353-371.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F., and French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.
- Fama, E. F., and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.
- Fama, E. F., and French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, 51(1), 55-84.
- Fama, E. F., and French, K. R. (2005). Financing decisions: who issues stock? *Journal of Financial Economics*, 76(3), 549-582.
- Fama, E. F., and French, K. R. (2008). Dissecting Anomalies. *Journal of Finance*, 63(4), 1653-1678.
- Fama, E. F., and French, K. R. (2013). A five-factor asset pricing model. *Fama-Miller working paper, University of Chicago, Dartmouth College and NBER*.
- Fama, E. F., and MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), 607-636.

- Fan, Joseph P. H., and Lang, Larry H. P. (2000). The Measurement of Relatedness: An Application to Corporate Diversification. *The Journal of Business*, 73(4), 629-660.
- Faulkender, M., and Wang, R. (2006). Corporate financial policy and the value of cash. *The Journal of Finance*, 61(4), 1957-1990.
- Fazzari, S., Hubbard, R. G., and Petersen, B. C. (1988). Financing constraints and corporate investment: National Bureau of Economic Research Cambridge, Mass., USA.
- Fernandez, P. (2014). CAPM: An Absurd Model. Available at SSRN 2505597.
- Flannery, M. J., and Rangan, K. P. (2006). Partial adjustment toward target capital structures. *Journal of Financial Economics*, 79(3), 469-506.
- Francis, J., LaFond, R., Olsson, P., and Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics*, 39(2), 295-327.
- Frank, M. Z., and Goyal, V. K. (2003). Testing the pecking order theory of capital structure. *Journal of Financial Economics*, 67(2), 217-248.
- Fries, S., Miller, M., and Perraudin, W. (1997). Debt in industry equilibrium. *Review of Financial Studies*, 10(1), 39-67.
- Gamba, A., and Triantis, A. (2008). The value of financial flexibility. *Journal of Finance*, 63(5), 2263-2296.
- Goetzmann, W. N., and Wachter, S. M. (1995). Clustering Methods for Real Estate Portfolios. *Real Estate Economics*, 23(3), 271-310.
- Goldstein, R., Ju, N., and Leland, H. (2001). An EBIT-Based Model of Dynamic Capital Structure*. *The Journal of Business*, 74(4), 483-512.
- Gordon, A. (1999). *Classification*. 1999. Chapman&Hall, CRC, Boca Raton, FL.
- Gow, I. D., Ormazabal, G., and Taylor, D. J. (2010). Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review*, 85(2), 483-512.
- Graham, J. R., and Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, 60(2), 187-243.
- Grauer, R. R., and Janmaat, J. A. (2009). On the power of cross-sectional and multivariate tests of the CAPM. *Journal of banking & Finance*, 33(5), 775-787.
- Grauer, R. R., and Janmaat, J. A. (2010). Cross-sectional tests of the CAPM and Fama-French three-factor model. *Journal of banking & Finance*, 34(2), 457-470.

- Gray, P., and Johnson, J. (2011). The relationship between asset growth and the cross-section of stock returns. *Journal of banking & Finance*, 35(3), 670-680.
- Grinold, R., Rudd, A., and Stefek, D. (1989). Global factors: Fact or fiction? [Article]. *Journal of Portfolio Management*, 16(1), 79-88.
- Guenther, D. A., and Rosman, A. J. (1994). Differences between COMPUSTAT and CRSP SIC codes and related effects on research. *Journal of Accounting and Economics*, 18(1), 115-128.
- Gunny, K. A. (2005). What are the consequences of real earnings management? : University of California, Berkeley.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L., and Black, W. C. (1998). *Multivariate Data Analysis*. New Jersey: Prentice-Hall, Inc.
- Hartigan, J. A. (1975). *Clustering algorithms*. New York: John Wiley.
- Harvey, C., Liu, Y., and Zhu, H. (2013). ... and the cross-section of expected returns. Available at SSRN 2249314.
- Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, 56(4), 1533-1597.
- Hirshleifer, D., Kewei, H., Teoh, S. H., and Yinglei, Z. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, 38, 297-331.
- Hoberg, G., and Phillips, G. (2010a). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, 23(10), 3773-3811.
- Hoberg, G., and Phillips, G. (2010b). Text-Based Network Industries and Endogenous Product Differentiation: National Bureau of Economic Research.
- Holthausen, R. W., and Larcker, D. F. (1992). The prediction of stock returns using financial statement information. *Journal of Accounting and Economics*, 15(2-3), 373-411.
- Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, 1127-1150.
- Hou, K., and Robinson, D. T. (2006). Industry concentration and average stock returns. *The Journal of Finance*, 61(4), 1927-1956.
- Hou, K., Xue, C., and Zhang, L. (2014a). Digesting Anomalies: An Investment Approach. *Review of Financial Studies*. doi: 10.1093/rfs/hhu068

- Hou, K., Xue, C., and Zhang, L. (2014b). *A Comparison of New Factor Models* (No. w20682). National Bureau of Economic Research.
- Houston, A. L., Jr., and Carol Olson, H. (1990). Financing with Preferred Stock. *Financial Management*, 19(3), 42-54.
- Hovakimian, A., Opler, T., and Titman, S. (2001). The Debt-Equity Choice. *The Journal of Financial and Quantitative Analysis*, 36(1), 1-24.
- Howe, J. S., and Lee, H. (2006). The long-run stock performance of preferred stock issuers. *Review of Financial Economics*, 15(3), 237-250.
- Huang, D., and Wang, F. (2009). Cash, investments and asset returns. *Journal of banking & Finance*, 33(12), 2301-2311.
- Hubbard, R. G. (1997). Capital-market imperfections and investment: National Bureau of Economic Research.
- Huber, P. J. (1967). *The behavior of maximum likelihood estimates under nonstandard conditions*. Paper presented at the Proceedings of the fifth Berkeley symposium on mathematical statistics and probability.
- Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, 323-329.
- Jensen, M. C., and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305-360.
- Jensen, R. E. (1971). A cluster analysis study of financial performance of selected business firms. *Accounting Review*, 36-56.
- Jobson, J. D. (1992a). *Applied multivariate data analysis* (Vol. 2): Springer Nueva York.
- Jobson, J. D. (1992b). *Applied Multivariate Data Analysis Volume II: Categorical and Multivariate Methods* (Vol. II). New York: Springer-Verlag.
- Ju, N., Parrino, R., Poteshman, A. M., and Weisbach, M. S. (2005). Horses and rabbits? Trade-off theory and optimal capital structure. *Journal of Financial and Quantitative Analysis*, 40(2), 259-281.
- Kahle, K. M., and Walkling, R. A. (1996). The Impact of Industry Classifications on Financial Research. *Journal of Financial and Quantitative Analysis*, 31(03), 309-335.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

- Kahneman, D., and Tversky, A. (2000). *Choices, values, and frames*: Cambridge University Press.
- Kendall, M. G., and Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1), 11-34.
- Kim, D., and Qi, Y. (2010). Accruals quality, stock returns, and macroeconomic conditions. *The Accounting Review*, 85(3), 937-978.
- King, B. F. (1966). Market and Industry Factors in Stock Price Behavior. *The Journal of Business*, 39(1), 139-190.
- Kogan, L., and Papanikolaou, D. (2014). Growth opportunities, technology shocks, and asset prices. *The Journal of Finance*, 69(2), 675-718.
- Krishnan, J., and Press, E. (2003). The North American Industry Classification System and Its Implications for Accounting Research. [Article]. *Contemporary Accounting Research*, 20(4), 685-717.
- Lam, F., and Wei, K.-C. (2011). Limits-to-arbitrage, investment frictions, and the asset growth anomaly. *Journal of Financial Economics*, 102(1), 127-149.
- Leary, M. T., and Roberts, M. R. (2005). Do firms rebalance their capital structures? *The Journal of Finance*, 60(6), 2575-2619.
- Lehmann, B. N., and Modest, D. M. (1988). The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics*, 21(2), 213-254.
- Leland, H. E. (1998). Agency costs, risk management, and capital structure. *Journal of Finance*, 53(4), 1213-1243.
- Lessard, D. R. (1974). World, National, and Industry Factors in Equity Returns. *The Journal of Finance*, 29(2), 379-391.
- Lev, B., and Thiagarajan, S. R. (1993). Fundamental Information Analysis. *Journal of Accounting Research*, 31(2), 190-215.
- Lewellen, J., and Resutek, R. (2013). The predictive power of investment and accruals. *Tuck School of Business Working Paper*(2012-99).
- Lewellen, S. (2012). Firm-Specific Industries: Working Paper.
- Lewis, C. M., Rogalski, R. J., and Seward, J. K. (2001). The long-run performance of firms that issue convertible debt: an empirical analysis of operating characteristics and analyst forecasts. *Journal of Corporate Finance*, 7(4), 447-474.

- Li, X., and Sullivan, R. N. (2011). The Limits to Arbitrage Revisited: The Accrual and Asset Growth Anomalies. [Article]. *Financial Analysts Journal*, 67(4), 50-66.
- Lin, X., and Zhang, L. (2013). The investment manifesto. *Journal of Monetary Economics*.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 13-37.
- López de Silanes, F., La Porta, R., Shleifer, A., and Vishny, R. (1998). Law and finance. *Journal of Political Economy*, 106, 1113-1155.
- Lusardi, A., and Tufano, P. (2009). Debt literacy, financial experiences, and overindebtedness: National Bureau of Economic Research.
- Lyandres, E., Sun, L., and Zhang, L. (2008). The New Issues Puzzle: Testing the Investment-Based Explanation. *The Review of Financial Studies*, 21(6), 2825-2855.
- MacKie-Mason, J. K. (1990). Do taxes affect corporate financing decisions? *The Journal of Finance*, 45(5), 1471-1493.
- Maksimovic, V. (1988). Capital structure in repeated oligopolies. *The RAND Journal of Economics*, 389-407.
- Maksimovic, V., and Zechner, J. (1991). Debt, agency costs, and industry equilibrium. *The Journal of Finance*, 46(5), 1619-1643.
- Marchica, M. T., and Mura, R. (2010). Financial flexibility, investment ability, and firm value: evidence from firms with spare debt capacity. *Financial Management*, 39(4), 1339-1365.
- Mauer, D. C., and Ott, S. H. (2000). Agency costs, underinvestment, and optimal capital structure: The effect of growth options to expand. *Project Flexibility, Agency, and Competition (Oxford University Press, Oxford)*.
- Mauer, D. C., and Sarkar, S. (2005). Real options, agency conflicts, and optimal capital structure. *Journal of banking & Finance*, 29(6), 1405-1428.
- Mauer, D. C., and Triantis, A. J. (1994). Interactions of corporate financing and investment decisions: A dynamic framework. *The Journal of Finance*, 49(4), 1253-1277.
- McLean, R. D., and Pontiff, J. (2014). Does academic research destroy stock return predictability.
- McNichols, M. F., and Stubben, S. R. (2008). Does earnings management affect firms' investment decisions? *The Accounting Review*, 83(6), 1571-1603.

- Mehran, H., Taggart, R. A., and Yermack, D. (1999). CEO ownership, leasing, and debt financing. *Financial Management*, 28(2), 5-14.
- Mello, A. S., and Parsons, J. E. (1992). Measuring the agency cost of debt. *The Journal of Finance*, 47(5), 1887-1904.
- Mello, A. S., Parsons, J. E., and Triantis, A. J. (1995). An integrated model of multinational flexibility and financial hedging. *Journal of International Economics*, 39(1), 27-51.
- Miao, J. (2005). Optimal capital structure and industry dynamics. *The Journal of Finance*, 60(6), 2621-2659.
- Modigliani, F., and Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American Economic Review*, 261-297.
- Morellec, E. (2004). Can managerial discretion explain observed leverage ratios? *Review of Financial Studies*, 17(1), 257-294.
- Moyen, N. (2007). How big is the debt overhang problem? *Journal of Economic Dynamics and Control*, 31(2), 433-472.
- Myers, S. C. (1974). Interactions of corporate financing and investment decisions-implications for capital budgeting. *The Journal of Finance*, 29(1), 1-25.
- Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, 5(2), 147-175.
- Myers, S. C., and Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2), 187-221.
- Newey, W. K., and West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 777-787.
- Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), 1-28.
- Officer, M. S. (2011). Overinvestment, corporate governance, and dividend initiations. *Journal of Corporate Finance*, 17(3), 710-724.
- Ogneva, M. (2012). Accrual quality, realized returns, and expected returns: the importance of controlling for cash flow shocks. *The Accounting Review*, 87(4), 1415-1444.
- Ou, J. A., and Penman, S. H. (1989). Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics*, 11(4), 295-329.

- Pastor, L., and Stambaugh, R. F. (2003). Liquidity Risk and Expected Stock Returns. *Journal of Political Economy*, 111(3).
- Penman, S. H. (2001). On Comparing Cash Flow and Accrual Accounting Models for Use in Equity Valuation: A Response to Lundholm and O'Keefe (CAR, Summer 2001)*. *Contemporary Accounting Research*, 18(4), 681-692.
- Parrino, R., Poteshman, A. M., and Weisbach, M. S. (2005). Measuring Investment Distortions when Risk-Averse Managers Decide Whether to Undertake Risky Projects. *Financial Management*, 34(1), 21-60.
- Parrino, R., and Weisbach, M. S. (1999). Measuring investment distortions arising from stockholder–bondholder conflicts. *Journal of Financial Economics*, 53(1), 3-42.
- Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1), 435-480.
- Petkova, R. (2006). Do the Fama–French factors proxy for innovations in predictive variables? *Journal of Finance*, 61(2), 581-612.
- Pinches, G. E., and Mingo, K. A. (1973). A Multivariate Analysis of Industrial Bond Ratings. *The Journal of Finance*, 28(1), 1-18.
- Piotroski, J. D. (2000). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research*, 38, 1-41.
- Polk, C., and Sapienza, P. (2009). The Stock Market and Corporate Investment: A Test of Catering Theory. *The Review of Financial Studies*, 22(1), 187-217.
- Pontiff, J., and Woodgate, A. (2008). Share Issuance and Cross-Sectional Returns. *Journal of Finance*, 63(2), 921-945.
- PORTA, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Legal determinants of external finance. *The Journal of Finance*, 52(3), 1131-1150.
- Prechter Jr, R. R. (2001). Unconscious herding behavior as the psychological basis of financial market trends and patterns. *The Journal of Psychology and Financial Markets*, 2(3), 120-125.
- Prombutr, W., Phengpis, C., and Zhang, Y. (2012). What explains the investment growth anomaly? *Journal of banking & Finance*, 36(9), 2532-2542.
- Pukthuanthong, K., and Roll, R. (2014). A Protocol for Factor Identification.
- Rajan, R. G., and Zingales, L. (1995). What do we know about capital structure? Some evidence from international data. *The Journal of Finance*, 50(5), 1421-1460.

