CORPORATE LEVERAGE IN THE US: ECONOMIC EXPLANATION AND MACHINE LEARNING EXPLORATION

by

Fanghua Lin

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Financial Services Analytics

Fall 2021

CORPORATE LEVERAGE IN THE US: ECONOMIC EXPLANATION AND MACHINE LEARNING EXPLORATION

by

Fanghua Lin

Approved: ______________________________

Bintong Chen, Ph.D.
Chair of Financial Services Analytics

Approved: ______________________________

Bruce Weber, Ph.D.
Dean of Alfred Lerner College of Business & Economics

Approved: ______________________________

Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ______________________________

Paul Laux, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ______________________________

Gonzalo R. Arce, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ______________________________

David A. Edwards, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ______________________________

Xiao Fang, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

David R. Stockman, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Mei Wang, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my dissertation advisor, Prof. Paul Laux, for the invaluable guidance in my Ph.D research. His vision, immense knowledge and empathy helped me in all the time of research in finding dissertation topics, writing this dissertation and presenting the research works as clearly as possible. Since I started my full-time job in 2017, it has been really difficult to find enough time in workdays to work on my dissertation, but Prof. Laux has been really kind to spend his weekends commenting and discussing with me in how to develop the dissertation topics with rigorous studies. I could not have imagined having a better advisor and mentor for my Ph.D study.

I also cannot express enough thanks to other dissertation committee for their encouragement and insightful comments: Prof. Gonzalo R. Arce, Prof. David A. Edwards, Prof. Xiao Fang, Prof. David R. Stockman, and Prof. Mei Wang. It has been an extremely great opportunity to learn from them.

My sincere gratitude also go to Prof. Bintong Chen for the tremendous efforts and kindness in offering so many precious opportunities in connecting with academic professors and industry practitioners such that we can get well-prepared for our career academically and at the same time stay tuned with industry practices regarding how these cutting-edge theories and studies have been applied in industry.

I also want to thank Dr. Yuntao Zhu, my manager when I worked in JPMorgan Chase Bank, Dr. Yuting Jia and Dr. Kiyoung Yang, my managers in Microsoft, for offering me full-time jobs such that I can sharpen my understanding of machine learning theories when applying these techniques in practice to tackle industry problems.

Last but not the least, I would like to thank my parents and siblings for supporting me spiritually throughout my life such that I can be fully dedicated to my

research and work. I also want to thank all the people who have supported me directly or indirectly in this journey with heartfelt thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation studies corporate leverage in the United States from the dual perspectives of economic explanation and machine learning prediction. Corporate leverage in our study is defined as the ratio of debt to assets. Our economic study proposes and investigates an explanation for the coexistence of two leverage patterns across US firms: low (or even zero) leverage for some firms and also some firms with very high leverage, even as the mean leverage ratio remains fairly stable over time. These patterns of the cross-sectional distribution of leverage have become even more pronounced in recent years.

The explanation we propose centers on production flexibility, defined as the ability of firms' to quickly adjust their output in response to demand conditions. Production flexibility is found to have increased in recent years, and is a companion pattern to the low-leverage pattern and the increasing high-leverage pattern. We posit that the combined influence of production flexibility along with the state of a firm's financial constraints and its lenders' concerns about being taken advantage of by managers or shareholders leads to a plausible explanation for the patterns in leverage. The core idea is that when financial constraints do not exist, high production flexibility indicates a low operation risk, which enables firms to take a high financial risk from debt as a substitute. Furthermore, firms where agency problem concerns are not so great can easily benefit from production flexibility to achieve growth, and the growth leads to lower leverage. However, for firms where agency problems are of greater concern to lenders, the production flexibility might be used to the detriment of lenders, leading to a positive relationship between leverage and production flexibility in the cross-section.

Our model predicts that, as production flexibility has increased over the decades, we should observe a stronger low-leverage pattern for firms with less serious agency

problems, and we should observe growing leverage for firms with more serious agency problems in extreme cases. We conduct an empirical investigation of these issues, and find results consistent with this reasoning, especially for firms listed in Nasdaq. We do not propose that growing production flexibility is the only reason for the leverage patterns we investigate, but we believe the reasoning and evidence suggests it is part of the explanation.

Secondly, we conduct a study from a machine learning perspective. Our empirical finance study is conducted using standard panel-data regressions methods. Such methods typically involve controlling for a large set of unexplained "fixed effects" in the cross-section, leaving only part of the variation to be explained by economic factors. The fixed effects are, mechanically, dummy variables whose coefficients need to be estimated. In the context of relatively short time series relative to the breadth of the cross-section, panel-data regressions are not usually considered to be useful for prediction. For prediction purposes, we reason that it may be useful to avoid using the information in the data to estimate pre-specified fixed effects, but instead to introduce clustering dummies only as needed in the model training process sequentially. The training algorithm we suggest is a variation on the traditional machine learning method "Gradient Boosting Machine" (GBM for short) and is called "PanelGBM" in our study. This approach relaxes assumptions of linearity which are inherent in a standard panel-data regression approach as well as other assumptions regarding clusters and data in other machine learning models that are designed for panel data analysis. Via a simulation study, we establish that our algorithm provides good accuracy and stability relative to other prediction methods for a variety of underlying true models and these clustering dummy variables that are learned in the model training process help improving model performance significantly. The algorithm could have applications in a variety of settings; in the context of this dissertation, we apply the algorithm to predict leverage for US firms and find that it performs much better than other models. Even a linear model with the clustering dummies learned from the PanelGBM performs much better than traditional panel regression model.

Chapter 1

INTRODUCTION AND MOTIVATION

This dissertation is motivated by different trending patterns in the financing of US corporations: the proportion of companies adopting zero/low-leverage have been increasing over the years. At the same time, high-leverage firms have also been increasing, but the mean leverage has remained stable. In our study, corporate leverage is defined as the ratio of debt to assets.

While the average leverage ratio has been stable, the dollar value of corporate debt overall has been growing for decades (Graham et al., 2015). Corporate debt has been an important component of financial systems designed to be a buffer to absorb economic shocks (Dudley, 2010). However, International Monetary Fund (IMF for short) has cautioned that high volume of company's debt could pose a threat to the financial system instead of being a risk mitigation tool (Robertson, 2019). The growing level of corporate debt can be seen from the first two panels of Figure 1.1,[1] which are constructed with Compustat data covering the years from 1962 to 2017. Panel (a) shows that the overall debt level of non-financial firms has grown from 68 billion in 1962 to 5.8 trillion in 2017 even though the number of covered companies had fallen from around 7,000 in 1996 to around 2,000 in 2017. Panel (b) provides a different perspective of the same phenomenon showing the average liability of Compustat firms, which grew from 0.05 billion in 1962 to 5 billion in 2017.[2] Even so, Panel (c) shows leverage ratios measured as the ratio of total liability to total asset, the ratio of total

[1] Three measures are explored according to the different definitions of debt in traditional research: total liability, total debt, and long-term debt

[2] The dollars are not inflation-adjusted dollars

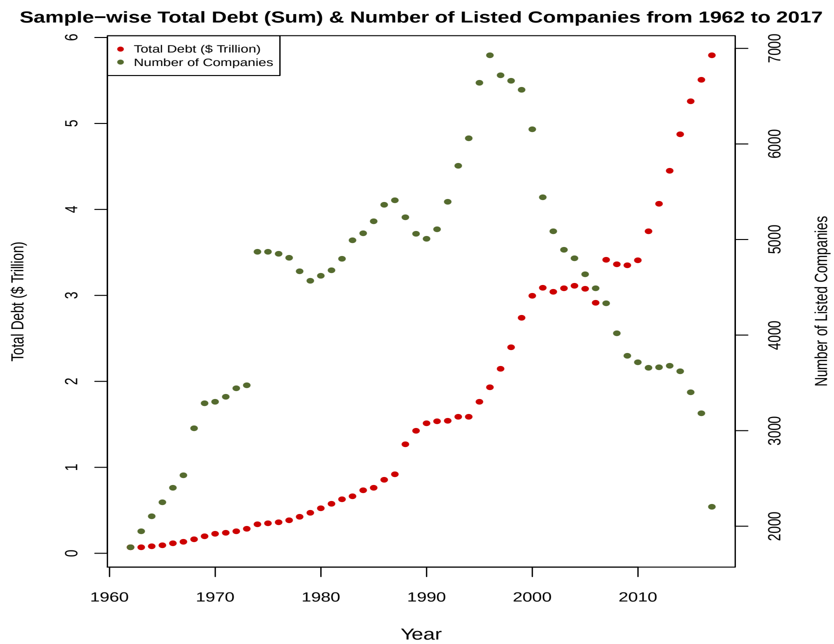
debt to total asset, and the ratio of total long-term debt to total asset, are fairly stable over the years.[3]

Changes in the tails of the leverage distributions can be seen when the full distribution of leverage in the cross-section of firms, not just its mean level, is considered. Figure 1.2 exhibits the rolling comparison of leverage distributions for each decade under study. For Figure 1.2, the leverage is measured as the ratio of total debt to total assets. Each plot in each panel presents the cross-sectional distribution for a particular year estimated using kernel density method. The 57 years from 1962 to 2017 are divided into six roughly decade long periods: 1962 to 1970, 1971 to 1980, 1981 to 1990, 1991 to 2000, 2001 to 2010, and 2011 to 2017. Each timespan is represented with the same color in all panels.