- Ramnath, S. (2002). Investor and Analyst Reactions to Earnings Announcements of Related Firms: An Empirical Analysis. *Journal of Accounting Research*, 40(5), 1351-1376.
- Rangan, S. (1998). Earnings management and the performance of seasoned equity offerings. *Journal of Financial Economics*, 50(1), 101-122.
- Rauh, J. D., and Sufi, A. (2010). Capital Structure and Debt Structure. *Review of Financial Studies*, 23(12), 4242-4280. doi: 10.1093/rfs/hhq095
- Rauh, J. D., and Sufi, A. (2012). Explaining corporate capital structure: Product markets, leases, and asset similarity. *Review of Finance*, 16(1), 115-155.
- Richardson, S. (2006). Over-investment of free cash flow. *Review of Accounting Studies*, 11(2-3), 159-189.
- Richardson, S., and Sloan, R. (2003). External financing and future stock returns. *Working Paper, University of Pennsylvania*(03-03).
- Richardson, S. A., Sloan, R. G., Soliman, M. T., and Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3), 437-485. doi: 10.1016/j.jacceco.2005.04.005
- Roberts, H. V. (1959). Stock-Market "Patterns" And Financial Analysis: Methodological Suggestions. *The Journal of Finance*, 14(1), 1-10.
- Rogers, W. (1983). Analyzing complex survey data, Rand Corporation memorandum. *Santa Monica, CA*.
- Roll, R. (1992). Industrial Structure and the Comparative Behavior of International Stock Market Indices. *The Journal of Finance*, 47(1), 3-41.
- Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3), 335-370.
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19(3), 425-442.
- Shleifer, A. (2000). *Inefficient markets: An introduction to behavioral finance*: Oxford university press.
- Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review*, 71(3), 289-315.
- Sorensen, D. E. (2000). Characteristics of merging firms. *Journal of Economics and Business*, 52(5), 423-433.

- Spiegel, M., and Tookes, H. (2013). Dynamic competition, valuation, and merger activity. *The Journal of Finance*, 68(1), 125-172.
- Spiess, D. K., and Affleck-Graves, J. (1999). The long-run performance of stock returns following debt offerings. *Journal of Financial Economics*, 54(1), 45-73.
- Stein, J. C. (1992). Convertible bonds as backdoor equity financing. *Journal of Financial Economics*, 32(1), 3-21.
- Stulz, R. (1990). Managerial discretion and optimal financing policies. *Journal of Financial Economics*, 26(1), 3-27.
- Subrahmanyam, A. (2010). The Cross-Section of Expected Stock Returns: What Have We Learnt from the Past Twenty-Five Years of Research? *European Financial Management*, 16(1), 27-42.
- Subramanyam, K. (1996). The pricing of discretionary accruals. *Journal of Accounting and Economics*, 22(1), 249-281.
- Sullivan, M., and Zhang, A. (2011). Are investment and financing anomalies two sides of the same coin? *Journal of Empirical Finance*, 18(4), 616-633.
- Teoh, S. H., Welch, I., and Wong, T. J. (1998). Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics*, 50(1), 63-99.
- Thaler, R. H. (1993). *Advances in Behavioral Finance* (Vol. 1): Russell Sage Foundation.
- Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1), 1-10.
- Titman, S., and Tsyplakov, S. (2007). A dynamic model of optimal capital structure. *Review of Finance*, 11(3), 401-451.
- Titman, S., Wei, K. C. J., and Xie, F. (2004). Capital Investments and Stock Returns. *Journal of Financial and Quantitative Analysis*, 39(04), 677-700.
- Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543): Wiley-Interscience.
- Tufano, P. (2001). HBS-JFE conference volume: complementary research methods. *Journal of Financial Economics*, 60(2), 179-185.
- Walkshäusl, C., and Lobe, S. (2011). The Alternative Three-Factor Model: An Alternative beyond US Markets? *European Financial Management*, 20(1), 33-70.
- Welch, I. (2004). Capital structure and stock returns. *Journal of Political Economy*, 112(1), 106-132.

- Williams, J. T. (1995). Financial and industrial structure with agency. *Review of Financial Studies*, 8(2), 431-474.
- Xing, Y. (2008). Interpreting the Value Effect through the Q-Theory: An Empirical Investigation. *The Review of Financial Studies*, 21(4), 1767-1795.
- Zang, A. Y. (2006). *Evidence on the tradeoff between real manipulation and accrual manipulation*: ProQuest.
- Zhang, L. (2005). The value premium. *Journal of Finance*, 60(1), 67-103.
- Zhang, X. (2006). Information uncertainty and stock returns. *The Journal of Finance*, 61(1), 105-137.

APPENDIX A
FOR ESSAY ONE

Appendix A.1

Table A.1 Definitions, Data source, Mnemonics and Labels for the Original Input

Dataset Items

Variable	Mnemonic	Definition	Annual data number	Data Sources
A1	CHE	This item represents cash and all securities readily transferable to cash as listed in the Current Asset section.	A1	COMPUSTAT
A2	CH	This item represents any immediately negotiable medium of exchange. It includes money and any instruments normally accepted by banks for deposit and immediate credit to a customer's account.	A162	COMPUSTAT
A3	IVST	This item represents currently marketable investments as presented in the current asset section of the Balance Sheet. Such investments may be converted to cash within a relatively short period of time.	A193	COMPUSTAT
A4	INVT	This item represents merchandise bought for resale and materials and supplies purchased for use in production of revenue.	A3	COMPUSTAT
A5	ACO	This item represents current assets that are not included in cash, cash equivalents, receivables or inventory on the Balance Sheet.	A68	COMPUSTAT
A6	ACT1	This item represents cash and other assets that are expected to be realized in cash or used in the production of revenue within the next 12 months.	A4	COMPUSTAT
A7	PPENT	This item represents the cost, less accumulated depreciation, of tangible fixed property used in the production of revenue.	A8	COMPUSTAT
A8	DPACT	This item represents the total portion of asset cost written off by periodic depreciation charges since the assets were acquired.	A196	COMPUSTAT
A9	IVAEQ	This item represents long-term investments and advances to unconsolidated subsidiaries and affiliates in which the parent company has significant control, as stated in the consolidated financial statements.	A31	COMPUSTAT
A10	PPEGT1	This item represents the cost of tangible fixed property used in the production of revenue.	A7	COMPUSTAT
A11	IVAO	This item represents long-term receivables and other investments and advances including investments in unconsolidated companies in which there is no control.	A32	COMPUSTAT
A12	INTAN	This item represents the unamortized value of intangible assets.	A33	COMPUSTAT
A13	AO	This item represents those long-term assets that are not property, plant and equipment, investments and advances, or intangibles.	A69	COMPUSTAT
A14	AT	This item represents current assets <i>plus</i> net property, plant, and equipment <i>plus</i> other noncurrent assets (including intangible assets, deferred items and investments and advances).	A6	COMPUSTAT
A15	DC	This item represents assets that derive value because they represent long-term prepayments of expenses that contribute to the production of revenue in the future.	A152	COMPUSTAT
A16	XPP	This item represents advance payments for services or benefits that are to be received within one operating cycle.	A160	COMPUSTAT
A17	RECT	This item represents claims against others (after applicable reserves) collectible in money, generally within one year.	A2	COMPUSTAT
L1	DLC	This item represents the total amount of short-term notes and the current portion of long-term debt, which is also stated as Notes Payable (NP) + Debt Due in 1st Year (DD1).	A34	COMPUSTAT
L2	NP	This item represents the total amount of short-term notes.	A206	COMPUSTAT
L3	AP	This item represents only trade obligations due within one year or the normal operating cycle of the company.	A70	COMPUSTAT
L4	TXP	This item represents the accrued tax liability on income due within one year.	A71	COMPUSTAT
L5	LCO	This item is a residual item representing those current liabilities that are not debt, trade accounts payable, or income taxes payable.	A72	COMPUSTAT
L6	LCT1	This item represents all liabilities due within one year, including short-term debt and the current portion of long-term debt. This item is not available for banks.	A5	COMPUSTAT

L7	DLTT	The item represents debt obligations due more than one year from the company's balance sheet date.	A9	COMPUSTAT
L8	DT	This concept is the sum of Total Long-Term Debt, which is defined as debt obligations due more than one year from the company's balance sheet date, <i>plus</i> Debt in Current Liabilities, which is defined as the total amount of short-term notes and the current portion of long-term debt (debt due in one year).	DLTT+D LC	COMPUSTAT
L9	LO	This item represents all noncurrent liabilities that are not debt, deferred taxes, investment tax credits, minority interest or shareholders' equity.	A75	COMPUSTAT
L10	TXDITC	This item represents the accumulated tax deferrals due to timing differences between the reporting of revenues and expenses for financial statements and tax forms and investment tax credit. This item excludes deferred taxes reported as current liabilities (included in Current Liabilities - Other). This item is not available for banks.	A35	COMPUSTAT
L11	TXDB	This item represents the accumulated tax deferrals due to timing differences between the reporting of revenues and expenses for financial reporting and tax purposes.	A74	COMPUSTAT
L12	ITCB	This item represents accumulated tax deferrals of investment tax credits generated by new capital investments. This data item is not available for banks.	A208	COMPUSTAT
L13	MIB	This item represents par or stated value of the subsidiary preferred or common stock not owned by the parent company, plus the minority interest's equity in the surplus of the subsidiary. Dividends in arrears on subsidiary preferred stock not owned by the parent company are included.	A38	COMPUSTAT
L14	LT1	This item represents current liabilities plus long-term debt plus other noncurrent liabilities (including deferred taxes, investment tax credit, and minority interest).	A181	COMPUSTAT
L15	PSTK	This item represents the net number of preferred shares at year-end multiplied by the par or stated value per share as presented in the company's Balance Sheet.	A130	COMPUSTAT
L16	PSTKR	This item represents any stock which the issuer undertakes to redeem at a fixed or determinable price on a fixed or determinable date or dates by operation of a sinking fund or other methods.	A175	COMPUSTAT
L17	PSTKN	This item represents the number of the company's nonredeemable preferred shares issued at yearend multiplied by the par or stated value per share.	A209	COMPUSTAT
L18	CEQ	This item represents the common shareholders' interest in the company. This item includes ... 1. Common stock outstanding (including treasury stock adjustments). 2. Capital surplus. 3. Retained earnings. 4. Treasury stock adjustments for both common and non-redeemable preferred stock	A60	COMPUSTAT
L19	CSTK	This item represents the total par or carrying value of common shares issued. Par usually equals the carrying value. However, when they are not identical, Standard & Poor's uses the carrying value to calculate this item.	A85	COMPUSTAT
L20	CAPS	This item represents a group of capital accounts other than capital stock or retained earnings.	A210	COMPUSTAT
L21	RE	This item represents the cumulative earnings of the company less total dividend distributions to shareholders. Stock adjustments made to this item relate to unissued shares.	A36	COMPUSTAT
L22	TSTK	This item represents the dollar amount of a company's common stock and nonredeemable preferred stock which has been reacquired and placed into treasury. A company uses the cost method to account for this stock and shows it on the Balance Sheet as a deduction to equity.	A88	COMPUSTAT
L23	XACC	This item represents expenses incurred with the passage of time.	A135	COMPUSTAT
BS1	CEQL	This item represents the common shareholders' interest in a company in the event of liquidation of company assets. Common equity is adjusted by the preferred stockholders' legal claims against the company.	A235	COMPUSTAT
BS2	CEQT	This item represents common stock plus the following items ... 1. Retained earnings. 2. Capital surplus. 3. Self-insurance reserves (when included in the Equity section). 4. Capital stock premium.	A11	COMPUSTAT
BS3	CSHO	This item represents the net number of all common shares outstanding at yearend, excluding treasury shares and scrip. Common treasury shares carried on the asset side of the balance sheet are netted against the number of common shares issued. Common shares paid in stock dividends are included when the ex-dividend date falls within the year and the payment date the next year.	A25	COMPUSTAT
BS4	DCPSTK	This item represents the balance sheet amount of outstanding convertible debt and the carrying value of convertible preferred stock that is reported as being convertible into common stock. This item is not available for banks.	39	COMPUSTAT
BS5	DCVT	This item represents all long-term debt convertible to the company's common or preferred stock and not subordinated to any other long-term debt.	A79	COMPUSTAT
BS6	DD	This item represents long-term debt containing a promise to pay a specific amount of money on a fixed date - usually more than 10 years after issuance and a promise to pay interest on stated dates.	A82	COMPUSTAT
BS7	DS	This item represents long-term debt with claims to interest and principal junior to	A80	COMPUSTAT

		that of ordinary debt. All debt specifically labeled subordinated but not convertible is included.		
BS8	DCVSUB	This item represents the portion of long-term debt that is both convertible and subordinated.	A154	COMPUSTAT
BS9	ICAPT	This item represents the sum of these data items. 1. Long-Term Debt – Total. 2. Preferred Stock - Carrying Value. 3. Minority Interest (Balance Sheet). 4. Common Equity – Total. This item may include the current portion of long-term debt when long-term debt is footnoted indicating long-term debt includes the current portion.	A37	COMPUSTAT
BS10	PSTKL	This item represents the total dollar value of the net number of preferred shares outstanding in the event of involuntary liquidation (such as, bankruptcy) multiplied by the per share involuntary liquidating value.	A10	COMPUSTAT
BS11	TLCF	This item represents the portion of prior and current year losses applied as a reduction of taxable income in the next succeeding year or years. When available and applicable, this item is usually reported in the notes to financial statements.	A52	COMPUSTAT
BS12	WCAP	This item represents the difference between total current assets minus total current liabilities as reported on a company's Balance Sheet. This item is not available for banks or utilities.	A179	COMPUSTAT
RET	RET	Daily change in the total value of an investment using prices or bid/ask averages if prices are not available. Dividends are reinvested on the ex-date.		CRSP
PRC	PRC	The last non-missing daily closing price or bid/ask average of security. If price is unavailable, the number in price is replace by bid/ask average.		CRSP
SHR	SHR	The unadjusted number of publically held share on NYSE, AMEX and NASDAQ exchanges and recorded in 000s.		CRSP
CAP	CAP	Closing price*shares outstanding (in 000s) at the end of previous period. If an index, capitalization is the total market value of issues used in the index at the beginning of previous period.		CRSP
TVOL	TVOL	Total volume traded within selected output calendar. For example, weekly calendar will sum the 5 trading days within each week.		CRSP