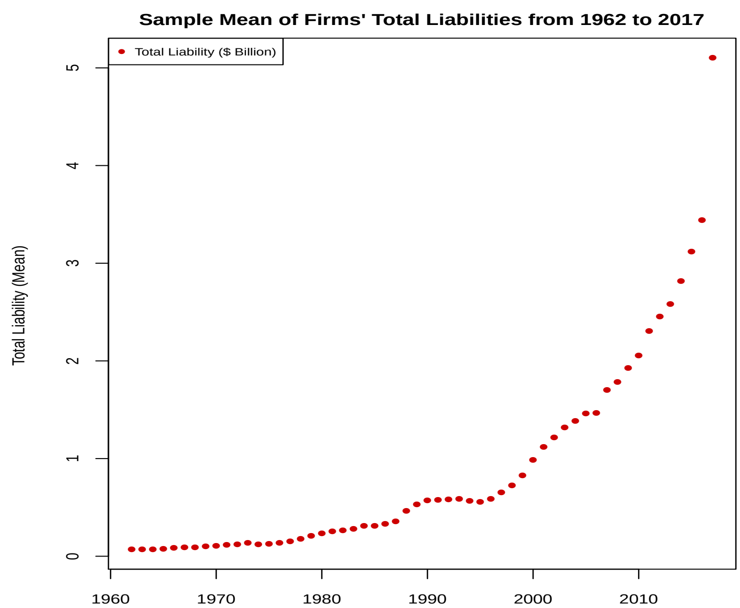
The full comparison across all six decades is shown in Panel (a), and comparisons between each pair of consecutive decades are shown in Panel (b) through Panel (f). From Panel (b) to Panel (f), dotted lines represent leverage distributions in later decades while solid lines represent leverage distributions in previous decades. Years classified as one group are represented by a same color.[4] Based on the comparison of leverage distributions across decades, it is clear that the left tail of the leverage distribution has been growing while the middle section has been decreasing since the 1970's (in orange). For instance, in Panel (c), the densities of years in the 1980's (in blue) have fatter left tails than those of the 1970's. Moreover, in the middle ranges of the densities, the densities for the 1980's are at a lower level than those of the 1970's. Finally, in the right-most ranges, the leverage distributions of the 1980's show slightly heavier tails than those of the 1970's. The same story is also observed in the comparisons of leverage distribution for other consecutive decades from Panel (d) to Panel (f). However, in Panel (d) and Panel (e), the trends of right tails of leverage

[3] Details of these statistics refer to Appendix Table B.1

[4] More details about density color: densities of leverage from 1962 to 1970 are in lime green, densities of leverage from 1971 to 1980 are in orange, densities of leverage from 1981 to 1990 are in blue, densities of leverage from 1991 to 2000 are in pink. densities of leverage from 2001 to 2010 are in moderate green, densities of leverage from 2011 to 2017 are in yellow.

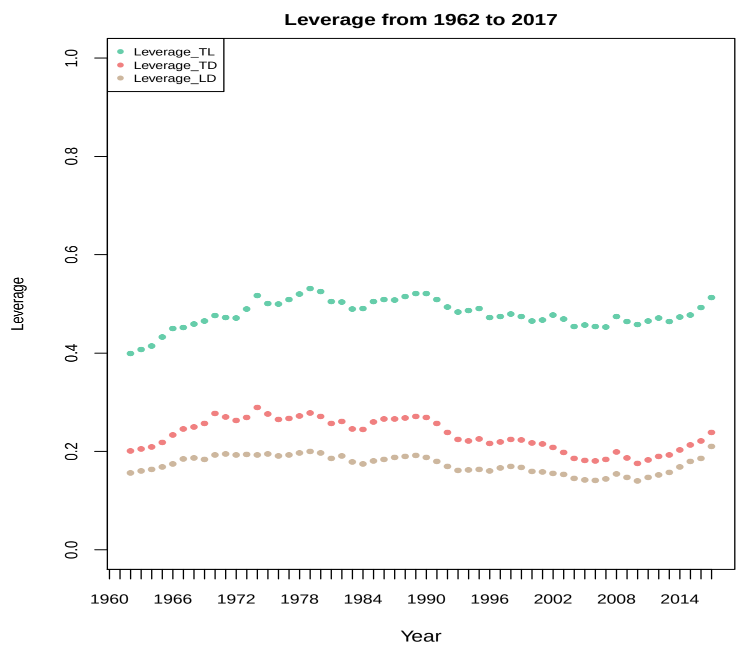

Sample−wise Total Debt (Sum) & Number of Listed Companies from 1962 to 2017
Total Debt ($ Trillion)
Number of Companies
Total Debt ($ Trillion)
Number of Listed Companies
Year
0
1
2
3
4
5
6
2000
3000
4000
5000
6000
7000
1960
1970
1980
1990
2000
2010

(a)


Sample Mean of Firms' Total Liabilities from 1962 to 2017
Total Liability ($ Billion)
Total Liability (Mean)
0
1
2
3
4
5
1960
1970
1980
1990
2000
2010

2.pdf

(b)



(c)

Figure 1.1: Debt and leverage evolution from 1962 to 2017

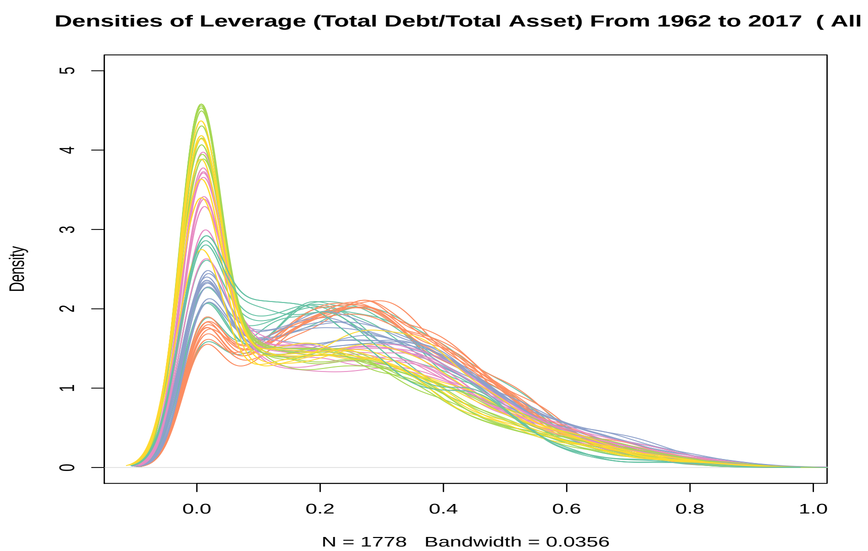
This figure illustrates the trends of debt and leverage from 1962 to 2017. Debts are measured in three different ways based on the research in the literature. The sample-wide sum and firm mean of debt, measured as total liabilities, are shown in Panel (a) and Panel (b) respectively, while the evolution of leverage is measured in three different ways, i.e., the ratio of total liability to the book value of total asset, the ratio of total debt to the book value of total asset, and the ratio of total long-term debt to the book value of total asset. The corresponding lines are marked as "Leverage_TL", "Leverage_TD", and "Leverage_LD" in Panel (c). For Panel (b) and Panel (c), the leverage ratio is calculated at company level and then averaged across the sample.

distributions in related decades seem to be decreasing, but it then turns increasing in the most recent decade as shown in Panel (f). Clearly, the evolution toward fatter right tails is not as clearly evident as that toward the left tails. The growing fatter left tails of the leverage densities reflecting the growing proportion of firms that have low leverage are one manifestation of the so-called zero/low-leverage phenomenon that has been studied(Strebulaev and Yang, 2013). The fatter right tails of the recent leverage densities are one source of industrial concerns over potential decrease in financial stability (Torry and Goldfarb, 2017).

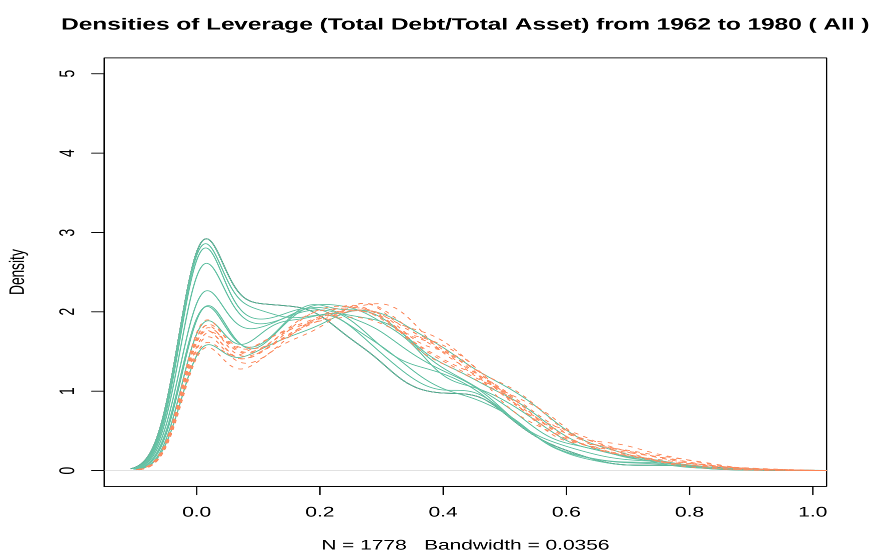
The characteristics of leverage distributions above motivate a study of the evolution of extreme leverage policies by US firms.[5] In this dissertation, I first explore one economic mechanism behind the increasing proportions of low-leverage firms. I develop a theoretical model and an empirical panel data study that show how the interaction of production flexibility[6] and lenders' concerns about being taken advantage of affect leverage decisions under different financial constraint situations. In this context, production flexibility means the ability of firms to adjust output based on market demand. Following this study from an economic perspective, I propose a machine learning study that modifies an existing machine learning training process to conquer some limitations of the panel data regressions used in the first study. The purpose is to introduce dummy variables that capture heterogeneous effects efficiently and improve model prediction performance. The proposed modified method is called "PanelGBM" (GBM is short for "Gradient Boosting Machine"). The main modification is in the way dummy variables, which represent individual effects, are introduced. I show that PanelGBM can improve predictive power by relaxing the linearity assumption of parametric models and categorizing different individual units efficiently. The method is shown to work well with both simulated data and in an application to the real leverage

[5] Chapter II provides a more detailed summary of the descriptive statistics

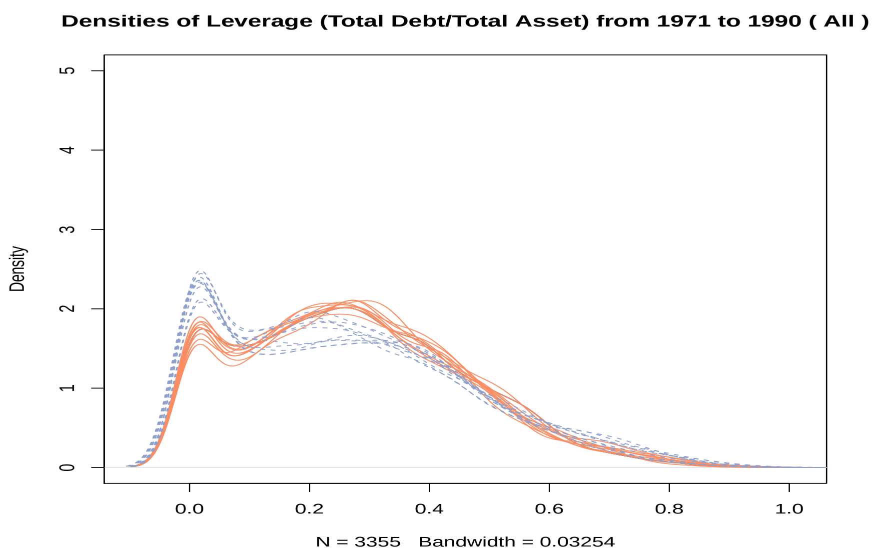
[6] In our essay, we narrow down production flexibility to volume flexibility, one type of production flexibility, when it comes to concrete measurement and calculation, but we will use production flexibility to show the concept since volume flexibility is a too narrowed concept to discuss about the idea thoroughly.


Densities of Leverage (Total Debt/Total Asset) From 1962 to 2017 (All)
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1778 Bandwidth = 0.0356

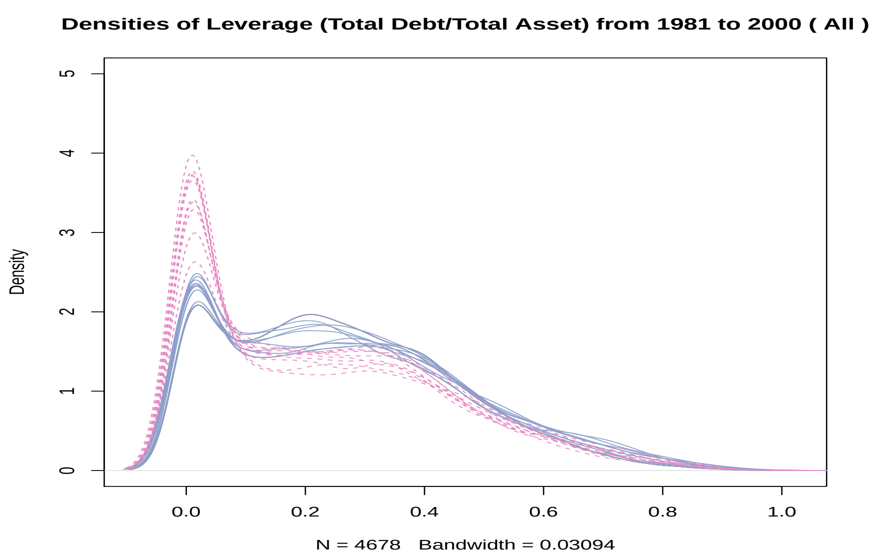
(a)


Densities of Leverage (Total Debt/Total Asset) from 1962 to 1980 (All)
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1778 Bandwidth = 0.0356

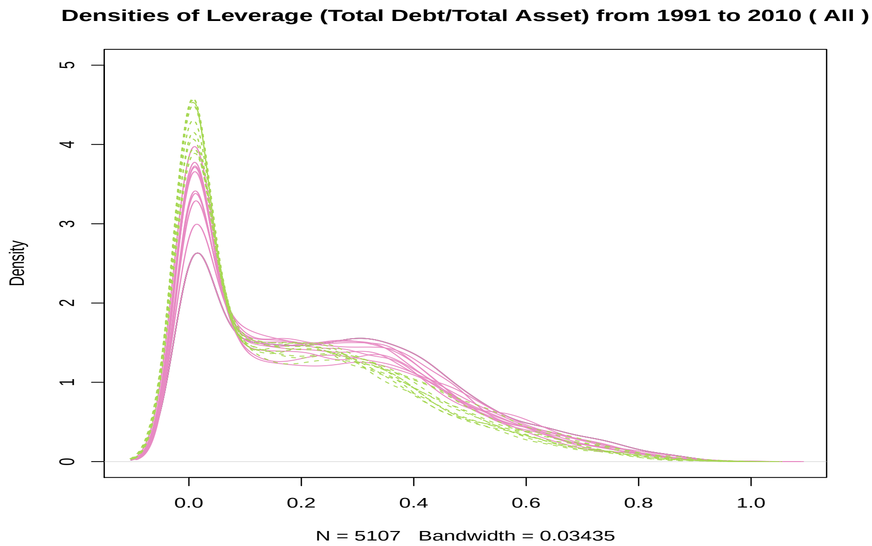
(b)


Densities of Leverage (Total Debt/Total Asset) from 1971 to 1990 (All)
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 3355 Bandwidth = 0.03254

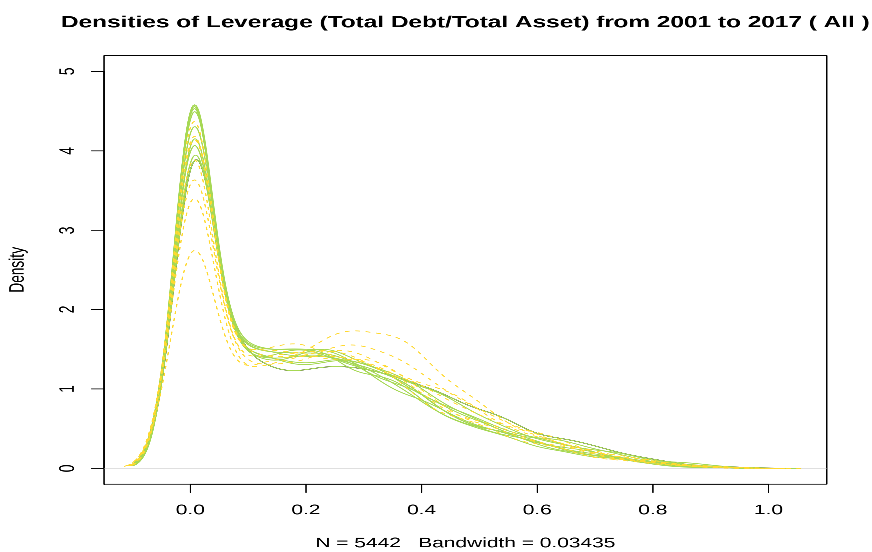
(c)


Densities of Leverage (Total Debt/Total Asset) from 1981 to 2000 (All)
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 4678 Bandwidth = 0.03094

(d)



(e)



(f)

Figure 1.2: Rolling comparison of leverage (total debt/total assets) from 1962 to 2017 for all firms

This figure exhibits a rolling comparison of leverage distributions every ten years. The leverage here is measured as the ratio of total debt to total asset. The general comparison is shown in (a), and the comparisons between every two ten years are shown in Panel (b) to Panel (f). From Panels (b) through (f), dotted lines represent years in later decades while solid lines represent years in previous decades and years classified as one group are represented using the same color. Seen from these figures, fatter low tails can be easily observed over decades while fatter high tails could be clearly observed in recent decades. Details about density color: densities of leverage from 1962 to 1970 are in lime green, densities of leverage from 1971 to 1980 are in orange, densities of leverage from 1981 to 1990 are in blue, densities of leverage from 1991 to 2000 are in pink. densities of leverage from 2001 to 2010 are in moderate green, densities of leverage from 2011 to 2017 are in yellow.

distributions. The next two subsections introduce these two separate studies in turn.

1.1 Finance Study

The economic study explores the interaction of agency problems, production flexibility and a financial constraints on leverage decisions. For our purposes, the agency problem in our paper is defined as the conflicts of interest between existing shareholders/managers and debtholders where shareholders/managers could invest in some self-benefits-seeking projects that could adversely affect debtholders. The financial constraint is modeled as the inability to fully finance production activities with equity. The study consists of a theoretical model and development of empirical evidence. The theoretical model illustrates the role of the three mentioned factors and the empirical study supports the logic illustrated in the theoretical study.

The theoretical model builds on an existing model outlined in existing research (Aivazian and Berkowitz, 1998) with some modifications. Key modifications include different optimization goals, introduction of the agency problem concern and introduction of financial constraint. As argued below, these are important because they increase the realism and empirical applicability of the model. There exist two different optimization goals of shareholders: one is firm value optimization and the other is equity value optimization. Different goals could lead to different leverage decisions, and thus are both discussed in the dissertation. The extent of agency problems and financial constraints varies across firms. The model shows that the relationship between production flexibility and leverage varies based on the firm's objective function, on whether there is a constraint in financial markets, and on whether investors are concerned about agency problems. If managers aim to maximize equity value, whether there exist agency problem concerns determines whether financial constraints affect the relationship between production flexibility and leverage. When there are agency problem concerns, production flexibility is positively related with leverage since high production flexibility exacerbates lenders' concerns and decreases firm value. The positive relationship holds when there is no agency problem concern or financial constraint since high production flexibility enables firms to consume more financial risk. However, when there is no concern with agency problems but exist financial constraints, higher production flexibility leads to lower leverage since high production flexibility helps increasing firm value and enables firms to borrow less to satisfy financial needs. Henceforth, the theoretical model builds up a possible mechanism that points to the existence of the extreme leverages mentioned earlier.

In addition to the theoretical exploration, an empirical study is undertaken to test the role of the three mentioned factors. The data comes from annual financial data from Compustat and Center for Research in Security Prices (CRSP). The time period of the data is from 1962 to 2017 and only non-financial companies that are publicly listed are kept in the data. In general, production flexibility has been shown to correspond to lower leverage over time. Further analysis based on the segmentation of market access to financial market suggests that production flexibility plays an important role in leverage only when the access to financial market is constrained. Furthermore, whether there exist concerns due to agency problem helps to explain the different roles of production flexibility with respect to leverage. When that concern is high, production flexibility is positively related with leverage, meaning that higher production flexibility leads to higher leverage. On the other hand, when the concern is low, production flexibility is negatively related with leverage, meaning that higher production flexibility leads to higher firm value and thus, lower leverage. The same results are found for low-leverage decisions and low-leverage transition decisions where low-leverage transition decisions are the decisions about firms going from non low-leverage state to low-leverage state (for more details, refer to "Empirical study" section). Reasonably, the roles of production flexibility on high-leverage decisions and high-leverage transition decisions are opposed to what we find in the low-leverage study where high-leverage transition decisions are the decisions about firms going from non high-leverage state to high-leverage state (for more details, refer to "Empirical study" section). These results remain robust even after controlling more firm characteristics and macro conditions.