Appendix A.2: Definition of variables

- ASSETG*: the total assets growth rate, measured as the year-to-year change in total assets (item A6) $[(\text{item } A6_{t-1} - \text{item } A6_{t-2}) / \text{item } A6_{t-2}]$ as defined by Cooper, et al. (2008). Of note, the *ASSETG* for fiscal year 1985 with June as year end, is the difference between fiscal year 1984 ($t-2$) and 1985 ($t-1$). *ASSETG* and other accounting measures follow this notation. For the return predictability test, the accounting information at fiscal year $t-1$ is matched after a gap of t fiscal years (July of Calendar year 1985 to June of Calendar year 1986) with the future monthly stock returns for fiscal year $t+1$ (July of Calendar year 1986 to June of Calendar year 1987).
- ACCR*: accounting accruals equals to the $[(\Delta \text{item } A4 - \Delta \text{item } A1) - (\Delta \text{item } A5 - \Delta \text{item } A34 - \Delta \text{item } A71) - \text{item } A14]$ scaled by the $[(\text{item } A6_{t-1} + \text{item } A6_{t-2})/2]$ as used in Cooper et al. (2008).
- Accruals: accounting accruals equals to the $[(\Delta \text{item } A4 - \Delta \text{item } A1) - (\Delta \text{item } A5 - \Delta \text{item } A34 - \Delta \text{item } A71) - \text{item } A14]$ scaled by the lagged total assets (item $A6_{t-2}$).
- NOA*: net operating asset that is the difference between operating assets (OA) and operating liabilities (OL) in fiscal year $t-1$ scaled by the lagged total assets as used in Hirshleifer et al. (2004). $NOA = [(\text{item } A6 - \text{item } A1) - (\text{item } A6 - \text{item } A34 - \text{item } A9 - \text{item } A38 - \text{item } A130 - \text{item } A60)] / \text{item } A6_{t-2}$.
- I_A*: investment-to-assets is the sum of change in inventories (item3) and change in gross property, plant, and equipment (item8) in fiscal year $t-1$ scaled by lagged total assets (item6) as used in Chen et al. (2011).
- LEV: leverage is the sum of long-term debt (item9) and debt in current liabilities (item34), scaled by total assets (item A6) in fiscal year $t-1$ as defined by Cooper, et al. (2008).
- ROA*: the measure of profitability (*ROA*) is net income (item A178) divided by the total assets (item A6) in fiscal year $t-1$.
- LSIZE*: market size is the log of the market value of equity that is price times the common stocks outstanding at the June of year $t-1$.
- LBTM*: the log book to market ratio, before taking the logarithm the book to market ratio estimated as in Fama and French (1992). The market equity is estimated at the end of December falling in fiscal year $t-1$.
- BHRET6*: the last six months buy-and-hold return, where we skip the most recent month for microstructure biases.
- BHRET36*: the last three years buy-and-hold returns skipping the most recent month.
- Tobin's *Q*: the market value of the firm measured in fiscal year $t-1$ that is the market equity (at June $t-1$ a share outstanding times the share price) plus total assets (item 6) minus book equity (item 60) and divided by total assets (item 6) as defined in Hou and Robinson (2006).

TA (\$m):	the total assets (item 6) in million dollars for fiscal year $t-1$.
MV (\$m):	the market value of equity (shares outstanding times the price) in million dollars at the June $t-1$.
Earnings:	the income from continued operations (item 178) scaled by the lagged total assets (item 6) in fiscal year $t-1$.
CFlows:	the difference of the Earnings and the accruals in fiscal year $t-1$.
Cash/TAt-1:	the ratio of the cash and equivalents (item 1) in fiscal year $t-1$ to lagged total assets (item 6).
Div/K:	the sum of the common (item 21) and preferred dividends (item 19) scaled by the lagged plant, property, equipment (item 8) in fiscal year $t-1$.
Inv_mv:	the inverse of the market value of equity, where the market value of equity is obtained at the June of fiscal year $t-1$.
Inv_growth:	the ratio of the research and development expenditure (item 46) to the sale-net (item 12) in fiscal year $t-1$.
Access_fm:	the dummy for access to financial market equals to 1 if dividends-common (item 21) greater than zero and 0 otherwise in fiscal year $t-1$.
SIZE:	the accounting size, estimated as the logarithm of the total assets in fiscal year $t-1$.
NegY:	the negative earnings dummy, it is one if income before extraordinary items (item A18) for fiscal year $t-1$ is negative and zero otherwise.
YB:	the annual profitability measured as the income before extraordinary items (item A18) for fiscal year $t-1$ divided by the book value of equity (item A60) at the fiscal year end $t-2$.
NegACCR:	the negative accruals equal $ACCR$ if $ACCR$ for fiscal year $t-1$ is negative and zero otherwise.
PosACCR:	the positive accruals equal $ACCR$ if $ACCR$ for fiscal year $t-1$ is positive and zero otherwise.
NoDiv:	the no dividend dummy that is one if the common dividends (item A21) for the fiscal year $t-1$ is zero and zero otherwise.
Div_B:	the book dividend yield that is a common dividend (item A21) for fiscal year $t-1$ divided by the book value of equity (item A60) at fiscal year $t-2$.
Ret12:	the 12 month cumulative raw stock returns skipping the latest month over the last year.
Ret36:	the 24 month cumulative raw stock returns starting from last 13th month to 36th month.

Age:	the number of years a stock has appeared in the CRSP database at the end of December year t .
INV_AGE:	it is just inverse of the firm age.
SG:	the growth in net sales (item A12) measured as change in net sales from fiscal year $t-2$ to fiscal year $t-1$ divided by sales in fiscal year $t-1$.
RDI:	ratio of R&D expense (item 46) sales (item 12) is measure of R&D intensity.
FirmSize:	the natural logarithm of total assets converted to 2004 dollars using CPI index is used for SA index.
SG_N:	the sales growth measure adjusted for inflation $(\text{Item } 12_t / \text{CPI}_t - \text{Item } 12_{t-1} \times \text{CPI}_{t-1}) / (\text{Item } 12_{t-1} \times \text{CPI}_{t-1})$.
ZSCORE:	it is unlevered Z-score (bankruptcy measure) introduced by MacKie-Mason (1990). $3.3 \times [(\text{Item } 18) + (\text{Item } 15) + (\text{Item } 16)] + (\text{Item } 12) + 1.4 \times (\text{Item } 36) + 1.2 \times [(\text{Item } 4) - (\text{Item } 5)] / (\text{Item } 6)$.
PayDiv:	it equals 1 if cash dividends (Item 127) are greater than 0, and 0 otherwise.
Lev_Median:	it is the median leverage of the same 3-digit SIC industry.
ISG:	it is the real sales growth rate of the same 3-digit SIC industry.
SA index:	the size and age based financial constraint index measured as $SA = -0.737 \times \text{FIRMSIZE} + 0.043 \times \text{SIZE}^2 - 0.040 \times \text{AGE}$.
CFVOL:	represents cash flow volatility, which is the standard deviation of cash flow from operations scaled by the average total assets in the past 60 months (requiring at least 36 months' data).
ACCR_NEW:	is a measure of accounting accruals as in Richardson et al. (2005).
CASH:	Item 1/ lagged Item 6.
Equity:	Item 60 / lagged Item 6.
Debt:	NOA – (Equity + CASH).
INCOME:	Item 18/ average Item 6.
FCF:	is measure of free cash flow $\text{FCF} = \text{INCOME} - \text{ACCRUALS}$.
CCASH:	Change in Item 1/ average Item 6.
*	empirical tests where dependent variables are firm value, future earnings, future cash flows, and other known return anomalies the notation $t-1$ equal t and $t-2$ equal $t-1$.

Appendix A.3: Robustness Tests

Size Effect

The market capitalisation break points are used at the end of June each year to allocate stocks to three size samples. Stocks under the 20th percentile are allocated to a micro stock sample, stocks that are between the 20th and 50th percentiles constitute a small stock sample, and stocks above the median constitute a large stock sample.

Panel A of Table A.3.1 presents the Fama and MacBeth regression results for the large stock sample. In Model (1) all latent growth factors appear to be negatively associated with future stock returns, but only LT_INV has significant coefficients (*t*-statistics) of -0.015 (-2.35). In the other four models, it is found that only LT_INV and NOA are able to explain variations in future stock returns. NOA appears to subsume the effects of LT_INV in model (3) but after accounting for variations in NOA explained by latent factors, LT_INV persists in explaining the future stock returns indicated in the following section.

For small stocks the cross-sectional regression results in Panel B of Table A.3.1 suggest that all latent factors except PSK_USE have a negative association with future stock returns. The factors that have significant coefficients (*t*-statistics) in Model (1) are ST_CREDIT with -0.0026 (-1.96), LT_INV with -0.0015 (-2.23), and PSK_USE with 0.0022 (1.68). Among the other control variables ACCR, I_A, ROA, and BHRET36 are able to explain future returns.

Table A.3.1

Relations between latent factors and subsequent monthly raw stock returns

$$RET_{it+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \beta_2 Controls_{i,t-1} + \varepsilon_{i,t+1}.$$

Panel A: Large Sample										
	MODEL1		MODEL2		MODEL3		MODEL4		MODEL5	
Intercept	0.0061	(0.96)	0.0077	(1.19)	0.0097	(1.47)	0.0057	(0.88)	0.0053	(0.77)
<i>FIN_FLEX</i>	-0.0003	(-0.18)	0.0033	(0.81)	-0.0002	(-0.10)	-0.0004	(-0.23)	0.0003	(0.18)
<i>ST_CREDIT</i>	-0.0010	(-0.85)	0.0014	(0.76)	-0.0004	(-0.31)	-0.0009	(-0.76)	-0.0009	(-0.72)
<i>LT_INV</i>	-0.0015**	(-2.35)	0.0005	(0.25)	0.0003	(0.31)	-0.0016**	(-2.56)	-0.0006	(-0.55)
<i>CVT_DEBT</i>	-0.0002	(-0.50)	0.0004	(0.79)	0.0001	(0.15)	-0.0002	(-0.47)	-0.0001	(-0.18)
<i>PSK_USE</i>	-0.0012	(-0.81)	-0.0009	(-0.62)	-0.0013	(-0.89)	-0.0008	(-0.55)	-0.0009	(-0.62)
<i>ASSETG</i>			-0.0039	(-0.96)						
<i>NOA</i>					-0.0053***	(-2.90)				
<i>ACCR</i>							-0.0017	(-0.72)		
<i>I_A</i>									-0.0054	(-1.08)
<i>ROA</i>									0.0077	(1.23)
<i>LSIZE</i>	0.0001	(0.11)	0.0001	(0.10)	0.0000	(0.03)	0.0001	(0.12)	0.0000	(0.05)
<i>LBTM</i>	0.0024	(0.81)	0.0022	(0.76)	0.0036	(1.23)	0.0023	(0.81)	0.0032	(0.95)
<i>BHRET6</i>	-0.0024	(-0.49)	-0.0026	(-0.54)	-0.0028	(-0.58)	-0.0018	(-0.37)	-0.0028	(-0.58)
<i>BHRET36</i>	-0.0006	(-0.85)	-0.0006	(-0.83)	-0.0006	(-0.78)	-0.0006	(-0.80)	-0.0005	(-0.75)
Adj R-Sq	0.069***		0.071***		0.072***		0.075***		0.076***	
Panel B: Small Sample										
	MODEL1		MODEL2		MODEL3		MODEL4		MODEL5	
Intercept	0.0029	(0.34)	0.0041	(0.48)	0.0049	(0.56)	0.0022	(0.26)	0.0054	(0.66)
<i>FIN_FLEX</i>	-0.0008	(-0.59)	0.0010	(0.25)	-0.0001	(-0.03)	-0.0007	(-0.54)	0.0002	(0.12)
<i>ST_CREDIT</i>	-0.0026*	(-1.96)	-0.0015	(-0.78)	-0.0019	(-1.39)	-0.0018	(-1.44)	-0.0020	(-1.47)
<i>LT_INV</i>	-0.0015**	(-2.23)	-0.0007	(-0.41)	-0.0003	(-0.37)	-0.0015**	(-2.16)	-0.0001	(-0.15)
<i>CVT_DEBT</i>	-0.0004	(-1.01)	-0.0003	(-0.51)	-0.0004	(-0.79)	-0.0004	(-0.95)	-0.0005	(-1.02)
<i>PSK_USE</i>	0.0022*	(1.68)	0.0024*	(1.81)	0.0022*	(1.67)	0.0024*	(1.84)	0.0021	(1.55)
<i>ASSETG</i>			-0.0009	(-0.31)						
<i>NOA</i>					-0.0026	(-1.55)				
<i>ACCR</i>							-0.0045**	(-2.02)		
<i>I_A</i>									-0.0077*	(-1.72)
<i>ROA</i>									0.0113*	(1.78)
<i>LSIZE</i>	0.0006	(0.53)	0.0005	(0.48)	0.0006	(0.54)	0.0005	(0.51)	0.0001	(0.11)
<i>LBTM</i>	0.0016	(0.53)	0.0016	(0.53)	0.0019	(0.69)	0.0013	(0.46)	0.0015	(0.52)
<i>BHRET6</i>	-0.0001	(-0.02)	-0.0002	(-0.05)	-0.0000	(-0.01)	0.0003	(0.09)	-0.0001	(-0.02)
<i>BHRET36</i>	-0.0011*	(-1.92)	-0.0011*	(-1.82)	-0.0012*	(-1.97)	-0.0011*	(-1.87)	-0.0011*	(-1.92)
Adj R-Sq	0.035***		0.036***		0.037***		0.039***		0.046***	
Panel C: Micro Sample										
	MODEL1		MODEL2		MODEL3		MODEL4		MODEL5	
Intercept	0.0371***	(5.91)	0.0371***	(6.04)	0.0415***	(5.88)	0.0362***	(5.61)	0.0353***	(6.24)
<i>FIN_FLEX</i>	-0.0045***	(-3.92)	-0.0038	(-1.13)	-0.0027*	(-1.73)	-0.0045***	(-3.96)	-0.0034***	(-2.86)
<i>ST_CREDIT</i>	-0.0020*	(-1.88)	-0.0017	(-1.22)	-0.0005	(-0.45)	-0.0018	(-1.65)	-0.0018*	(-1.66)
<i>LT_INV</i>	-0.0026***	(-3.98)	-0.0028*	(-1.85)	0.0004	(0.40)	-0.0025***	(-3.82)	-0.0018**	(-2.10)
<i>CVT_DEBT</i>	-0.0011**	(-2.43)	-0.0010*	(-1.93)	-0.0010*	(-1.96)	-0.0012**	(-2.48)	-0.0011**	(-2.42)
<i>PSK_USE</i>	0.0004	(0.44)	0.0003	(0.33)	0.0005	(0.52)	0.0004	(0.44)	-0.0000	(-0.02)
<i>ASSETG</i>			-0.0006	(-0.23)						
<i>NOA</i>					-0.0071***	(-3.45)				
<i>ACCR</i>							-0.0032**	(-2.24)		
<i>I_A</i>									-0.0029	(-0.96)
<i>ROA</i>									-0.0084**	(-2.02)
<i>LSIZE</i>	-0.0067***	(-7.06)	-0.0066***	(-7.04)	-0.0067***	(-7.04)	-0.0066***	(-6.93)	-0.0064***	(-7.09)
<i>LBTM</i>	0.0005	(0.24)	0.0004	(0.19)	0.0019	(0.89)	0.0004	(0.17)	0.0020	(1.11)
<i>BHRET6</i>	-0.0040	(-1.50)	-0.0040	(-1.50)	-0.0043	(-1.62)	-0.0040	(-1.50)	-0.0044*	(-1.69)
<i>BHRET36</i>	-0.0031***	(-3.65)	-0.0032***	(-3.68)	-0.0030***	(-3.53)	-0.0032***	(-3.70)	-0.0026***	(-3.33)
Adj R-Sq	0.022***		0.022***		0.023***		0.023***		0.026***	

In this table across three size samples, using Fama and MacBeth (1973) cross-sectional regressions, the individual monthly raw future stock returns (RET_{it+1}) are regressed on the Latent growth factors and the other characteristics for the period of July, 1986 to December, 2010. The size samples are constructed following Fama and French (2008). The header of Table 1 defines the *LatentGrowth* (include *FIN_FLEX*, *ST_CREDIT*, *LT_INV*, *CVT_DEBT*, *PSK_USE*) and the *Controls* (include *LBTM*, *LSIZE*, *BHRET6*, *BHRET36*, *ASSETG*, *NOA*, *ACCR*, *I_A*, and *ROA*) are defined in appendix 1. The table shows average slope coefficients with significance (***) for 1%, ** for 5% and * for 10%) based on *t*-statistics.