The theoretical and empirical studies constitute my economic study on the evolution of leverage in the economy. However, the low predictive power of the traditional panel data analysis method is noticed. Thus, in the second section, we try to modify an existing machine learning training algorithm for better model prediction in panel data.

1.2 Machine Learning Exploration

As described above, the first section of this dissertation explores the long-run evolution of corporate leverage, using economic reasoning to develop a theoretical model and then using panel data regressions to develop inference for the model. In keeping with the tool-building focus of the Financial Services Analytics Ph.D. Program, the second section of this dissertation extends the panel data regression method in a machine learning direction with a focus on prediction. The proposed new method has predictive performance that is superior to panel regressions when applied to the leverage data. Moreover, the new machine learning method also outperforms panel regression in a simulation study, suggesting that this new tool may have usefulness in other situations as well.

There are several reasons why the panel regression methods that are commonly applied for inference about underlying economic mechanisms might be less successful for the purpose of prediction. Such regression models are usually specified based on economic theory or intuitive reasoning about how regressors should relate to an outcome. The models or intuitions often do not imply strictly linear relationships. As shown in the first section of this dissertation, effects that vary across the domains of the regressors may be present, as may interactive effects. Yet a linear conditional mean function is often estimated, whether due to lack of full knowledge of the true model or based on the attractive statistical properties of least squares regression methods. Moreover, models often reason about only part of what is going on in an economic process, with other parts of the process perhaps being better understood with a different model. One practical reaction to such model incompleteness is to estimate fixed-effects regressions, in which cross-sectional or time-series effects that cannot be well-understood with the model are effectively “dummied out”. Dummy variables are commonly defined to account for the cross-group outcome differences of nominal data groupings or categories within which the economic model is thought to be valid.

When it comes to prediction, both the emphasis on linear effects and the use of categorical dummy variables can potentially limit performance. The limitations flow essentially from the fact that the regression does not apply the true data generating process or true model. A linearized prediction rule applied in the context of a non-linear prediction process must produce biased predictions, according to the difference between the prediction model and the true model. Additionally, a researcher’s best assessment of the categorical dummy variables will often differ from the set that generates predictions with minimum error, unless the researcher is able to choose just the right number of dummy variables (dividing categories neither too finely nor too coarsely), and also just the right qualitative definitions of categories (industries, time periods, geographical locations, etc.). Too coarse a set of dummies forces bias onto the predictions, whereas too fine a set brings undue estimation noise. Incorrect qualitative definitions bring both bias and noise.

The method that is developed and tested in the second section of this dissertation is designed to overcome these limitations. Essentially, the proposed method overcomes bias in predictions by using a flexible machine learning algorithm as a curve fitter to optimize out-of-sample prediction, controlling over-fitting in the usual way via cross-validation using training-testing splits of the data. For specificity, I choose a gradient boosting method, though other methods could also be used as a starting place. I then augment the method to create a “PanelGBM” method. PanelGBM chooses how many and which nominal groupings to consider in the predictive model as part of the repetitive training and testing routine.

PanelGBM utilizes GBM to exploit the prediction power of independent variables and generates dummy variables, which are proxies of fixed effects dummies, in the model training process to cluster different sections into groups. The prediction improvement of “PanelGBM” comes from the relaxation of linearity assumption and efficient introduction of dummy variables that deal with specific effects related with individual sections. The limitation of linearity assumption in traditional panel data method is in the ignorance of potential non-linearity relationship between independent variables and dependent variables and of the interaction between independent variables and dummy variables. Furthermore, the heuristic method of introducing one dummy variable for each section is not efficient nor necessary since some sections may have same effects. Realizing these limitations of current method, we first choose Gradient boosting machine (GBM for short henceforth), which is famous for their exploitation of non-linearity, to mitigate the limitation of linearity assumption. Gradient boosting is a machine learning technique that utilize ensemble methods based on weak prediction models to solve regression and classification problems. Each weak prediction model is built sequentially in a stage-wise fashion such that the weak model in the next step is trained to minimize the errors not learned in the previous steps. We further modifies GBM by introducing dummy variables into the model sequentially to capture individual effects. These dummy variables are integrated into the training process for grouping sections with similar section-level fixed effects. This modified GBM method is called PanelGBM in our study.

The effectiveness of the PanelGBM method is explored via simulation studies on data generated by various true models. For every data generation process examined, the PanelGBM outperforms the panel regression for prediction. The outperformance is most substantial for cases where non-linearity is most prominent and for cases where an extensive set of fixed effects would bring mainly estimation noise. Furthermore, PanelGBM also outperforms other existing machine learning methods that have been proposed in panel data analysis by leveraging less assumptions in grouping, which will be discussed in detail in the machine learning section.

1.3 Summary

The first section of the dissertation explores an economic mechanism behind the growing proportions of zero/low-leverage firms and that of high-leverage firms, while the second section proposes a modified method to improve model prediction performance. As I write these words during the late stages of work on my dissertation, the COVID-19 pandemic is sweeping the world. Facing the economic disruption triggered by the COVID-19 disease, high levels of corporate debt remain a center of worry since they might amplify later problems and worsen the economy. I hope that adding to the understanding of the seemingly conflicting zero/low-leverage pattern exists in the context of growing high-leverage companies can help practitioners and governments in decision making and corporate regulation.

The rest of the research is organized as follows. Part 1 contains the whole process of the traditional finance study described above including theoretical research and empirical study. Part 2 summarizes the proposed modified machine learning algorithm and model results. Finally, in Part 3, we conclude this dissertation with a summary of our finding and some thoughts on future work.

Part I

FINANCE STUDY

Chapter 2
INTRODUCTION

This chapter extends the consideration of some facts presented earlier, especially focusing on the difference between NYSE and Nasdaq firms. The extended consideration of these facts suggests an intuitive motivation for a theoretical consideration of changing leverage ratios driven by changes in the flexibility of firms' production processes. I present the additional facts and intuitive motivation in separate sections.

2.1 More Facts

As illustrated in the Introduction (Figure 1.1), the debt of public U.S. nonfinancial corporations has been increasing over a long period of time, even as the number of listed companies has been decreasing in recent decades. In contrast, the mean and median corporate leverage ratios have been fairly stable. While these patterns may at first seem different, examination of the cross-sectional distribution of leverage as it evolves over time shows how they fit together. As shown in Figure 1.2 in the Introduction, the proportion of firms in the left (low-leverage) tail of the distribution has been increasing over time, while the proportion in the middle ranges of leverage has been decreasing substantially. The proportion in the right (high-leverage) tail has also been increasing, though the magnitude of the change is much smaller. All things considered, the U.S. now has relatively more companies with low and high leverage (debt-to-assets) and relatively fewer companies with mid-level leverage.

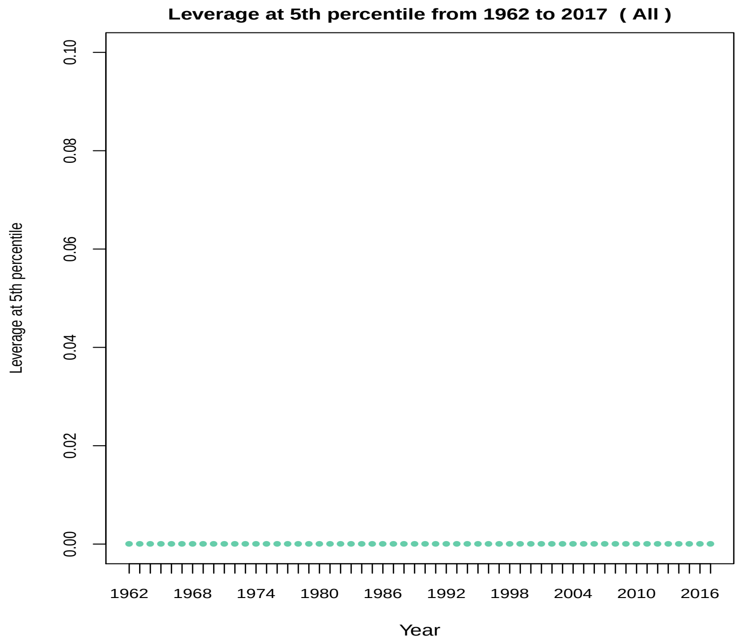
The striking fact is that extreme leverage policies for U.S. firms have become much more common. Using data from Standard and Poor's Compustat, Figure 2.1 characterizes the evolution over time in terms of some representative percentiles of the leverage distribution over 1962 through 2017. As to the left tail of the distribution, Panel (a) shows that there has long been a cohort of zero-leverage firms.[1] The representative firm at the 5th percentile of the leverage distribution has zero leverage over the entire span of 1962 to 2016. Panel (b), however, shows a sharp increase in low-leverage firms. While in 1962 the representative firm at the 25th percentile had about a 5% leverage ratio measured as total debt to assets (where assets is measured as Compustat item AT), and in 1974 the representative firm had about a 15% leverage ratio, the representative firm after 2014 has had an approximate zero leverage.

As to the right tail of the distribution, Panels (c) and (d) show that the representative 75th percentile and 95th percentile firms have maintained about 30% and 60% leverage ratios respectively. Overall, many more firms carry no or very little debt, while some firms carry extremely high leverage. It is the aim of this part of the dissertation to suggest and investigate a particular explanation for the coincidence of more very low leverage firms along with a few extremely high leverage firms.

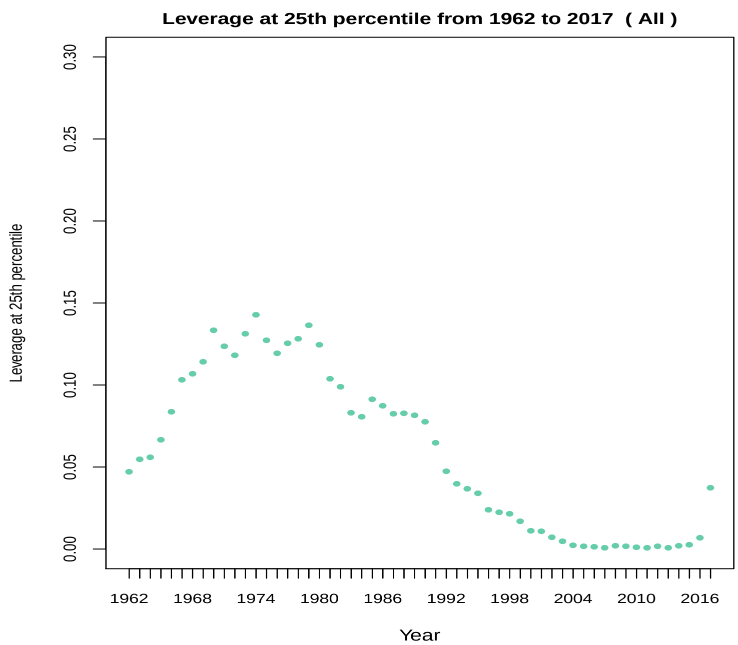
It is the aim of this dissertation to understand at least some of the reasons for these distributional changes over time. One intuitive motivation for our proposed explanation can be seen by considering NYSE and Nasdaq firms separately.[2] This split is interesting to consider because Nasdaq firms are, on the whole, smaller, younger, and more growth oriented than those on the NYSE (even thought these differences are not as stark in recent years).

[1] Zero-leverage firms are defined as those firms with Compustat items DLTT (long-term debt) and DLC (current debt) equal to zero. I refer to the sum of DLTT and DLC as total debt or total book debt in this chapter.

[2] The rolling comparison of leverage (total debt/total asset) distribution from 1962 to 2017 for NYSE firms and Nasdaq firms are illustrated in Appendix Figure A.1 and Appendix Figure A.2, respectively.

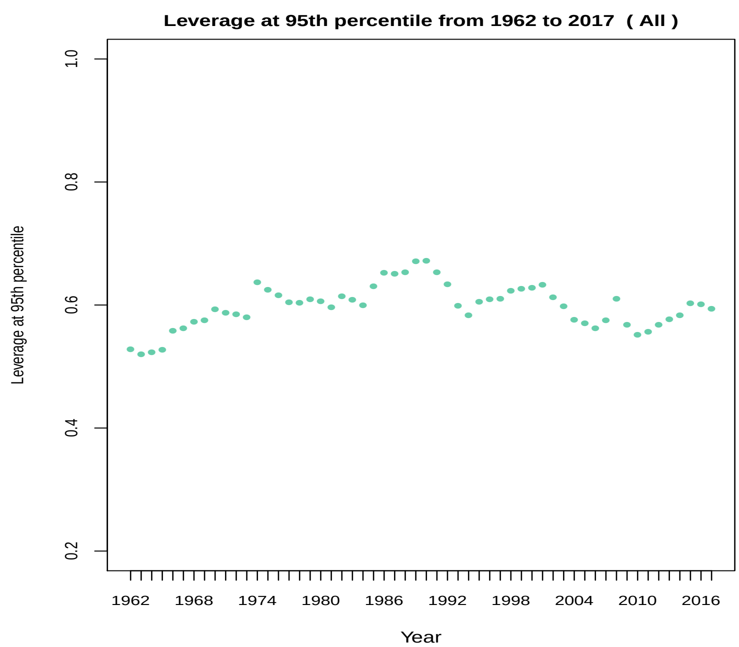

Leverage at 5th percentile from 1962 to 2017 (All)
Leverage at 5th percentile
0.00
0.02
0.04
0.06
0.08
0.10
1962
1968
1974
1980
1986
1992
1998
2004
2010
2016
Year

(a)


Leverage at 25th percentile from 1962 to 2017 (All)
Leverage at 25th percentile
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1962
1968
1974
1980
1986
1992
1998
2004
2010
2016
Year

(b)

(c)



(d)

Figure 2.1: Evolution of leverage (total debt/total asset) from 1962 to 2017 for all listed firms at 5th percentile, 25th percentile, 75th percentile and 95th percentile

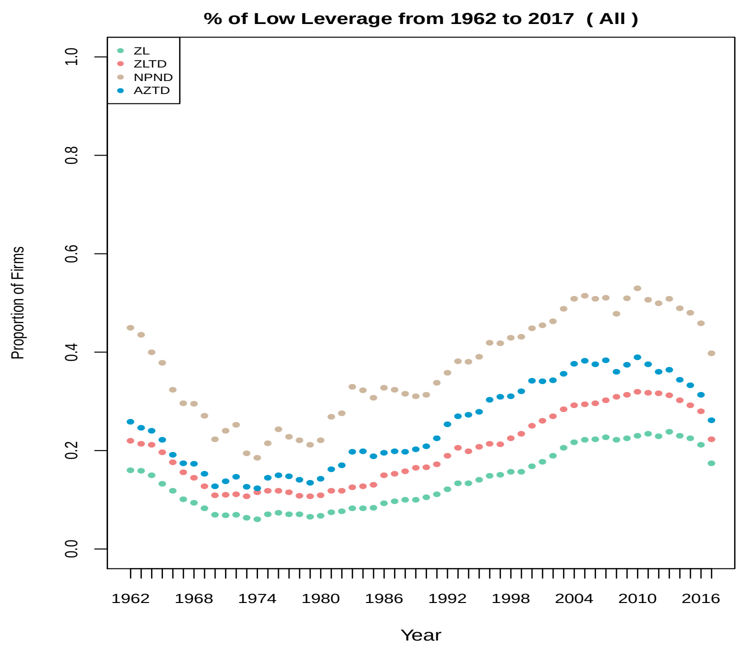
This figure exhibits the evolution of percentile statistics of leverage distribution over years. The leverage is measured as the ratio of total debt to total assets. Each Panel represents the evolution of leverage across years at a certain percentile. The percentile levels from top to bottom are 5th percentile, 25th percentile, 75th percentile and 95th percentile, respectively.

Figure 2.2 illustrates the changes over time in the proportions of zero-leverage and low-leverage firms in different exchange marketplaces, i.e., NYSE and Nasdaq, separately. To show the set-up of the graphical analysis, Panel (a) plots the evolution of the overall proportion of low-leverage firms –NYSE and Nasdaq together—as measured in several alternative ways. The Panel plots the proportions of

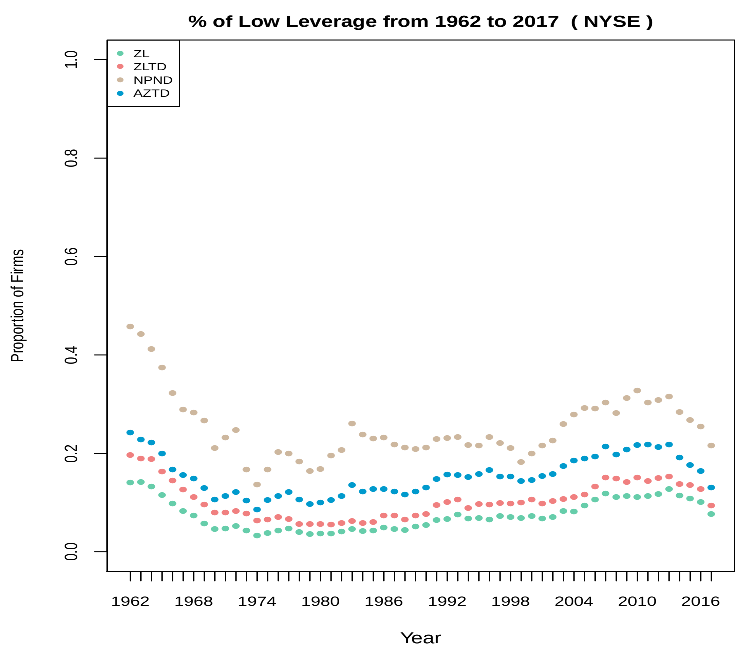
- zero leverage firms (ZL in the figure), defined as firms with zero Compustat total debt (DLTT + DLC) (Strebulaev and Yang, 2013),
- zero long term leverage firms (ZLTD in the figure), defined as firms with zero Compustat long term debt (DLTT) (Strebulaev and Yang, 2013),
- non-positive net debt firms (NPND in the figure), defined as firms with non-positive Compustat total debt (DLTT + DLC) less cash and equivalents (CHE) (Strebulaev and Yang, 2013), and
- almost-zero total debt firms (AZTD in the figure), defined as firms with leverage of less than 5% in the firm-year, with leverage defined as Compustat total debt (DLTT + DLC) scaled by assets (AT) (Strebulaev and Yang, 2013).

The same general patterns established above are evident. For all four measures of leverage, a modest decrease in the proportion of zero or low leverage firms over 1962 to 1974 is evident, followed by a sustained increase over subsequent years up to 2017.

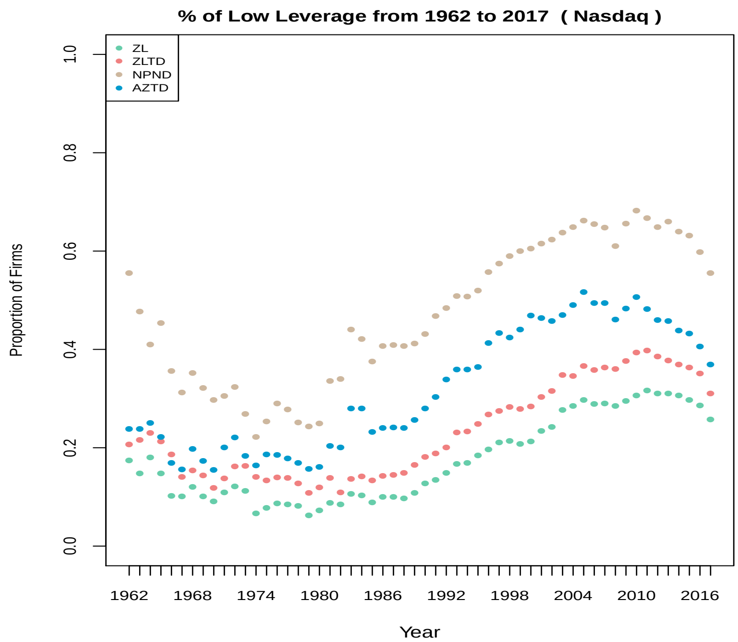
Panel (a) of Figure 2.2 serves as a full-sample benchmark. Panels (b) and (c) provide the analogous plots for non-financial NYSE and Nasdaq firms, respectively. For NYSE firms, the plot of proportions of zero and low leverage firms over time is nearly flat, although modest declines in the early years and increases in the later years are evidenced. For example, the proportion of zero leverage firms on the NYSE is around 30% in most years past 1974. In sharp contrast, plots of all four measures of the proportions of zero and low leverage firms in Nasdaq are sharply higher after 1974. For example, about 20% of Nasdaq firms had zero leverage in 1974, while in 2016 the