Panel C of Table A.3.1 indicates regression results for the micro stock sample. Results for Model (1) indicate that FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT are significant negative predictors of future stock returns. The relative coefficients (*t*-statistics) for FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT are -0.045 (-3.92), -0.020 (-1.88), -0.026 (-3.98), and -0.0011 (-2.43), respectively. Consistently, our latent factors are able to capture return effects of the total asset growth measure (ASSETG). Among other control variables, I_A, LBTM, and BHRET6 become flat. NOA subsumes the effect of ST_CREDIT and LT_INV but the rest of the latent factors persist across the models. It appears that the return predictive power of latent growth factors is robustly prevalent in micro stocks, with no substantial influence of the control variables.

Overall, size sample based regression results suggest that all latent factors have an expected negative association with future stock returns except PSK_USE, which is positive. It appears that return effects of total asset growth (ASSETG), book-to-market (LBTM), and short-term price momentum (BHRET6) are captured by the latent factors. The association for large and medium sized firms is not as strong as in the full sample, showing micro stocks influence, which is true for the set of control variables too. The evidence is consistent with the Bernard and Thomas (1990), and Fama and French (2008), who observe that investment effects are stronger in small firms than in big firms.

Table A.3.2**Relations between return anomalies and latent growth factors**

$$X_{i,t} = \alpha + b \text{LatentGrowth}_{i,t} + \varepsilon_{it}$$

Dependent Variable	Intercept	<i>FIN_FLEX</i>	<i>ST_CREDIT</i>	<i>LT_INV</i>	<i>CVT_DEBT</i>	<i>PSK_USE</i>	Adj R-Sq
<i>ASSETG</i>	0.3569*** (28.18)	1.1465*** (24.29)	0.4462*** (32.76)	0.5142*** (45.62)	0.0639*** (16.91)	0.0478*** (4.00)	0.825***
<i>NOA</i>	0.7320*** (115.32)	0.3448*** (16.95)	0.2105*** (21.46)	0.4208*** (48.39)	0.0341*** (11.97)	0.0351*** (3.35)	0.480***
<i>ACCR</i>	-0.2447*** (-107.75)	0.0365*** (9.31)	0.1016*** (22.22)	-0.0111*** (-4.48)	0.0120*** (14.07)	-0.0124*** (-3.35)	0.041***
<i>I_A</i>	0.0824*** (37.06)	0.1437*** (25.57)	0.1058*** (23.68)	0.2509*** (65.78)	-0.0021* (-1.70)	0.0188*** (6.17)	0.648***
<i>ROA</i>	0.0322*** (25.39)	0.0521*** (15.73)	-0.0243*** (-9.79)	0.0285*** (30.77)	-0.0092*** (-21.07)	-0.0115*** (-7.21)	0.044***
<i>LBTM</i>	0.5063*** (95.29)	-0.1023*** (-53.70)	-0.0394*** (-18.83)	-0.0325*** (-26.50)	-0.0090*** (-18.52)	-0.0033** (-2.16)	0.034***
<i>LSIZE</i>	5.1540*** (123.88)	0.2872*** (18.01)	-0.0910*** (-7.19)	0.2911*** (44.44)	0.0010 (0.17)	-0.0391*** (-4.75)	0.018***
<i>BHRET6</i>	0.0682*** (6.69)	-0.0303*** (-9.81)	-0.0033 (-1.19)	-0.0196*** (-9.63)	-0.0030*** (-4.02)	0.0088*** (5.22)	0.007***
<i>BHRET36</i>	0.5728*** (20.97)	0.8266*** (19.04)	0.1154*** (11.23)	0.1895*** (21.49)	-0.0115*** (-2.93)	-0.0489*** (-3.83)	0.060***

This table reports the estimated coefficients and *t*-statistics of latent growth factors when regressed on prior return anomalies ($X_{i,t}$) running the Fama and MacBeth (1973) cross-sectional regressions for the sample period of 1985 to 2009. Where $X_{i,t}$ include *ASSETG*, *NOA*, *ACCR*, *I_A*, *ROA*, *LBTM*, *LSIZE*, *BHRET6*, and *BHRET36* as dependent variables and *LatentGrowth* include *FIN_FLEX*, *ST_CREDIT*, *LT_INV*, *CVT_DEBT*, *PSK_USE* latent growth factors. Both sets of variables are defined in header of Table 2. The table shows average slope coefficients with significance (***) for 1%, ** for 5% and * for 10%) based on *t*-statistics. See appendix 1 for variable definition.

Table A.3.3 Cross-sectional regressions of subsequent raw monthly stock returns on the latent growth factors and residuals other characteristics

$$X_{i,t} = \alpha + b \text{LatentGrowth}_{i,t} + \varepsilon_{it}$$

$$RET_{it+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-1} + \beta_2 \text{ResidControls}_{i,t-1} + \mu_{i,t+1}$$

Panel A: Full sample							Panel B: Large sample						
	M-1	M-2	M-3	M-4	M-5	M-6		M-1	M-2	M-3	M-4	M-5	M-6
Intercept	0.0096*** (2.84)	0.0097*** (2.87)	0.0096*** (2.84)	0.0096*** (2.85)	0.0095*** (2.82)	0.0098*** (2.92)	Intercept	0.0071* (1.79)	0.0072* (1.78)	0.0071* (1.80)	0.0072* (1.80)	0.0069* (1.74)	0.0066 (1.59)
FIN_FLEX	-0.0059*** (-5.18)	-0.0056*** (-4.04)	-0.0065*** (-5.50)	-0.0060*** (-5.20)	-0.0061*** (-5.32)	-0.0054*** (-4.98)	FIN_FLEX	-0.0010 (-0.51)	-0.0018 (-0.60)	-0.0027 (-1.36)	-0.0011 (-0.57)	-0.0009 (-0.46)	-0.0008 (-0.39)
ST_CREDIT	-0.0022*** (-2.60)	-0.0018** (-2.17)	-0.0023*** (-2.73)	-0.0022*** (-2.59)	-0.0022*** (-2.64)	-0.0021** (-2.52)	ST_CREDIT	-0.0012 (-0.97)	-0.0005 (-0.35)	-0.0015 (-1.19)	-0.0012 (-0.97)	-0.0015 (-1.13)	-0.0014 (-1.13)
LT_INV	-0.0035*** (-5.93)	-0.0035*** (-5.55)	-0.0035*** (-6.22)	-0.0034*** (-5.79)	-0.0037*** (-6.03)	-0.0035*** (-5.92)	LT_INV	-0.0016** (-2.43)	-0.0016* (-1.97)	-0.0020*** (-2.78)	-0.0017*** (-2.60)	-0.0019** (-2.25)	-0.0015** (-2.15)
CVT_DEBT	-0.0006** (-2.36)	-0.0005* (-1.87)	-0.0006** (-2.24)	-0.0007** (-2.43)	-0.0007** (-2.46)	-0.0006** (-2.22)	CVT_DEBT	-0.0002 (-0.52)	0.0001 (0.29)	-0.0001 (-0.27)	-0.0002 (-0.55)	-0.0002 (-0.52)	-0.0001 (-0.40)
PSK_USE	0.0002 (0.38)	0.0004 (0.57)	0.0002 (0.36)	0.0003 (0.45)	0.0003 (0.40)	0.0000 (0.04)	PSK_USE	-0.0012 (-0.81)	-0.0011 (-0.74)	-0.0018 (-1.18)	-0.0008 (-0.54)	-0.0008 (-0.54)	-0.0015 (-1.00)
Resid_ASSETG		0.0004 (0.29)					Resid_ASSETG		-0.0039 (-0.96)				
Resid_NOA			-0.0054*** (-3.59)				Resid_NOA			-0.0053*** (-2.79)			
Resid_ACCR				-0.0041*** (-2.87)			Resid_ACCR				-0.0017 (-0.72)		
Resid_IA					-0.0033 (-1.47)		Resid_IA					-0.0046 (-0.94)	
Resid_ROA						-0.0070* (-1.66)	Resid_ROA						0.0074 (1.20)
Resid_SIZE	-0.0022*** (-3.80)	-0.0022*** (-3.80)	-0.0022*** (-3.86)	-0.0022*** (-3.78)	-0.0022*** (-3.80)	-0.0019*** (-3.98)	Resid_SIZE	0.0001 (0.11)	0.0001 (0.10)	0.0001 (0.13)	0.0001 (0.12)	0.0000 (0.09)	0.0000 (0.08)
Resid_BM	0.0022 (0.98)	0.0021 (0.96)	0.0031 (1.56)	0.0019 (0.88)	0.0021 (0.97)	0.0030 (1.59)	Resid_BM	0.0024 (0.81)	0.0022 (0.76)	0.0035 (1.17)	0.0023 (0.81)	0.0020 (0.69)	0.0035 (1.04)
Resid_BHRET6	-0.0039 (-1.33)	-0.0039 (-1.33)	-0.0040 (-1.38)	-0.0039 (-1.32)	-0.0039 (-1.34)	-0.0041 (-1.43)	Resid_BHRET6	-0.0024 (-0.49)	-0.0026 (-0.54)	-0.0027 (-0.56)	-0.0018 (-0.37)	-0.0027 (-0.56)	-0.0025 (-0.52)
Resid_BHRET36	-0.0020*** (-3.11)	-0.0020*** (-3.08)	-0.0021*** (-3.21)	-0.0020*** (-3.14)	-0.0020*** (-3.07)	-0.0017*** (-2.86)	Resid_BHRET36	-0.0006 (-0.85)	-0.0006 (-0.83)	-0.0006 (-0.82)	-0.0006 (-0.80)	-0.0005 (-0.75)	-0.0006 (-0.86)
Adj R-Sq	0.026***	0.027***	0.027***	0.028***	0.027***	0.031***	Adj R-Sq	0.069***	0.071***	0.072***	0.075***	0.071***	0.074***

Table A.3.3: Continued

Panel C: Small sample							Panel B: Micro sample						
	M-1	M-2	M-3	M-4	M-5	M-6		M-1	M-2	M-3	M-4	M-5	M-6
Intercept	0.0066*	0.0072*	0.0071*	0.0069*	0.0064	0.0067*	Intercept	0.0009	0.0007	0.0006	0.0010	0.0007	0.0012
	(1.69)	(1.84)	(1.83)	(1.77)	(1.63)	(1.71)		(0.23)	(0.18)	(0.16)	(0.27)	(0.20)	(0.33)
FIN_FLEX	-0.0016	-0.0009	-0.0021	-0.0017	-0.0019	-0.0011	FIN_FLEX	-0.0088***	-0.0088***	-0.0093***	-0.0089***	-0.0090***	-0.0079***
	(-1.04)	(-0.37)	(-1.37)	(-1.08)	(-1.21)	(-0.74)		(-5.12)	(-3.70)	(-4.73)	(-5.21)	(-5.17)	(-4.73)
ST_CREDIT	-0.0015	-0.0010	-0.0012	-0.0012	-0.0023	-0.0016	ST_CREDIT	-0.0020*	-0.0020*	-0.0022**	-0.0021**	-0.0021**	-0.0019*
	(-1.03)	(-0.65)	(-0.84)	(-0.84)	(-1.48)	(-1.16)		(-1.92)	(-1.78)	(-2.07)	(-2.01)	(-1.98)	(-1.78)
LT_INV	-0.0017**	-0.0013	-0.0014*	-0.0016**	-0.0021**	-0.0016**	LT_INV	-0.0052***	-0.0057***	-0.0053***	-0.0050***	-0.0054***	-0.0050***
	(-2.19)	(-1.54)	(-1.76)	(-2.02)	(-2.50)	(-2.01)		(-6.89)	(-5.77)	(-7.35)	(-6.64)	(-6.93)	(-6.84)
CVT_DEBT	-0.0005	-0.0004	-0.0005	-0.0006	-0.0007	-0.0005	CVT_DEBT	-0.0009*	-0.0009*	-0.0010**	-0.0010**	-0.0009*	-0.0008*
	(-1.22)	(-0.88)	(-1.10)	(-1.31)	(-1.59)	(-1.13)		(-1.93)	(-1.65)	(-2.01)	(-2.03)	(-1.91)	(-1.73)
PSK_USE	0.0017	0.0017	0.0018	0.0018	0.0019	0.0013	PSK_USE	0.0008	0.0007	0.0008	0.0008	0.0007	0.0005
	(1.28)	(1.28)	(1.38)	(1.42)	(1.40)	(0.94)		(0.86)	(0.74)	(0.80)	(0.90)	(0.79)	(0.61)
Resid_ASSETG		-0.0020					Resid_ASSETG		-0.0005				
		(-0.64)							(-0.19)				
Resid_NOA			-0.0025				Resid_NOA			-0.0075***			
			(-1.47)							(-3.57)			
Resid_ACCR				-0.0043*			Resid_ACCR				-0.0033**		
				(-1.93)							(-2.29)		
Resid_IA					-0.0083*		Resid_IA					-0.0023	
					(-1.83)							(-0.78)	
Resid_ROA						0.0130**	Resid_ROA						-0.0089**
						(2.02)							(-2.13)
Resid_SIZE	0.0005	0.0004	0.0006	0.0004	0.0004	-0.0001	Resid_SIZE	-0.0067***	-0.0067***	-0.0068***	-0.0066***	-0.0067***	-0.0064***
	(0.42)	(0.39)	(0.54)	(0.35)	(0.37)	(-0.10)		(-7.10)	(-7.07)	(-7.12)	(-6.97)	(-7.08)	(-7.14)
Resid_BM	0.0016	0.0017	0.0021	0.0013	0.0017	0.0014	Resid_BM	0.0006	0.0005	0.0019	0.0004	0.0006	0.0021
	(0.51)	(0.55)	(0.71)	(0.44)	(0.55)	(0.45)		(0.26)	(0.21)	(0.93)	(0.19)	(0.26)	(1.16)
Resid_BHRET6	0.0010	0.0009	0.0013	0.0013	0.0008	0.0011	Resid_BHRET6	-0.0041	-0.0041	-0.0044	-0.0040	-0.0041	-0.0044*
	(0.24)	(0.21)	(0.33)	(0.34)	(0.19)	(0.28)		(-1.51)	(-1.52)	(-1.64)	(-1.51)	(-1.53)	(-1.68)
Resid_BHRET36	-0.0011*	-0.0010	-0.0012*	-0.0010	-0.0010	-0.0011*	Resid_BHRET36	-0.0033***	-0.0033***	-0.0032***	-0.0033***	-0.0033***	-0.0027***
	(-1.67)	(-1.61)	(-1.81)	(-1.59)	(-1.54)	(-1.86)		(-3.89)	(-3.91)	(-3.84)	(-3.94)	(-3.85)	(-3.63)
Adj R-Sq	0.036***	0.037***	0.038***	0.040***	0.038***	0.046***	Adj R-Sq	0.021***	0.022***	0.023***	0.023***	0.022***	0.025***