% of Low Leverage from 1962 to 2017 (All)
ZL
ZLTD
NPND
AZTD
Proportion of Firms
0.0
0.2
0.4
0.6
0.8
1.0
1962
1968
1974
1980
1986
1992
1998
2004
2010
2016
Year

(a)


% of Low Leverage from 1962 to 2017 (NYSE)
ZL
ZLTD
NPND
AZTD
Proportion of Firms
0.0
0.2
0.4
0.6
0.8
1.0
1962
1968
1974
1980
1986
1992
1998
2004
2010
2016
Year

(b)



(c)

Figure 2.2: Trends of low-leverage statistics for non-financial firms from 1962 to 2017 by market segment

This figure exhibits the comparisons of low-leverage measurements between firms in NYSE and those in Nasdaq. Panels (a) to (c) illustrate the evolution of three low-leverage measurements over years for "Non-financial firms", "firms listed in NYSE", and "firms listed in Nasdaq", respectively. ZL firms have zero book debt (DLTT + DLC = 0, where DLTT and DLC refer to the Compustat long-term debt and debt in current liabilities respectively). ZLTD firms have zero long-term debt (DLTT = 0). Almost all zero-leverage (AZTD) firms have book leverage not exceeding 5% in a given year. NPND firms have non-positive net debt in a given year, i.e., $DLTT + DLC - CHE \leq 0$.

proportion was about 50%. The takeaway is that Nasdaq firms appear the main source of the growth in the lower tail of the leverage distribution in the full sample.

Figure 2.3 has a similar structure to the previous figure, and illustrates the evolution of the upper tail of the leverage distribution for NYSE firms (Panel (b)) and Nasdaq firms (Panel (c)), with a benchmark plot for the full sample as well (Panel (a)). Specifically, the plots show the proportion of firms over time that satisfy the following definitions, respectively:

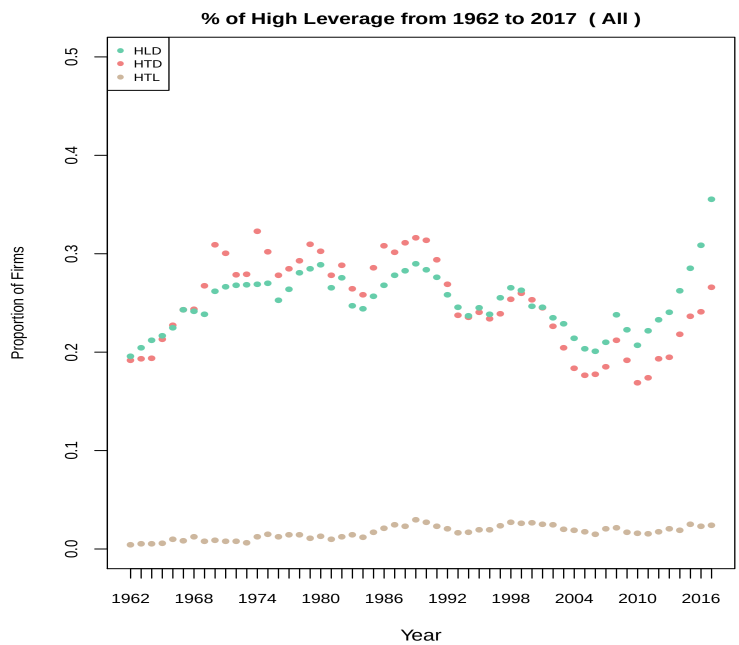
- high leverage firms (HLD in the figure), defined as being above the full time span 75th percentile for the full sample in total leverage based on debt (Compustat items DLTT plus DLC, scaled by assets (AT),

- high long term leverage firms (HTD in the figure), defined as above as being above the full time span 75th percentile for the full sample in leverage as computed with long-term debt only (DLTT) scaled by assets (AT), and

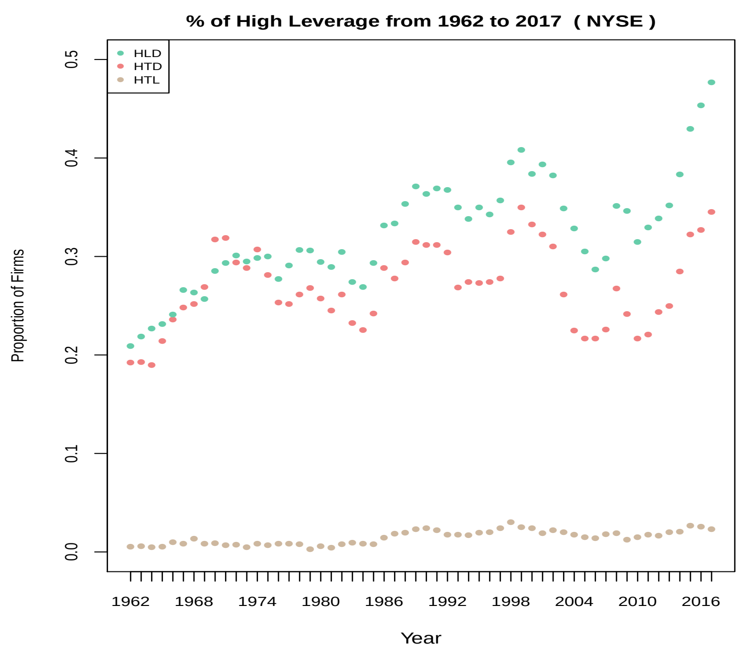
- high total liabilities leverage firms (HTL in the figure), defined as above as being above 0.9 for the full sample in leverage as computed with total liabilities (TL) scaled by assets (AT).

From these plots, it is evident that the source of any growth in the upper tail of the leverage distribution for non-financial public U.S. corporations lies within the set of NYSE firms. In Plot (b), an upward trend in the proportion of high leverage firms as assessed with total debt and long-term debt over 1962-1974 is evident, and that trend seems to continue over 1974-2016 based on leverage assessed with long-term debt. Less evident due to the scaling of the plot is the fact that the proportion of firms with high leverage as assessed with total liabilities has also trended up slightly for NYSE firms. For Nasdaq firms, the pattern is very different. The proportion of high leverage firms based on either measure of debt is reasonably stable until about 1980, then declining until around 2010 before recently increasing back to its 1974 level.

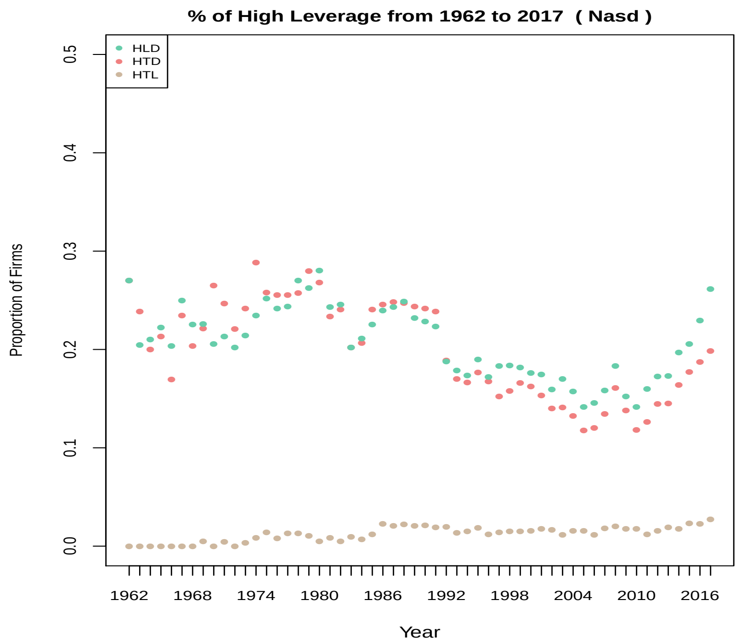
Overall, the descriptive analysis in the introduction of this dissertation along with the additional facts developed in this section provide an summary impression of


% of High Leverage from 1962 to 2017 (All)
HLD
HTD
HTL
Proportion of Firms
Year

(a)


% of High Leverage from 1962 to 2017 (NYSE)
HLD
HTD
HTL
Proportion of Firms
Year

(b)



(c)

Figure 2.3: Trends of high-leverage statistics for non-financial firms from 1962 to 2017 by market segment

This figure exhibits the comparisons of high-leverage measurements between firms in NYSE and those in Nasdaq. Panels (a) to (c) exhibit the evolution of three high-leverage measurements over years for “non-financial firms”, “firms listed in NYSE”, and “firms listed in Nasdaq”, respectively. HLD, HTD, and HTL are the three high-leverage measurements. They are dummy variables indicating whether they exceed certain threshold. For HTD and HTL, the threshold is set as the top 75% of the whole related sample, which is 0.367 for HTD and 0.280 for HLD while the threshold is set at 0.9 for HTL. The difference among HTD, HTL and HLD is that they use different numerators: total debt for HTD, total liability for HTL and long-term debt for HLD.

the evolution of the leverage distribution for non-financial U.S. public corporations. A larger proportion of firms are engaged in zero or low leverage balance sheet situations, with Nasdaq firms being the central source of this pattern. Less strongly, the data suggest that a larger proportion of firms are engaged in high leverage balance sheet situation, with NYSE firms being central to the source of this pattern.

2.2 Motivation for Theoretical and Empirical Work

The reasons for the distributional shifts documented above are not well understood. To investigate them theoretically and empirically is the aim of this part of the dissertation.

Of course, corporate leverage has been studied before in thousands of papers. Literatures have been developed that specifically investigates the phenomenon of low leverage firms, and one that considers how firms can get stuck at high leverage. Moreover, system wide changes, such as the development of private equity, junk bonds, syndicated loans, leveraged loans, collateralized debt obligations, etc. have facilitated higher leverage for firms that choose these tools. These literatures are surveyed in a subsequent chapter, but a few notes are useful at the outset in order to see how these factors play their roles.

In a sweeping study, Graham et al. (2015) examine the evolution of debt for U.S. non-financial firms. They conclude that changes in firm-specific characteristics are unlikely to be the drivers of this evolution, and point to such macroeconomic factors as changes in government lending and borrowing, uncertainty of macroeconomic conditions, and financial sector development. However, macroeconomic policies and economy-wide conditions seem unlikely to be sufficient to explain the disparity of distributional changes we observe, with some firms taking less leverage and some firms taking much more. Moreover, their study focuses on only a few central firm characteristics such as book-to-market ratio, profitability, and size. In this dissertation, I consider if the evolution of firm characteristics might be a part of the explanation.

As mentioned, separate literatures have developed regarding the phenomena of low leverage firms and high leverage firms. In an influential paper, Strebulaev and Yang, 2013 focuse on the phenomenon of low leverage firms. He documents that central firm characteristics such as size and profitability are relevant to firms engaging in a low leverage over protracted periods, as are more situation-specific factors such as being credit constrained and being family owned. Other papers consider the phenomenon of high leverage firms, one prominent recent example being Admati et al., 2018 who argue that a "ratchet effect" leads shareholders to choose only increases in leverage, even when a decrease would add to firm value, because some of the value would accrue to an increase in the market value of the debtholders' claims but shareholders alone would bear the full cost.[3] These papers suggest that firm characteristics could be important at least for the low-leverage part of the evolution of the leverage distribution

In this context, I contribute a perspective with both theory and evidence that 1) exploits changes in an aspect of firms' characteristics that has been little-studied within the context of leverage, and 2) jointly considers the phenomena of low leverage and high leverage firms. Specifically, my perspective focuses on changes in firms' production process flexibility. Changes in technology, communication, data processing, and many other areas, along with labor market changes such as the decline of unionization and input market processes such as more extensive supply chains have all combined to make firms' production processes more flexible and nimble than in the past (Fragapane et al., 2020). One result has been the protracted increase in labor productivity over the past century. The facts in the previous subsection encourage me to investigate using the production flexibility focus. The Nasdaq market is likely the exchange home to firms where increases in production flexibility have been the most substantial over time since firms listed in Nasdaq market is much smaller and easier to adopt new technology than those listed in NYSE market, and it is the home of firms that have disproportionately

[3] The basic idea of the ratchet effect was first laid out by Black and Scholes (1973). In other settings, this is sometimes called the debt overhang effect, and it is known to be strongest when debt levels are already high.

avoided leverage.

Additionally, there is various existing literature on production flexibility and leverage indicates variously that production flexibility can be a complement or a substitute for leverage. On the one hand, production flexibility can reduce default risk by increasing firms' ability to respond to shocks (Reinartz and Schmid, 2016), thus allowing for more borrowing and a higher leverage ratio. On the other hand, it can exacerbate fears of agency conflicts of interest for the managers, and shareholders could use the flexibility to the detriment of debt providers. Anticipating this, debtholders would step back or charge higher interest rates, resulting in less borrowing and a lower leverage ratio (MacKay, 2003).

The reasoning focuses on firms' decisions about borrowing. Existing literature regarding financial flexibility and leverage largely presumes that firms can access debt and equity as desired, i.e., they do not face constraints in financing and only raise debts and equity as needed. Beginning with Fazzari Hubbard and Petersen (Fazzari et al., 1988), however, a large literature has established that firms do face financial constraints and that they have substantial effects on firms' decisions. Such firms cannot make free choices about borrowing. Furthermore, multiple financing methods are not independent on each other. For example, the fact that firms are able to utilize external equity would send positive signal about firm value to potential debtholders and thus help debt financing activities .

Firms that face financial constraints but do not have serious agency problems can use production flexibility profitably, raising firm value and producing a lower debt ratio as a side effect. On the other hand, for financially constrained firms subject to agency problems, high production flexibility would exacerbate managers' ability to abuse free cash flow and, most importantly for my purposes, shareholders' ability to game debtholders. Firm value would be reduced as a consequence, leading to a higher leverage ratio. Strictly speaking, this reasoning applies to leverage ratios computed in market value terms, but would also be reflected in book value leverage ratios to the extent that high (low) market values allow for firms to grow (or restrict) balance sheet assets.

To state these scenarios together most simply, financially unconstrained firms can take high financial risk when they have high production flexibility due to lower default risk, unless agency problems are present that worry lenders. The existence of agency problem pushes leverage up due to the decreased firm value. Financially constrained firms can also use production flexibility to increase firm value unless agency problems get in the way. In this case, leverage is then lower due to the increased firm value.

Studies that have not considered financial constraints have perhaps also not considered to focus on production flexibility as a partial explanation for the evolution of leverage and the changing incidence of high and low leverage firms. After all, Nasdaq firms are more likely to follow low leverage policies and likely have greater production flexibility. But when one considers that Nasdaq firms are also more likely to be financially constrained due to being, on average, smaller and younger than NYSE firms, the possibility that production flexibility can help to explain both low leverage firms seems more plausible. Conversely, it can also help with understanding high leverage policies for NYSE firms.

In the chapters that follow, I develop a formal theory that embodies this intuition. I then develop an empirical investigation to understand the importance of this reasoning for the evolution of leverage ratios, considering that firms vary in the cross section and over time both in terms of the extent to which they are subject to financial constraints and the extent to which they are subject to agency problems. In the empirical work, I delineate between NYSE and Nasdaq firms, with the notion that Nasdaq firms may have more production flexibility (and may differ from NYSE firms in terms of agency problems as well).

Chapter 3

LITERATURE REVIEW

Capital structure has been an important and interesting topic in finance studies over decades. At the beginning of this section, key concepts and commonly accepted theories of capital structure research are reviewed (Section 3.1) since they lay the foundation of the different logic strands this study explores and explain why certain factors are brought in as control variables in our study. We then move further to cover the works that have been conducted on the low-leverage phenomenon (Section 3.2) and high-leverage phenomenon (Section 3.3) since these two phenomena motivate our study. Finally, we go through the studies that explore the role of operational factors (like production flexibility) on financial decisions in Section 3.4.

3.1 General Theory

Capital structure studies how a firm finances its operation activities using a mixture of equity, debt, and other hybrid securities. Equity mainly consists of common stock, retained earnings, and preferred stock. Corporate debt is usually raised in the form of bank loans and firm bonds, both of which could be further categorized as long-term debt or short-term debt based on the term of the debt.

Both debt and equity carry some strengths and shortcomings in different perspectives. Debt seems to be a better choice than equity for its lower financing cost because debt bears lower risk than equity and has tax reduction benefits from the pre-tax interest payments. What makes debt more attractive is that it allows existing company shareholders to retain their ownership and control rights. However, just as a coin has two sides, the potential bankruptcy risk brought by debts due to committed regular interest payments are not negligible. Furthermore, the agency problem brought by debt, such as debt overhang and risk shifting issues, also makes debt financing less attractive. The advantages and disadvantages that debt financing bears are reversed when it comes to equity-financing; equity is riskier and more expensive than debt since the payment to equity investors come from the residual claims of companies. Moreover, existing shareholders would be adversely affected by newly raised equity since their control rights and cash flow rights would get diluted.

Many theories are proposed to explain debt/equity policies, which is often quantified as debt-to-equity (D/E) ratio. A high D/E ratio usually implies an aggressive financing policy. In a frictionless market, capital structure plays no role in firm-value-creation decisions and thus would not matter according to the theory proposed by Modigliani and Miller (Modigliani and Miller, 1958). However, the frictionless market does not exist since none of these assumptions to generate a perfect capital market are easy to meet in the real world. To explain industry practices with market frictions are incorporated, many theories have been proposed. So far, the following are accepted as plausible leverage explanations: trade-off theory, pecking order theory, and market timing theory.

These different theories explain leverage practices from different perspectives. Trade-off theory claims that an optimal financial leverage is the result of the trade-off between costs and benefits of financing choices. The costs here include fund-raising costs, costs brought by potential financial distress and agency costs that arise from conflicts of interest among stakeholders. The agency costs constitute the external costs brought by debt because of the interest conflicts between shareholders and debt holders. Agency concerns are often lumped into the trade-off framework. Pecking-order theory (Myers and Majluf, 1984) believes that firms tend to prioritize the "least effort" funding source when facing information asymmetry. In this way, a financing hierarchy of retained earnings, debt, and equity would be chosen, i.e., firms would prefer internal financing if available and then choose debt over equity once retained earnings are exhausted. Thus the ways in which funds are raised could act as signals that indicate a company's need for external finance. The market timing theory focuses on the timing of stock issues and repurchases. According to this theory, new shares are issued when the management believe that company stocks are overvalued, but they are repurchased when believed to be undervalued. These theories are reviewed below.

1. Modigliani-Miller (MM) Capital Structure Propositions

The first proposition of Modigliani-Miller is the irrelevance theorem proposed in 1958 (Modigliani and Miller, 1958), which laid the foundation of empirical capital structure research in modern finance. The theorem was developed based on perfect capital market assumptions and claims that debt ratio does not matter in both firm value and financing cost. However, considering the tax benefits of debt interests, we would reach a conclusion that firm value would be maximized if a firm is fully financed by debt. However, we can hardly find a firm that only uses debt to support its operation and growth in reality and thus MM theory fails to explain realistic industry practices. The reason for this failure is that the assumptions on which the theory is developed are too strict and idealized. However, the MM theorem still bears value since it depicts the circumstance where different financing means do not matter, which further sheds light on the reasoning that imperfections in market conditions render financing do matter in reality.

2. Trade-Off Theory

The trade-off theory claims that capital structure is the result of the trade-off between the benefits and costs brought by debt, equity and other financing sources. The benefits of debt mainly include tax benefits, i.e., tax shield of debt, and the monitoring role of regular periodic interest payments in mitigating agency concerns brought by free cash flow (Jensen, 1986a; Jensen and Meckling, 1976). The costs of debt mainly include fund-raising costs, dead-weight costs of bankruptcy, and external costs arising from exacerbated shareholder-debtholder conflicts (Stulz, 1990). The optimal capital structure is determined as the level of debt that maximizes firm value by making a trade-off between two firm states: solvency and insolvency (Kraus and Litzenberger, 1973). Myers and Majluf (Myers and Majluf, 1984) pointed out that the trade-off theory provides a target for a firm to move towards (Flannery and Rangan, 2006), even though the existence of that target is difficult to prove and its quantity is not easy to measure.