New vs. Known Factors and Return Predictability

Table A.3.2 reports the Fama and MacBeth (1973) cross-sectional regression results of latent accounting growth factors on the well-established determinants of average stock returns. The aim of this test is twofold: first, to examine whether latent factors are able to explain variations in return anomalies or not. In fact, it is found that latent factors do contain information relating to these anomalies. Second, if latent factors are able to explain variations in return anomalies, then after separating the common information contained in the known determinants of the stock returns, do they still possess enough information about the subsequent stock returns? For the first objective, we use cross-sectional regressions of the following form:

$$X_{i,t} = \alpha + b_i \text{LatentGrowth}_{i,t} + \varepsilon_{it} \quad (\text{A1})$$

Where $X_{i,t}$ includes two sets of dependent variables: first, market variables that include LSIZE as the log of market value of equity, LBTM as the log of book equity to market equity, BHRET6 as the 6-month buy-and-hold returns over January (t) to June (t), and BHRET36 as the 3-year buy-and-hold returns over July ($t-3$) to June (t). The second set of accounting variables include: ASSETG as a measure of the total asset growth, ACCR as a measure of accounting accruals, the NOA representing cumulative accruals, I_A as a measure of investment-to-asset ratio and ROA as a measure of profitability. The independent variables are the five latent accounting growth measures.

Table A.3.2 indicates that the five latent factors are robust determinants of these known return anomalies. The latent factors, being accounting measures by construction, have high adjusted R -squared values in explaining accounting anomalies other than those of market anomalies. The highest adjusted R -squared values are 83% when explaining ASSETG, 64% for I_A, and 48% for NOA. For the rest of the variables the adjusted R -squared values range from 6% for BHRET36 to just 0.7% for BHRET6.

For the second objective, we obtain the residual of all models given in Table A.3.2 and run Fama and MacBeth's (1973) cross-sectional regressions of monthly raw stock returns on the latent growth factors and the residual factors over the whole sample period and for the three capitalisation based samples. The regressions are of the following form:

$$RET_{it+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \beta_2 ResidControls_{i,t-1} + \varepsilon_{i,t+1} \quad (A2)$$

Where $ResidControls_{i,t-1}$ are the residuals of factors (LBTM, LSIZE, BHRET6, BHRET36, ASSETG, NOA, ACCR, I_A, and ROA) obtained by running Fama-MacBeth (1973) regressions of the latent growth factors (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE as in equation (A2)). Panel A of Table A3 indicates that for the whole sample period, both coefficients (t -statistics) of the latent factors have improved and they are robust across all models as compared to the results given in Table 5. This study is able to claim that the factors can capture all stock return-related effects of total asset growth (ASSETG), investment-to-assets (I_A), logarithm of book-to-market ratio (LBTM), and

short-term price momentum (*BHRET6*). Among control residuals, those of NOA, ACCR, ROA, LSIZE, and BHRET36 maintain their return predictability. The results of the whole sample period are consistent with the results of the micro firms shown in Panel D.

Panel B of Table A.3.3 indicates that, consistent with results for the large sample given in Table A.3.1, only one latent factor – long-term investment (LT_INV) – appears to predict future stock returns after controlling for other known return determinants. Among controls only the residual of NOA possesses a significant negative relationship with future stock returns, but it does not affect LT_INV. However, in Panel C of Table A.3.3, LT_INV persists and the residual of NOA becomes insignificant for small size stocks.

Appendix A.4: Full Sample Factor Analysis Output Factors based Return Predictability Tests

This study started with application of the factor analysis model on the full sample instead of the annual sample to test the conjecture that there exist systematic shifts in the firms' corporate finance decisions. Similar to chapter 2 in this thesis, the overall sample factor structure obtained via factor analysis was similar to one presented in Table 2.1. In the overall, one can obtain same five factors with same factor components.

The purpose of this appendix is to give an overview of the implication of the full sample based latent growth factors of the return predictability. Table A.4.1 present the results of Fama-MacBeth (1973) cross-sectional regressions of annual raw buy-and-hold returns on the latent growth factors and prominent controls for the period of nineteen years (July, 1993 to June 2011). This study employs following regression models:

$$RET_{i,t+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \varepsilon_{it} \quad (A.4.1)$$

$$RET_{i,t+1} = \alpha + \beta_1 LatentGrowth_{i,t-1} + \beta_2 Controls_{i,t} + \varepsilon_{it} \quad (A.4.2)$$

For full sample, in Panel A, Table A.4.1, evidence shows that all latent growth factors but PSK_USE, have significant negative associations with future returns. The coefficients (t-statistics) of FIN_FLEX, ST_CREDIT, LT_INV and CVT_DEBT growth measures in a model (1) without controls are -0.041 (-5.17), -0.023 (-3.54), -0.021 (-2.22), and -0.013 (-3.40) respectively. Among control variables only LSIZE has a significant negative association with future returns, but this factor has no substantial effect on the

predictive power of the latent factors. The book to market ratio is insignificant may be due to increase in firm size during our research period.⁶³ The empirical results support our argument that the newly created latent variables are better represented than the simple Asset Growth used in the previous literature.

For cross-sectional regressions, the results for large stocks are reported in Panel B, Table A.4.1, the results are consistent with the growth effect seen in decile portfolio returns. Among latent growth factors except LT_INV and PSK_USE, the FIN_FLEX, ST_CREDIT, and CVT_DEBT have a significant coefficients (t-statistic) -0.023 (-2.34), -0.014 (2.30), and -0.06 (-1.98) respectively. However, among other control variables no factor has significant future stock returns predictive power. In the regression model which includes all latent growth measures, when augmented by ASSETG, there exist no factor possessing return predictive power for large stocks. The results suggest that the prediction power of ASSETG shown in previous literature for large stock is originated from growth in financial flexibility, short-term credit, and debt convertible.

For small stocks cross-sectional regression results in Panel C, Table A.4.1 suggests that the FIN_FLEX, ST_CREDIT, and PSK_USE significantly predict future returns, their coefficient (t-statistics) estimates are -0.031 (-4.13), -0.027 (-2.84), and -0.013 (-1.80) respectively. Among the other control variables the 36 months lagged buy-and-hold returns (BHRET36) variable marginally explains the future returns with a coefficient (t-statistics)

⁶³ See, Chordia, et al. (2012) “The book-to-market ration has declined over time, and this decline is consistent with the increase in firm size.

of -0.025 (-1.91). Overall, results for small stocks suggest that return prediction power of financial flexibility, short-term credit growth, and preferred stock usage is an independent phenomenon.

Then micro firm cross-sectional results in Panel D, Table A.4.1 shows that FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT are significant negative predictors of the future stock returns. The relative coefficients (t-statistics) for FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT are -0.059 (-4.01), -0.021 (-1.85), -0.040 (-2.72), and -0.028 (-2.66) respectively. Among other control variables, the ASSETG becomes flat. However, in such a regression, ACCR, LSIZE and BHRET36 show significant association with future stock returns. Additionally, these control variables have effect on the ST_CREDIT but the rest of the latent factors remains significant. It appears that the return predictive power of latent growth factors is prevalent in micro stocks robustly with no influence of the control variables.

For sub-sample tests, this study divides the data period in two sub-samples. The first sub-sample constitutes of fiscal year 1992 to the year 2000, and second sub-sample constitutes of fiscal year 2001 to 2009. The construction of regression models remains same. For first sub-sample, the cross-sectional results are given in Panel E in Table A.4.1, the latent factor regression models show significant coefficients (t-statistics) for FIN_FLEX, ST_CREDIT, LT_INV, and CVT_DEBT is -0.041 (-3.44), -0.024 (-2.57), -0.036 (-2.68), and -0.014 (-2.05) respectively. Moreover, only BHRET36 is a robust

determinant of the future stock returns in among control variables. However, it does not subsume the predictive power of latent factors.

Compared to first sub-sample, in Panel F, Table A.4.1 reports that the significance of the latent growth effect is reduced significantly in the base regression model, only FIN_FLEX, ST_CREDIT, and CVT_DEBT is significant in explaining future stock returns in the second sub-sample period (2001-2009). The coefficients (t-statistics) for these latent factors are -0.040 (-3.24), -0.018 (-1.87), and -0.010 (-1.88) respectively. Among control variables, ACCR is only robust in explaining the future stock returns. Again, we find the latent growth factors consistently superior in return prediction as compared to other well-known factors.

Table A.4.1 Fama-MacBeth Regressions of Annual Stock Returns on Latent Growth Factors and Other Variables

$$RET_{it+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-1} + \varepsilon_{it}$$

$$RET_{it+1} = \alpha + \beta_1 \text{LatentGrowth}_{i,t-1} + \beta_2 \text{Controls}_{i,t} + \varepsilon_{it}$$

Panel A: ALL FIRMS															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.157***	-0.041***	-0.023***	-0.021**	-0.013***	0.000									0.014***
M-2	0.162***	-0.037***	-0.018***	-0.018	-0.011**	0.002	-0.010								0.014***
M-3	0.195**	-0.040***	-0.022***	-0.019**	-0.019**	0.002		-0.061							0.023***
M-4	0.146***	-0.040***	-0.020**	-0.021**	-0.012**	0.002			-0.056*						0.016***
M-5	0.160***	-0.033***	-0.021***	-0.020**	-0.012***	0.002				-0.014					0.023***
M-6	0.276***	-0.034***	-0.021***	-0.015*	-0.010**	-0.003					-0.023**	0.005			0.028***
M-7	0.154***	-0.030***	-0.019***	-0.017*	-0.012**	0.003							0.011	-0.021	0.022***

Panel B: LARGE FIRMS															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.124***	-0.023**	-0.014**	-0.007	-0.006*	0.003									0.027**
M-2	0.127***	-0.014	-0.008	-0.002	-0.004	0.003	-0.025								0.027**
M-3	0.142*	-0.024***	-0.015**	-0.006	-0.006	0.003		-0.033							0.040***
M-4	0.113***	-0.022**	-0.014*	-0.009	-0.006*	0.002			-0.042						0.030**
M-5	0.110**	-0.021*	-0.013*	-0.008	-0.006**	0.007				0.107					0.036***
M-6	0.140**	-0.021**	-0.014**	-0.007	-0.006*	0.002					-0.003	0.011			0.036***
M-7	0.117***	-0.015	-0.016**	-0.005	-0.005	0.003							0.035	-0.005	0.046***

Panel C: SMALL FIRMS															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.166***	-0.031***	-0.027**	-0.017	-0.011	-0.013*									0.010***
M-2	0.169***	-0.024**	-0.022**	-0.012	-0.009	-0.014*	-0.020								0.010***
M-3	0.189*	-0.029***	-0.026**	-0.016	-0.010	-0.014**		-0.042							0.021***
M-4	0.153**	-0.030***	-0.025**	-0.017	-0.011	-0.014*			-0.058						0.012***
M-5	0.161***	-0.027**	-0.024**	-0.021*	-0.009	-0.006				0.101					0.021***
M-6	0.298**	-0.028***	-0.027**	-0.016	-0.010	-0.014**					-0.031	0.010			0.017***
M-7	0.148***	-0.032***	-0.026***	-0.009	-0.012	-0.016**							0.044	-0.025*	0.020***

Panel D: MICRO FIRMS															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.230***	-0.059***	-0.021*	-0.040**	-0.028**	0.008									0.007***
M-2	0.232***	-0.054**	-0.018	-0.039**	-0.027**	0.008	-0.017								0.006***
M-3	0.318**	-0.053***	-0.022*	-0.043***	-0.028**	0.008		-0.142							0.015***
M-4	0.209***	-0.058***	-0.019	-0.042**	-0.028**	0.008			-0.083*						0.009***
M-5	0.230***	-0.052***	-0.021*	-0.037**	-0.029**	0.007				-0.055					0.015***
M-6	0.604***	-0.051***	-0.016	-0.030**	-0.028**	0.005					-0.138***	-0.012			0.019***
M-7	0.204***	-0.036*	-0.012	-0.040**	-0.029**	0.013							-0.027	-0.099***	0.018***

Table A.4.1 Continued

Panel E: SUB-SAMPLE (1992-2000)															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.165***	-0.041***	-0.024**	-0.036**	-0.014*	0.005									0.017**
M-2	0.167***	-0.037**	-0.021**	-0.033**	-0.013	0.005		-0.011							0.017**
M-3	0.256	-0.038***	-0.026**	-0.032***	-0.011	0.003		-0.161							0.030***
M-4	0.160**	-0.040***	-0.024*	-0.036**	-0.014*	0.005			-0.023						0.019**
M-5	0.168**	-0.030	-0.024**	-0.034***	-0.014**	0.002				-0.055					0.030***
M-6	0.273**	-0.037**	-0.025**	-0.031**	-0.013*	0.001					-0.021	-0.007			0.029***
M-7	0.160***	-0.029*	-0.022*	-0.030**	-0.013*	0.006							0.062	-0.013**	0.023***
Panel F: SUB-SAMPLE (2001-2009)															
Model	Intercept	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	ASSETG	NOA	ACCR	ROA	LSIZE	LBTM	BHRET6	BHRET36	AdjR ²
M-1	0.155*	-0.040**	-0.018*	-0.005	-0.010*	-0.000									0.011***
M-2	0.157*	-0.036*	-0.016	-0.003	-0.009	0.000		-0.009							0.011***
M-3	0.135	-0.041**	-0.018	-0.006	-0.010*	0.001		0.040							0.016***
M-4	0.133	-0.039**	-0.016	-0.007	-0.009	-0.001			-0.090**						0.014***
M-5	0.153*	-0.037***	-0.019*	-0.006	-0.010*	0.002				0.028					0.016***
M-6	0.279*	-0.030***	-0.017	0.002	-0.007	-0.007					-0.025	0.017			0.028***
M-7	0.148	-0.030**	-0.015*	-0.004	-0.011	-0.000							-0.040	-0.028	0.020***

In this table, the annual geometric compounding future returns are regressed on the observed Latent growth measures and the other growth measures. For every year for period of 1992 to 2009, this study runs the Fama-MacBeth (1973) cross-sectional regressions at the individual stock level. Two regression models are employed. First model is labeled as base regression model; and in second model the base regression model is augmented by control variables. FIN_FLEX growth measure is the latent growth score obtained via factor analysis procedures and it proxy for financial flexibility. Similar is the construction of other latent growth factors. ST_CREDIT growth measure proxy for short-term credit of firms. LT_INV growths proxy for long-term capital investment. CVT_DEBT growth measure proxy for convertible debt usage asset. PSK_USE growth measure for preferred stock usage. For The five factors, original input variables are estimated by ratio of change in balance sheet variables to lagged total assets before processing through the factor analysis model. LSIZE is the log market value of equity at June of year t-1. LBTM is the log book equity to market equity (Fama and French, 1992).). BHRET6 is the 6-month buy-and-hold returns over January (t) to June (t). BHRET36 is the 3-year buy-and-hold returns over July (t-3) to June (t). In case of controls we have ASSETG as a measure of the total asset growth that the year-to-year change in total assets [(TASSETSt-TASSETSt-1)/TASSETSt-1] (Cooper, et al. 2008). ACCR (accruals) equals the change in accounts receivables plus the change in inventories plus the change in the other current assets minus the change in account payable minus the change in other current liabilities minus depreciation and scaled by lagged total assets as defined by Polk and Sapienza (2009). NOA represents cumulative accruals as defined by Hirshleifer et al. (2004). The measure of profitability (ROA) is net income (Compustat item A178) divided by the total assets (Compustat item A6). ***, **, and * show the significance of parameter coefficients at 1%, 5%, and 10% respectively.

APPENDIX C

Appendix C.1: Fama and French Industry Specifications based on SIC Codes

Under section 4.3 in Tables 4.4 and 4.5, this study presents the firm frequency across the 10 industries for the 12 clusters. The specifications of the 10 industries are borrowed from the 12 industry portfolios given in the Fama-French online data library⁶⁴. We use 10 industries out of 12 because utilities (with SIC codes 4900-4949) and financials (with SIC code 6000-6999) are not part of this research project.

Table C.1.1

No.	Industry	Label	Compositions	SIC Codes
1.	Consumer NonDurables	NoDur	Food, Tobacco, Textiles, Apparel, Leather, Toys	0100-0999, 2000-2399, 2700-2749, 2770-2799, 3100-3199, 3940-3989.
2.	Consumer Durables	Durbl	Cars, TVs, Furniture, Household Appliances	2500-2519, 2590-2599, 3630-3659, 3710-3711, 3714-3714, 3716-3716, 3750-3751, 3792-3792, 3900-3939, 3990-3999.
3.	Manufacturing	Manuf	Machinery, Trucks, Planes, Off Furn, Paper, Com Printing	2520-2589, 2600-2699, 2750-2769, 3000-3099, 3200-3569, 3580-3629, 3700-3709, 3712-3713, 3715-3715, 3717-3749, 3752-3791, 3793-3799, 3830-3839, 3860-3899.
4.	Oil, Gas, and Coal	Enrgy	Oil, Gas, and Coal Extraction and Products	1200-1399, 2900-2999.
5.	Chemicals	Chems	Chemicals and Allied Products	2800-2829, 2840-2899.
6.	Business Equipment	BusEq	Computers, Software, and Electronic Equipment	3570-3579, 3660-3692, 3694-3699, 3810-3829, 7370-7379.
7.	Telecommunication	Telcm	Telephone and Television Transmission	4800-4899.
8.	Shops	Shops	Wholesale, Retail, and Some Services (Laundries, Repair Shops)	5000-5999, 7200-7299, 7600-7699.
9.	Health	Hlth	Healthcare, Medical Equipment, and Drugs	2830-2839, 3693-3693, 3840-3859, 8000-8099.
10.	Other	Other	Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment	Excluding 4900-4949, 6000-6999, and all above SIC codes.

⁶⁴ See, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/changes_ind.html.

Table C.1.2
Industry Characteristics' pattern across clusters

	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE	SA	WW	I_A	ACCESS_FM	NODIV	CASH	EQUITY	DEBT	SG_N	ZSCORE	ACCR_NEW	CFVOL	INV_MV	INV_AGE	
CLUSTER 1																										
NoDur	5.92	2.08	-0.79	-0.61	-5.79	0.11	1.13	2.17	4.04	-0.23	1.50	-1.96	.	0.50	0.00	1.00	8.95	14.05	-17.08	.	-1.57	0.23	.	0.02	0.75	
Durbl	4.99	-0.89	1.84	-0.72	-6.23	0.01	0.93	0.26	3.18	-0.10	1.50	-1.56	.	7.75	0.00	1.00	10.36	18.54	-20.51	.	2.29	0.78	.	0.05	0.75	
Manuf	4.65	0.12	-0.38	-0.69	-5.87	0.02	1.22	0.17	4.69	-0.12	1.57	-2.14	.	0.57	0.00	1.00	3.45	4.98	-6.80	.	0.98	0.37	.	0.03	0.73	
Engry	2.98	-2.63	0.26	-0.58	-6.68	0.04	1.51	0.00	2.43	-0.01	7.50	-0.54	.	0.22	0.00	1.00	0.64	1.68	-1.20	.	-11.95	0.27	.	0.09	0.13	
Chems	4.73	-0.19	-1.19	0.06	-6.42	0.01	0.92	2.41	3.93	-0.20	1.46	-1.81	.	0.10	0.00	1.00	4.13	4.70	-8.19	.	0.14	0.13	.	0.05	0.77	
Buseq	5.03	0.22	-0.36	-0.30	-5.30	0.02	1.36	0.67	5.33	-0.06	1.70	-2.33	-0.29	0.43	0.00	1.00	5.71	6.90	-11.25	3.32	2.23	0.21	0.21	0.02	0.72	
Telecm	5.06	-0.26	0.53	-0.76	-5.65	0.01	1.21	1.13	4.88	-0.06	1.64	-2.03	.	0.72	0.00	1.00	8.31	9.62	-16.53	.	3.13	0.29	.	0.01	0.70	
Shops	5.10	-0.32	0.35	-0.60	-3.84	0.04	0.72	0.91	4.55	-0.11	1.60	-2.16	.	0.89	0.00	1.00	10.88	13.26	-21.49	.	2.47	0.38	.	0.03	0.72	
Hlth	5.25	-0.09	-0.66	-0.53	-5.62	0.05	1.12	2.82	4.69	-0.21	1.69	-2.13	.	0.25	0.00	1.00	5.04	5.37	-9.77	.	-1.76	0.13	0.12	0.02	0.73	
Others	5.47	-0.24	-0.06	-0.21	-5.89	0.05	1.37	1.75	5.22	-0.21	1.69	-2.27	.	0.35	0.00	1.00	7.54	7.81	-14.73	.	-0.40	0.16	0.16	0.01	0.74	
CLUSTER 2																										
NoDur	0.42	-0.05	0.72	-0.79	4.95	0.24	0.75	0.07	4.10	-0.07	5.38	-2.27	-0.24	0.37	0.22	0.77	0.31	0.53	0.36	5.82	1.66	0.28	0.17	0.02	0.24	
Durbl	0.16	0.09	0.57	-0.46	4.57	0.26	0.22	0.01	4.15	0.01	5.00	-2.56	-0.29	0.26	0.38	0.63	0.20	0.37	0.44	1.12	3.65	0.31	0.13	0.03	0.41	
Manuf	0.19	-0.16	0.00	-0.16	3.97	0.25	0.45	0.21	4.45	0.02	7.47	-2.51	0.04	0.13	0.35	0.65	0.15	0.46	0.26	2.61	3.26	0.15	0.08	0.03	0.24	
Engry	0.22	-0.40	0.43	-0.35	4.13	0.26	0.27	0.00	3.71	-0.07	7.71	-2.44	0.80	0.26	0.07	0.93	0.14	0.50	0.39	2.41	-0.05	0.19	0.11	0.07	0.22	
Chems	0.20	-0.35	0.25	0.16	3.74	0.25	0.64	0.07	4.22	-0.24	8.43	-2.41	-0.16	0.14	0.28	0.72	0.22	0.43	0.16	0.85	-3.14	0.16	0.13	0.04	0.23	
Buseq	0.53	-0.02	-0.12	0.01	4.62	0.11	0.93	0.60	3.93	-0.21	5.83	-2.11	-0.07	0.09	0.04	0.97	0.53	0.69	-0.36	3.43	-4.37	0.23	0.17	0.05	0.26	
Telecm	0.35	-0.21	0.83	0.05	3.92	0.31	0.64	0.32	4.83	-0.10	4.74	-2.45	-0.21	0.21	0.00	1.00	0.26	0.61	0.52	1.49	0.03	0.37	0.14	0.06	0.38	
Shops	0.51	0.33	0.52	-0.06	5.99	0.22	0.45	0.00	3.47	-0.03	7.28	-2.23	0.08	0.44	0.09	0.91	0.23	0.56	0.49	1.57	1.63	0.34	0.17	0.08	0.28	
Hlth	0.49	-0.26	-0.10	0.08	4.67	0.14	1.06	1.88	3.92	-0.36	6.33	-2.04	0.16	0.10	0.04	0.96	0.81	0.65	-0.81	5.46	-7.13	0.15	0.22	0.05	0.23	
Others	0.52	-0.02	0.81	0.03	4.33	0.27	0.59	0.21	3.69	-0.10	5.14	-2.19	-0.17	0.38	0.12	0.88	0.43	0.57	0.39	0.99	-0.74	0.29	0.10	0.06	0.29	
CLUSTER 3																										
NoDur	-0.12	-0.09	-0.07	-0.06	-0.01	0.25	0.24	0.02	3.98	0.07	9.79	-2.52	-0.11	0.05	0.31	0.69	0.12	0.58	0.06	1.50	6.77	0.06	0.08	0.02	0.17	
Durbl	-0.13	-0.07	-0.04	-0.06	-0.01	0.22	0.36	0.05	3.99	0.05	8.70	-2.45	-0.11	0.07	0.33	0.67	0.11	0.59	0.04	1.40	6.49	0.07	0.07	0.02	0.20	
Manuf	-0.13	-0.10	-0.07	-0.06	-0.01	0.22	0.26	0.04	3.95	0.06	10.13	-2.49	-0.12	0.05	0.41	0.59	0.12	0.57	0.02	1.34	6.13	0.05	0.07	0.02	0.17	
Engry	-0.15	-0.19	0.10	-0.12	-0.02	0.23	0.38	0.00	3.98	-0.02	10.04	-2.43	-0.11	0.06	0.13	0.87	0.13	0.61	0.01	12.86	1.00	0.05	0.09	0.02	0.17	
Chems	-0.17	-0.07	-0.04	-0.09	0.01	0.21	0.54	0.26	3.97	-0.01	10.18	-2.34	-0.09	0.05	0.35	0.65	0.18	0.60	-0.10	1.53	3.64	0.05	0.11	0.02	0.16	
Buseq	-0.16	-0.09	-0.12	-0.09	-0.01	0.11	0.50	0.22	3.95	-0.05	9.08	-2.31	0.00	0.02	0.11	0.89	0.29	0.68	-0.44	1.82	1.74	0.03	0.12	0.02	0.19	
Telecm	-0.18	-0.07	0.05	-0.13	0.02	0.31	0.44	0.02	4.15	-0.03	8.26	-2.52	-0.07	0.04	0.18	0.82	0.15	0.46	0.07	2.36	0.45	0.06	0.11	0.02	0.23	
Shops	-0.13	-0.06	-0.04	-0.08	-0.01	0.25	0.19	0.00	3.98	0.05	8.96	-2.58	-0.12	0.07	0.24	0.76	0.10	0.52	0.09	1.39	8.37	0.07	0.07	0.02	0.20	
Hlth	-0.21	-0.09	-0.06	-0.09	-0.02	0.14	0.83	1.40	3.98	-0.19	8.64	-2.16	0.06	0.04	0.07	0.93	0.37	0.70	-0.57	3.03	-3.77	0.05	0.15	0.02	0.20	
Others	-0.15	-0.12	-0.01	-0.08	-0.01	0.26	0.31	0.08	3.94	0.02	8.70	-2.49	-0.14	0.05	0.21	0.79	0.16	0.54	-0.02	9.40	4.46	0.06	0.09	0.02	0.20	
CLUSTER 4																										
NoDur	0.74	3.94	0.70	-0.46	0.39	0.36	0.51	0.00	3.64	0.07	5.94	-2.21	-0.55	1.33	0.10	0.90	0.20	1.18	0.92	13.78	5.55	0.61	0.11	0.04	0.33	
Durbl	1.09	4.47	0.97	-0.26	-0.14	0.37	0.66	0.00	3.61	0.02	5.23	-2.16	-0.17	1.16	0.08	1.00	0.24	1.67	0.96	4.03	2.56	0.84	0.11	0.04	0.34	
Manuf	0.76	4.16	1.92	-0.16	0.37	0.39	0.55	0.21	3.32	-0.04	5.55	-2.12	-0.74	2.05	0.07	0.93	0.41	1.51	1.71	23.56	2.20	0.63	0.11	0.06	0.32	
Engry	0.99	4.19	2.81	-0.47	0.02	0.43	0.55	0.00	3.35	-0.02	7.35	-2.05	1.19	2.89	0.12	0.88	0.27	1.57	2.10	22.32	-1.03	0.82	0.15	0.06	0.25	
Chems	0.76	3.41	1.83	-0.25	0.23	0.34	0.65	0.02	3.54	0.15	6.57	-2.23	-0.20	13.17	0.22	0.75	1.72	6.12	15.46	4.62	4.49	0.51	0.07	0.03	0.24	
Buseq	1.07	4.02	1.26	-0.44	0.14	0.28	0.81	0.16	3.57	-0.06	4.69	-1.98	-0.12	1.14	0.03	0.97	0.72	2.18	-0.06	33.26	2.53	0.58	0.08	0.07	0.37	
Telecm	1.03	3.88	1.78	-0.29	0.13	0.37	0.74	0.00	3.51	-0.08	3.63	-1.96	-0.15	0.95	0.05	0.95	1.23	1.86	-0.93	2.46	1.34	0.53	0.07	0.05	0.42	
Shops	0.68	4.25	1.17	-0.01	-0.07	0.37	0.63	0.00	3.59	0.03	4.70	-2.15	-0.21	3.37	0.08	0.93	0.84	2.61	0.96	3.05	6.09	0.67	0.08	0.06	0.38	
Hlth	1.11	4.02	1.73	-0.40	0.38	0.32	0.85	0.36	3.78	-0.06	5.77	-2.01	-0.12	1.56	0.06	0.94	0.79	2.16	0.68	2.88	-0.45	0.66	0.14	0.05	0.33	
Others	1.30	3.98	1.82	-0.28	0.31	0.31	0.76	0.07	3.69	-0.01	4.39	-2.11	-0.44	1.47	0.15	0.85	0.57	2.13	1.02	15.27	4.12	0.68	0.10	0.06	0.43	

FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE	SA	WW	I_A	ACCESS_FM	NODIV	CASH	EQUITY	DEBT	SG_N	ZSCORE	ACCR_NEW	CFVOL	INV_MV	INV_AGE	
CLUSTER 5																									
NoDur	-0.13	-0.10	-0.04	-0.03	-0.01	0.25	0.75	0.01	8.46	0.16	14.92	-3.21	-0.28	0.03	0.86	0.14	0.09	0.48	0.11	2.13	6.80	0.05	0.06	0.00	0.11
Durbl	-0.14	-0.09	-0.03	-0.05	-0.02	0.25	0.57	0.02	8.25	0.12	15.45	-3.21	-0.32	0.04	0.84	0.16	0.10	0.45	0.06	1.63	6.38	0.05	0.05	0.00	0.11
Manuf	-0.13	-0.11	-0.03	-0.07	-0.03	0.24	0.55	0.02	8.21	0.11	15.61	-3.21	-0.25	0.04	0.79	0.21	0.08	0.47	0.11	2.56	6.04	0.05	0.05	0.00	0.10
Enrgy	-0.19	-0.18	0.27	-0.12	-0.01	0.25	0.57	0.01	8.69	0.09	14.42	-3.21	-0.09	0.09	0.74	0.26	0.06	0.47	0.18	3.32	4.36	0.07	0.07	0.00	0.12
Chems	-0.16	-0.11	-0.03	-0.03	-0.02	0.27	0.63	0.02	8.38	0.12	14.97	-3.22	-0.29	0.03	0.89	0.11	0.07	0.41	0.18	1.94	5.85	0.04	0.05	0.00	0.11
Buseq	0.02	-0.07	-0.04	-0.11	0.00	0.12	1.07	0.11	8.45	0.11	12.51	-3.15	-0.17	0.05	0.30	0.70	0.39	0.82	-0.63	3.08	4.79	0.08	0.09	0.00	0.14
Telcm	-0.14	-0.14	0.12	0.00	-0.01	0.35	0.47	0.02	8.85	0.07	10.58	-3.20	-0.23	0.04	0.51	0.49	0.11	0.43	0.20	4.00	1.89	0.07	0.08	0.00	0.19
Shops	-0.10	-0.11	0.05	-0.10	0.00	0.19	0.76	0.00	8.34	0.14	14.72	-3.19	-0.24	0.08	0.65	0.35	0.11	0.56	0.02	2.56	10.09	0.07	0.05	0.00	0.11
Hlth	-0.05	-0.09	0.03	-0.08	-0.02	0.20	1.14	0.24	8.55	0.13	14.19	-3.16	-0.10	0.05	0.48	0.52	0.27	0.70	-0.28	2.89	4.98	0.09	0.10	0.00	0.11
Others	-0.11	-0.12	0.09	-0.05	-0.01	0.24	0.69	0.00	8.16	0.11	14.13	-3.18	-0.07	0.06	0.65	0.35	0.15	0.52	-0.01	5.58	5.14	0.07	0.07	0.00	0.12
CLUSTER 6																									
NoDur	2.05	-0.28	-0.31	-0.19	0.12	0.05	0.97	0.07	4.45	0.14	4.39	-2.16	-0.19	0.27	0.14	0.86	1.00	1.92	-1.89	0.70	6.99	0.27	0.07	0.02	0.51
Durbl	2.20	-0.24	-0.36	-0.46	0.13	0.04	1.33	0.38	4.23	-0.08	4.50	-1.86	-0.23	0.34	0.06	0.94	1.03	1.88	-1.96	7.44	2.86	0.27	0.16	0.03	0.41
Manuf	1.98	-0.17	-0.18	-0.29	0.13	0.05	1.00	0.13	4.57	0.09	4.29	-2.21	-0.16	0.31	0.15	0.85	0.95	1.84	-1.78	2.25	4.96	0.53	0.12	0.02	0.47
Enrgy	1.93	-0.40	-0.52	-0.32	0.02	0.07	1.03	0.17	4.02	-0.05	5.83	-1.88	-0.10	0.11	0.12	0.88	0.91	1.51	-1.69	1.53	-1.68	0.08	0.07	0.04	0.35
Chems	1.97	-0.22	-0.82	-0.01	0.22	0.06	1.43	0.90	4.33	-0.19	5.30	-1.84	-0.14	0.15	0.00	1.00	1.41	1.84	-2.72	1.99	-0.24	0.15	0.12	0.03	0.38
Buseq	2.22	-0.05	-0.49	-0.14	0.05	0.04	1.10	0.30	4.88	0.02	4.18	-2.29	-0.11	0.18	0.06	0.94	1.78	2.61	-3.46	2.20	2.70	0.22	0.15	0.01	0.49
Telcm	2.15	-0.14	-0.38	-0.19	0.13	0.09	0.91	0.22	4.97	-0.05	4.81	-2.34	0.01	0.20	0.04	0.96	1.32	2.00	-2.39	2.60	1.25	0.20	0.19	0.02	0.47
Shops	2.01	-0.11	-0.13	-0.38	0.22	0.05	1.06	0.02	4.54	0.07	2.86	-2.17	-0.09	0.40	0.20	0.80	0.91	2.00	-1.68	1.89	6.03	0.34	0.15	0.02	0.56
Hlth	2.13	-0.17	-0.65	-0.02	0.03	0.05	1.26	2.27	4.78	-0.18	6.19	-2.20	0.00	0.10	0.02	0.98	1.89	2.27	-3.65	5.24	-2.54	0.11	0.20	0.01	0.33
Others	2.13	0.00	-0.35	-0.30	0.13	0.05	1.13	0.36	4.67	0.02	4.38	-2.21	0.98	0.19	0.10	0.89	1.51	2.37	-2.87	2.18	3.22	0.25	0.15	0.02	0.49
CLUSTER 7																									
NoDur	0.20	0.27	0.61	5.37	-0.18	0.43	0.49	0.06	4.37	0.03	5.51	-2.45	-0.13	0.45	0.36	0.64	0.48	1.58	0.09	1.89	4.18	0.28	0.07	0.03	0.26
Durbl	0.18	-0.05	0.44	5.50	-0.06	0.47	0.44	0.34	4.37	0.00	9.27	-2.56	-0.20	0.30	0.21	0.79	0.71	0.69	-0.30	0.74	4.86	0.27	0.07	0.06	0.20
Manuf	0.09	0.11	0.41	5.14	-0.21	0.39	0.73	0.09	5.40	-0.01	8.29	-2.69	-0.18	0.26	0.21	0.80	0.28	0.63	0.18	4.50	3.75	0.29	0.09	0.02	0.21
Enrgy	0.12	-0.46	1.44	5.11	-0.32	0.44	0.60	0.00	5.41	-0.01	10.76	-2.69	0.40	0.69	0.16	0.83	0.36	0.80	0.27	5.31	0.31	0.35	0.11	0.02	0.16
Chems	0.69	-0.06	0.69	5.60	-0.27	0.46	0.62	0.12	4.95	-0.01	9.63	-2.70	0.31	0.37	0.07	0.93	0.71	0.76	-0.31	11.44	2.57	0.35	0.11	0.02	0.17
Buseq	0.42	0.04	0.23	5.69	0.02	0.37	0.80	0.22	5.89	-0.01	8.27	-2.75	-0.16	0.19	0.06	0.94	0.77	1.15	-0.67	2.47	0.73	0.27	0.14	0.01	0.21
Telcm	0.24	0.00	0.98	5.58	0.01	0.51	0.60	0.06	5.59	-0.05	5.72	-2.72	0.23	0.25	0.14	0.86	0.65	0.81	-0.03	5.71	0.41	0.33	0.09	0.01	0.31
Shops	0.33	0.13	0.46	5.54	-0.05	0.38	0.52	0.00	5.04	0.05	6.64	-2.64	-0.22	0.35	0.25	0.75	0.52	0.77	-0.01	0.84	6.61	0.32	0.08	0.03	0.23
Hlth	0.48	0.01	0.40	5.84	0.00	0.45	1.03	1.44	5.73	-0.09	8.22	-2.66	-0.04	0.15	0.05	0.95	0.99	1.02	-0.93	4.33	-1.25	0.26	0.18	0.01	0.20
Others	0.23	-0.08	0.71	5.31	-0.10	0.43	0.57	0.02	4.93	0.01	6.24	-2.62	-0.13	0.36	0.17	0.83	0.48	0.81	-0.02	3.58	2.93	0.32	0.10	0.02	0.27
CLUSTER 8																									
NoDur	-0.19	0.42	3.58	-0.04	-0.11	0.41	0.40	0.00	5.06	0.05	7.62	-2.75	-0.18	0.73	0.35	0.65	0.22	1.07	0.98	7.68	4.11	0.68	0.07	0.02	0.29
Durbl	-0.30	0.41	3.32	0.01	-0.16	0.47	0.37	0.01	5.20	0.10	7.44	-2.85	-0.29	0.68	0.30	0.70	0.08	0.76	1.10	0.90	5.17	0.64	0.05	0.01	0.22
Manuf	-0.17	0.40	3.74	0.04	-0.07	0.42	0.37	0.01	4.90	0.07	8.17	-2.74	-0.18	1.06	0.40	0.60	0.17	1.03	0.95	8.45	4.28	0.68	0.08	0.02	0.25
Enrgy	-0.25	-0.42	4.14	-0.47	-0.06	0.33	0.52	0.00	5.30	0.03	8.65	-2.72	0.18	1.28	0.17	0.83	0.14	1.22	0.59	7.28	1.02	0.62	0.10	0.02	0.25
Chems	-0.10	-0.06	4.21	-0.52	-0.25	0.39	0.35	0.01	4.97	0.03	7.84	-2.70	-0.14	1.22	0.34	0.65	0.32	1.19	0.59	4.04	2.28	0.68	0.16	0.03	0.24
Buseq	0.24	0.35	3.73	-0.31	-0.24	0.35	0.56	0.21	4.71	-0.05	6.20	-2.54	-0.15	0.94	0.06	0.93	0.67	1.94	-0.08	3.32	0.74	0.67	0.10	0.02	0.32
Telcm	-0.11	-0.12	4.15	0.37	-0.09	0.53	0.40	0.07	5.61	0.00	4.55	-2.89	-0.15	0.76	0.12	0.88	0.66	1.02	0.90	4.04	0.95	0.70	0.06	0.01	0.40
Shops	0.03	0.00	3.59	-0.41	-0.12	0.34	0.68	0.00	4.73	0.04	5.78	-2.50	-0.24	0.90	0.22	0.78	0.23	1.21	0.49	2.99	4.24	0.64	0.08	0.02	0.33
Hlth	0.00	0.26	3.73	-0.23	-0.16	0.35	0.71	0.36	5.00	0.02	5.34	-2.52	-0.16	0.77	0.07	0.93	0.35	1.41	0.32	4.85	2.14	0.72	0.08	0.02	0.33
Others	-0.02	-0.25	3.92	-0.33	-0.10	0.39	0.45	0.02	5.22	0.04	5.84	-2.74	-0.09	1.23	0.20	0.80	0.32	1.32	0.62	4.92	2.32	0.67	0.08	0.02	0.34

FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE	SA	WW	I_A	ACCESS_FM	NODIV	CASH	EQUITY	DEBT	SG_N	ZSCORE	ACCR_NEW	CFVOL	INV_MV	INV_AGE	
CLUSTER 9																									
NoDur	6.13	-0.74	0.02	-0.35	0.65	0.04	1.14	0.20	3.62	0.01	1.77	-1.65	-0.15	0.58	0.05	0.95	4.93	6.59	-9.64	1.29	4.08	0.44	0.00	0.05	0.67
Durbl	6.24	-0.41	0.31	-0.32	0.65	0.05	1.42	0.94	3.93	-0.26	1.45	-1.62		0.98	0.10	0.89	5.26	6.87	-10.36		-0.69	0.34		0.03	0.78
Manuf	6.42	-0.30	-0.33	-0.34	0.73	0.05	1.33	0.63	3.74	-0.11	2.56	-1.63	-0.19	1.74	0.00	1.00	10.66	13.02	-21.01	1.23	2.23	0.30	0.34	0.08	0.68
Engry	6.97	-0.45	1.03	-0.64	0.19	0.01	0.83	0.13	5.11	-0.12	5.08	-2.45	-0.24	2.35	0.00	1.00	7.06	9.66	-14.07	1.47	0.13	0.34	0.16	0.01	0.47
Chems	6.47	-0.34	0.03	-0.13	1.13	0.02	1.64	0.70	3.73	-0.43	3.49	-1.51	-0.14	0.49	0.00	1.00	6.40	7.09	-12.28	-0.22	-3.23	0.30	0.10	0.04	0.62
Buseq	6.30	-0.04	-0.41	-0.13	0.38	0.02	1.26	0.58	4.79	-0.07	2.30	-2.12	-0.04	1.16	0.04	0.96	8.14	13.44	-15.48	10.71	2.32	0.37	0.12	0.03	0.66
Telcm	5.90	-0.49	1.68	-0.22	0.35	0.17	0.83	0.14	4.76	-0.13	2.58	-2.47	-0.14	0.93	0.00	1.00	10.73	8.93	-15.47	2.26	0.98	0.47	0.25	0.02	0.60
Shops	6.35	-0.13	0.33	-0.38	0.71	0.03	1.01	0.02	3.85	-0.07	1.57	-1.86	-0.19	1.37	0.08	0.92	8.62	11.25	-17.04	1.73	3.96	0.37	0.00	0.06	0.74
Hlth	6.34	-0.23	-0.68	0.00	0.44	0.04	1.16	2.61	4.30	-0.18	4.52	-1.93	-0.12	0.34	0.00	1.00	6.83	7.48	-13.35	1.65	-1.43	0.17	0.23	0.03	0.51
Others	6.37	0.01	0.22	-0.63	0.64	0.04	1.13	0.65	4.19	-0.03	2.94	-2.06	0.25	0.79	0.10	0.89	7.58	9.18	-14.79	13.46	2.62	0.35	0.29	0.05	0.60
CLUSTER 10																									
NoDur	-0.16	-0.11	-0.15	-0.06	0.00	0.24	0.04	0.01	2.14	0.01	8.97	-1.87	-0.06	0.03	0.21	0.79	0.12	0.57	0.02	0.93	6.66	0.02	0.07	0.15	0.19
Durbl	-0.18	-0.10	-0.15	-0.04	-0.03	0.28	0.10	0.05	2.15	-0.05	7.96	-1.89	-0.02	0.03	0.12	0.88	0.12	0.50	0.07	1.94	4.57	0.02	0.10	0.14	0.21
Manuf	-0.17	-0.09	-0.16	-0.03	-0.01	0.26	0.04	0.04	2.07	0.00	9.16	-1.88	-0.05	0.02	0.17	0.83	0.11	0.50	0.06	0.85	5.33	0.01	0.07	0.18	0.17
Engry	-0.19	-0.13	-0.19	-0.03	-0.02	0.20	0.09	0.00	1.81	-0.10	8.26	-1.55	0.27	-0.01	0.05	0.95	0.14	0.61	-0.09	1.94	-1.13	-0.04	0.08	0.24	0.19
Chems	-0.16	-0.09	-0.13	0.00	-0.02	0.25	0.27	0.11	2.02	-0.01	9.38	-1.58	0.01	0.03	0.11	0.89	0.15	0.56	-0.03	1.38	3.94	0.03	0.08	0.19	0.19
Buseq	-0.20	-0.10	-0.19	-0.04	-0.02	0.17	0.18	0.14	2.08	-0.10	8.58	-1.67	0.03	0.00	0.08	0.92	0.19	0.58	-0.21	1.61	1.92	-0.01	0.10	0.18	0.19
Telcm	-0.17	-0.21	-0.17	-0.17	0.01	0.28	0.26	0.07	2.22	-0.16	6.94	-1.81	0.22	-0.01	0.02	0.98	0.18	0.44	-0.05	3.50	-4.43	-0.03	0.15	0.15	0.24
Shops	-0.16	-0.10	-0.15	-0.04	0.00	0.27	0.05	0.00	2.09	0.00	7.96	-1.94	-0.06	0.03	0.10	0.90	0.11	0.48	0.07	1.14	7.24	0.02	0.08	0.16	0.20
Hlth	-0.20	-0.07	-0.15	-0.06	0.00	0.19	0.38	0.48	2.20	-0.15	7.99	-1.59	0.06	0.02	0.03	0.97	0.23	0.59	-0.24	2.76	-1.30	0.02	0.12	0.14	0.21
Others	-0.18	-0.09	-0.17	-0.05	0.00	0.27	0.13	0.08	2.07	-0.03	8.12	-1.84	0.22	0.01	0.12	0.88	0.14	0.48	0.01	1.19	4.02	0.00	0.09	0.18	0.21
CLUSTER 11																									
NoDur	0.18	-0.27	0.15	-0.43	-4.44	0.26	0.68	0.03	4.31	0.06	4.05	-2.32	-0.16	0.12	0.08	0.92	0.22	0.79	-0.14	0.36	3.87	0.18	0.13	0.02	0.51
Durbl	0.62	0.07	-0.45	-0.06	-3.99	0.21	0.62	0.02	3.91	0.10	3.82	-2.26	-0.07	0.06	0.09	0.91	0.47	0.79	-0.67	2.30	5.57	0.02	0.10	0.04	0.44
Manuf	0.39	0.08	0.09	-0.24	-4.56	0.19	0.73	0.13	4.80	0.07	4.54	-2.48	-0.23	0.16	0.19	0.81	0.56	1.11	-0.83	0.41	4.61	0.17	0.09	0.06	0.47
Engry	-0.01	-0.19	0.24	-0.20	-4.23	0.32	0.45	0.00	4.53	0.00	7.93	-2.39	0.24	0.08	0.19	0.81	0.13	0.51	0.12	8.04	-1.68	0.04	0.10	0.06	0.28
Chems	0.75	-0.50	-0.21	-0.18	-5.96	0.07	1.08	0.90	4.40	-0.12	2.88	-2.04	0.03	0.14	0.22	0.75	0.92	1.27	-1.63	0.25	-1.16	0.15	0.43	0.03	0.63
Buseq	0.96	0.11	0.15	-0.34	-5.17	0.07	1.06	0.39	4.92	-0.07	2.73	-2.32	-0.02	0.18	0.02	0.98	1.50	2.05	-2.84	3.48	-0.21	0.18	0.17	0.02	0.62
Telcm	0.51	-0.07	0.29	-0.32	-5.31	0.21	0.86	0.10	5.33	-0.05	3.85	-2.54	-0.35	0.19	0.00	1.00	1.09	1.68	-1.85	3.36	-1.48	0.16	0.11	0.02	0.54
Shops	0.50	0.04	0.45	-0.46	-4.67	0.15	0.77	0.00	4.64	0.06	3.21	-2.38	0.47	0.35	0.07	0.93	0.48	1.16	-0.75	0.66	6.47	0.21	0.08	0.03	0.59
Hlth	0.83	-0.06	-0.05	-0.38	-5.61	0.11	1.14	2.36	4.94	-0.24	3.31	-2.31	0.36	0.09	0.01	0.99	1.21	1.42	-2.21	5.95	-5.34	0.07	0.27	0.01	0.58
Others	0.61	-0.10	0.35	-0.30	-4.70	0.16	1.03	0.37	5.01	-0.02	3.69	-2.37	-0.18	0.15	0.08	0.92	0.89	1.36	-1.53	0.32	1.81	0.16	0.13	0.02	0.56
CLUSTER 12																									
NoDur	-0.10	-0.10	-0.06	-0.08	-0.02	0.22	0.46	0.01	5.91	0.12	11.51	-2.95	-0.18	0.05	0.61	0.39	0.13	0.59	-0.01	1.97	7.69	0.06	0.06	0.00	0.16
Durbl	-0.10	-0.09	-0.06	-0.07	-0.02	0.20	0.48	0.02	5.97	0.12	12.54	-2.96	-0.17	0.05	0.61	0.39	0.12	0.59	0.01	2.03	7.70	0.06	0.06	0.00	0.14
Manuf	-0.12	-0.11	-0.06	-0.07	-0.01	0.21	0.44	0.03	5.92	0.10	11.79	-2.96	-0.19	0.05	0.61	0.39	0.12	0.57	0.00	1.80	6.56	0.05	0.06	0.00	0.15
Engry	-0.16	-0.21	0.22	-0.18	0.00	0.27	0.49	0.01	6.00	0.05	11.18	-2.96	0.10	0.09	0.34	0.66	0.10	0.57	0.12	3.05	2.90	0.07	0.08	0.00	0.16
Chems	-0.12	-0.10	-0.09	-0.03	0.00	0.23	0.51	0.03	6.16	0.12	11.74	-2.98	-0.18	0.03	0.77	0.23	0.11	0.52	0.04	2.14	6.82	0.04	0.06	0.00	0.15
Buseq	-0.05	-0.07	-0.06	-0.12	0.00	0.09	0.82	0.18	5.89	0.05	9.45	-2.81	-0.07	0.04	0.14	0.86	0.40	0.81	-0.68	2.71	3.44	0.07	0.12	0.00	0.18
Telcm	-0.15	-0.14	0.10	-0.11	-0.02	0.35	0.42	0.04	6.01	0.04	7.14	-2.98	-0.14	0.05	0.30	0.70	0.17	0.50	0.09	2.12	2.00	0.07	0.10	0.00	0.23
Shops	-0.09	-0.09	0.02	-0.11	-0.01	0.20	0.49	0.01	5.95	0.10	10.39	-2.96	-0.15	0.09	0.38	0.62	0.13	0.60	-0.02	2.44	9.14	0.08	0.06	0.00	0.17
Hlth	-0.07	-0.11	-0.04	-0.11	-0.01	0.13	1.10	1.16	5.86	-0.06	9.61	-2.71	0.02	0.04	0.15	0.85	0.44	0.79	-0.71	3.75	0.62	0.07	0.15	0.00	0.17
Others	-0.10	-0.13	0.04	-0.09	-0.01	0.24	0.49	0.05	5.96	0.08	9.82	-2.96	0.18	0.06	0.36	0.64	0.18	0.58	-0.08	2.88	5.42	0.07	0.08	0.00	0.19

The table reports the mean statistics of six clustering criteria and the distribution of industries across 12 clusters. Similar statistics for 19 other attributes of these clusters are also produced (see appendix A.2 for definitions). The six clustering criteria include FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, and LSIZE. The 10 industry groups are constituted using the SIC codes given in Fama and French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) and are defined in appendix C Table C.1.1. Financial and utilities industries are excluded. NoDur represents consumer non-durables. Durbl represents consumer durables. Manuf represents manufacturing firms. Engry includes oil, gas, and coal firms. Chems represents chemicals and allied products industries. BusEq represents business equipment. Telcm represents telephone and television transmission industries. Shops include wholesale, retail, and some service firms. Hlth includes health-related firms. And Others include mines, construction, building materials, transportation, hotels, business services, and entertainment industries.

Table C.1.3

Mean statistics of clustering criteria and other characteristics across industries

Full Sample												
Industry	No. Of Firms	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE
Nodur	5626	-0.06	-0.07	-0.01	-0.03	-0.02	0.24	0.42	0.01	5.21	0.09	10.39
Durbl	2509	-0.05	-0.06	-0.01	-0.01	-0.02	0.23	0.43	0.05	5.07	0.07	10.22
Manuf	11358	-0.09	-0.08	-0.02	-0.02	-0.02	0.23	0.36	0.04	5.06	0.07	10.72
Enrgy	3978	-0.13	-0.19	0.49	-0.07	-0.01	0.25	0.43	0.01	5.37	0.01	10.37
Chems	2327	-0.07	-0.09	-0.01	-0.01	-0.03	0.24	0.53	0.11	5.71	0.07	11.27
BusEq	19044	0.15	-0.05	-0.08	0.00	-0.11	0.12	0.67	0.20	4.87	-0.01	8.55
Telcm	2224	0.00	-0.09	0.37	0.03	-0.05	0.34	0.49	0.05	6.13	0.01	7.18
Shops	9356	-0.05	-0.04	0.05	-0.02	-0.02	0.23	0.39	0.01	4.98	0.07	9.31
Hlth	9488	0.20	-0.07	-0.03	0.07	-0.12	0.16	0.94	1.16	4.99	-0.10	8.60
Others	10797	-0.01	-0.09	0.18	-0.01	-0.03	0.25	0.45	0.07	5.06	0.04	8.92
Micro Firms												
Industry	No. Of Firms	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE
Nodur	3206	-0.031	-0.065	-0.032	-0.027	-0.012	0.241	0.248	0.019	3.636	0.049	9.186
Durbl	1491	-0.018	-0.059	-0.024	0.006	-0.022	0.230	0.341	0.062	3.820	0.030	8.649
Manuf	6365	-0.078	-0.074	-0.034	-0.025	-0.017	0.232	0.244	0.049	3.609	0.038	9.655
Enrgy	2045	-0.093	-0.173	0.454	-0.050	0.006	0.231	0.330	0.006	3.383	-0.045	9.128
Chems	984	0.005	-0.080	0.019	-0.005	-0.033	0.228	0.464	0.211	3.539	-0.012	9.577
BusEq	12450	0.075	-0.062	-0.105	-0.032	-0.093	0.123	0.481	0.223	3.715	-0.057	8.145
Telcm	939	0.054	-0.038	0.345	-0.056	-0.074	0.320	0.424	0.051	4.103	-0.049	5.896
Shops	5590	-0.052	-0.013	0.017	-0.028	-0.021	0.245	0.208	0.005	3.634	0.034	8.142
Hlth	6226	0.199	-0.075	-0.071	0.021	-0.143	0.144	0.821	1.386	3.919	-0.180	7.782
Others	6288	-0.013	-0.075	0.132	-0.014	-0.021	0.257	0.318	0.102	3.726	0.003	7.902

Small Firms

Industry	No. Of Firms	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE
Nodur	1027	-0.067	-0.085	0.028	-0.012	-0.048	0.238	0.498	0.010	6.007	0.121	10.124
Durbl	539	-0.064	-0.068	0.020	-0.007	-0.022	0.217	0.541	0.022	6.035	0.121	11.442
Manuf	2615	-0.078	-0.088	0.013	0.008	-0.018	0.220	0.473	0.031	5.990	0.108	10.649
Enrgy	796	-0.166	-0.251	0.769	-0.141	-0.051	0.287	0.503	0.014	6.099	0.061	9.712
Chems	516	-0.090	-0.094	-0.076	0.032	-0.038	0.227	0.521	0.044	6.109	0.111	11.633
BusEq	3585	0.338	-0.026	-0.055	0.055	-0.211	0.102	0.920	0.180	6.097	0.062	8.154
Telcm	503	0.052	-0.105	0.473	0.125	-0.022	0.358	0.518	0.056	6.100	0.027	5.776
Shops	2072	-0.003	-0.055	0.124	-0.002	-0.042	0.205	0.560	0.009	6.081	0.111	9.429
Hlth	1846	0.306	-0.063	-0.004	0.151	-0.133	0.159	1.156	1.075	6.067	-0.022	8.471
Others	2581	0.036	-0.115	0.262	0.009	-0.075	0.251	0.569	0.048	6.163	0.085	8.860

LargeFirms

Industry	No. Of Firms	FIN_FLEX	ST_CREDIT	LT_INV	CVT_DEBT	PSK_USE	LEV	LogMTB	RDI	LogSIZE	ROA	AGE
Nodur	1393	-0.121	-0.084	0.000	-0.039	-0.012	0.252	0.741	0.006	8.249	0.159	13.349
Durbl	479	-0.135	-0.065	-0.021	-0.046	-0.028	0.250	0.560	0.021	7.889	0.118	13.695
Manuf	2378	-0.123	-0.090	-0.006	-0.026	-0.029	0.241	0.550	0.023	7.935	0.116	13.639
Enrgy	1137	-0.171	-0.187	0.350	-0.064	-0.019	0.263	0.547	0.005	8.432	0.084	13.067
Chems	827	-0.147	-0.101	0.004	-0.029	-0.011	0.260	0.626	0.023	8.033	0.130	13.046
BusEq	3009	0.229	-0.030	-0.030	0.084	-0.070	0.124	1.152	0.122	8.190	0.103	10.665
Telcm	782	-0.086	-0.131	0.341	0.086	-0.035	0.355	0.562	0.045	8.591	0.067	9.629
Shops	1694	-0.076	-0.089	0.096	-0.037	-0.018	0.197	0.764	0.002	8.081	0.138	12.996
Hlth	1416	0.043	-0.066	0.096	0.184	-0.008	0.211	1.153	0.301	8.283	0.119	12.270
Others	1928	-0.069	-0.113	0.231	-0.039	-0.024	0.240	0.696	0.006	7.945	0.106	12.296

This table reports the mean statistics of six clustering criteria and the other firm characteristics (FIN_FLEX, ST_CREDIT, LT_INV, CVT_DEBT, PSK_USE, LEV, LogMTB, RDI, LogSIZE, ROA, AGE as defined in Appendix A.2) across 10 industries. The 10 industry groups are constituted using the SIC codes given in Fama and French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) and are defined in appendix C Table C.1.1. Financial and utilities industries are excluded. NoDur represents consumer non-durables. Durbl represents consumer durables. Manuf represents manufacturing firms. Enrgy includes oil, gas, and coal firms. Chems represents chemicals and allied products industries. BusEq represents business equipment. Telcm represents telephone and television transmission industries. Shops include wholesale, retail, and some service firms. Hlth includes health-related firms. And Others include mines, construction, building materials, transportation, hotels, business services, and entertainment industries.