There are two main strands that try to elaborate this theory: static trade-off theory and dynamic trade-off theory. The static trade-off theory incorporates positive personal taxes on both equity and debt income, positive non-debt tax shield, bankruptcy cost, and agency costs (Bradley et al., 1984; Titman and Roberto, 1988; Rajan and Zingales, 1995). The dynamic model proposes that the optimal capital structure is determined by balancing between tax shield and bankruptcy distress cost (Kane et al., 1984; Brennan and Schwartz, 1984; Hackbarth et al., 2017; Hackbarth and Mauer, 2012; Choi, 2013; Fisher et al., 1989; Leary and Roberts, 2005). However, inconsistency between industry practices and theory prediction are also founded in some studies(Graham, 2000; Lemmon et al., 2008; Hovakimian et al., 2012; Korajczyk and Levy, 2003).

Among all the costs considered, agency cost has attracted a special attention from researchers. there is a substream theory developed out of trade-off theory, which is called agency cost theory. Agency cost theory believes leverage is the result of agency problems caused by interest conflicts among stakeholders, including shareholders and debt holders. One of these conflicts is the conflict between shareholders and debt holders that arises from the concern that shareholders may maximize self-interests instead of acting for debt holders because of the different loss/profit distribution mechanisms between the two parties. The theory can explain “asset substitution effect” (Jensen and Meckling, 1976) and “debt overhang” phenomenon (Myers, 1977), etc.. Finally, some studies reveal that debt could be used as a disciplining tool for managers (Jensen, 1986b). It is found that firms would reduce their leverage when more financial information is required to release after the Sarbanes-Oxley Act (2004), and this would mitigate agency problems to some degree (Bertus et al., 2008). However, one criticism of this theory is that the overall effect of agency problems on the optimal debt level is difficult and complicated to quantify (Parrino and Weisbach, 1999).

3. Pecking Order Theory

The pecking order theory proposes a financing hierarchy because of adverse selection brought by information asymmetry. This theory (Myers and Majluf, 1984) argues that firms prioritize retained earnings when funds are needed for operation activities. If the retained earnings are not adequate to fulfil the operation needs, debt financing would be preferred over equity as the second choice. The information asymmetry here refers to the fact that firm owners know the operational conditions and true values of their firms best, and they are not willing to give out future profits to others. The logic is applicable in explaining stock issuance/repurchase behaviors in open capital markets (Anderson and Carverhill, 2012; Boot and Thakor, 2011; Antweiler and Frank, 2005; Dybvig and Zender, 1991a; Brennan and Kraus, 1987; Bharath et al., 2008). However, as sound and reasonable as this theory is, it has not proved applicable to all situations. Instead of following the pecking order hierarchy, other factors such as the capital structures of peers (Leary and Roberts, 2014), firm size (Frank and Goyal, 2003; Fama and French, 2005), and the term structure of debts (Dangl and Zechner, 2016) have been proved as dominant in determining capital structures in certain situations. Moreover, when comparing the explanatory power of trade-off theory and pecking-order theory in practice, researchers hold different opinions as to which one is better than the other (Fama and French, 2005; Shyam-Sunder and Myers, 1999; Fama and French, 2002; Harris and Raviv, 1991).

4. Market timing Theory

Market timing theory is supported in many surveys among managers (Graham and Harvey, 2001). The basic idea of this theory is that managers choose equity market or debt market based on how favorable these markets are at the time when financing is needed. If neither market looks favorable, issuance may be deferred or cancelled. Otherwise, if the current conditions look unusually favorable, funds may be raised no matter whether there is a need for new funds or not. Studies so far have revealed many firm practices that are consistent with market timing behavior (Lucas and McDonald, 1990; Baker and Wurgler, 2002).

These theories introduced above differ in their relative explanatory power in firms' financing behaviors across different industries over different time periods. Although they seem like substitutes to each other, many researchers attest to the fact that these theories are complementary to each other (Flannery et al., 2012; Cole, 2013). These theories are the foundations on which we develop our arguments in our dissertation. In order to ensure that the proposed new factor contributes extra prediction power on leverage prediction, factors that have been proven important in these theories will be introduced as control variables in our research.

3.2 Zero/low-leverage Puzzle

The zero-leverage phenomenon exists in publicly listed firms from United States (Strebulaev and Yang, 2013) and from the U.K. (Dang, 2011). Zero/low-leverage firms are those firms whose leverage is zero or lower than expected. Strebulaev and Yang, 2013 point out that about a tenth of large non-financial firms listed in the U.S. are debt-free, 30% of which keep this zero-leverage policy for at least five consecutive years over 1962-2009 period. The extreme debt conservatism has become more prevalent in recent years. For instance, nearly 20% of firms utilized zero-leverage policy over the period from 2000 to 2007 (Strebulaev and Yang, 2013). In the U.K. 12% of public non-financial firms were debt-free between 1980 and 2007 (Dang, 2011).

Many firm characteristics, especially cash reserve, have proven to play important roles on zero-leverage decisions. When firms are profitable with lots of cash at hand, they prefer to use internal retained earnings to support the cash needs, thereby resulting in a low leverage (Minton and Wruck, 2001). This logic is consistent with the phenomenon that conservative debt policies are usually observed in large and liquid firms with high profitability. Furthermore, bigger firms tend to use less debt as the financing method for new fund when debt is already in current capital structure (Ramalho et al., 2018). Researchers also point out that dividend-paying zero-leverage firms is likely to pay higher taxes, issue less equity (Strebulaev and Yang, 2013), and leave a considerable tax benefit of debt on the table. Following this logic, for zero-leverage firms that do not pay dividends, they would need less external financing, meaning that firms may keep staying in the zero-leverage state in the near future.

Other firm characteristics such as product market conditions, potential growth opportunities, future expenditure plan and credit-constrained situations also contribute to conservative debt usage (Graham, 2000). Product market characteristics, industry concentration and product uniqueness motivate firm to use debt conservatively (Titman, 1984; Titman and Roberto, 1988). Additionally, firms that have growth options or future acquisitions tend to use debt conservatively (Graham, 2000). In addition to those factors, many firms have been found to adopt zero-leverage policy due to credit-constrained situations, but some less constrained firms that pay dividends may also tend to avoid debt to mitigate investment distortions (Dang, 2011).

Besides firm characteristics, personalities and experiences of managers are also important in conservative leverage decisions. Managers who possess substantial stocks is less likely to do debt financing, hoping to keep their wealth safe (Strebulaev and Yang, 2013; Agrawal and Nagarajan, 1990; Berger et al., 2012). Furthermore, overconfident CEOs tend to adopt low leverage (Malmendier et al., 2011). What's more, family-controlled firms tend to be conserve and adopt low-leverage policies since family members tend to secure the well-being of the family members (Strebulaev and Yang, 2013; Becker, 1981; Bertrand and Schoar, 2006). Finally, the proportion of voting rights that managers possess and the family relationships among senior managers play roles in zero-leverage decisions (Agrawal and Nagarajan, 1990)

Though these factors play important roles on leverage decisions, they are shown to be inefficient in explaining leverage changes in many studies (Graham et al., 2015). Dramatic changes in leverage cannot be explained by such factors that are related with firm characteristics (Rajan and Zingales, 1995; Frank and Goyal, 2009), but can be explained by macroeconomic factors that altered firms' preferences in financing methods. For example, government leverage, the ratio of Federal debt held by the public to gross domestic product (GDP), is found to have the capability to explain dramatic leverage changes (Graham et al., 2015). The reason is that the supply of government debt may be a good proxy of unmeasured investment opportunities, which is an important factors to consider in financing. When the economic condition is not good for investments, firms do not have strong motivations to raise external money. At the same time, we would also observe an increasing treasury supply. Because debt is the primary source of external capital (Gorton and Winton, 2002), leverage falls as a result, Thus this macroeconomic factor could have a strong impact on leverage.

3.3 High-leverage Phenomenon

High-leverage phenomenon is commonly observed in financial companies, especially in banks. Many factors have been studied to pay roles on high leverage decisions.

Agency problems are important factors in high-leverage decisions. One agency problem concern is moral hazard. Moral hazard exacerbates the friction brought by additional lending and makes shareholders reluctant to repurchase debt but inclined to increase debt instead (Bizer and DeMarzo, 1992). Especially for firms whose creditor governance is weak, say banks, excessive leverage is more common (Becht et al., 2011). Another agency problem concern is leverage ratchet effect. The leverage ratchet effect is a phenomenon where the leverage of a firm has a tendency to climb up even though debt reduction helps enhance firm value. This effect exists because firms are unable to commit to future funding which makes shareholders pervasively resist leverage reductions no matter how much value such reductions may add to the firm. Instead, shareholders would even increase leverage even if the new debt is junior and would reduce firm value. With these factors considered, shareholders would always be reluctant to repurchase debt for self-interest maximization (Admati et al., 2018).

Another factor that has been shown to be important is the companies' role in the money supply chain. One typical example is the fact that banks hold higher leverage than firms in other industries in general. Banks, as money suppliers, adopt high leverages because of the advantage in reducing bank asset volatility and the risk-taking incentive from skewed return distributions (Gornall and Strebulaev, 2018). Additionally, the ability to offer lowest interest rates for highly levered financial intermediaries enables banks to take on high leverage safely. Thus, high bank leverage is the result of the joint capital structure decisions of banks and their borrowers. Even though we exclude financial companies in our study in later sections, this paper gives us some possible factors that we need to consider regarding the high leverage phenomenon.

Since most of the discovered factors can explain either high-leverage phenomenon only or low-leverage phenomenon only for firms in real industry but not both at the same time, thus some new factors need to be brought in to understand the possible mechanisms behind the coexistence of the low/zero-leverage and high-leverage phenomena. Motivated by the study results in (Dang, 2011), we frame our research within the context of financial constraints and agency problem, and combine these two with another factor, i.e., production flexibility, which we are going to review in a subsequent section.

3.4 Production Flexibility and Leverage

Capital structure decision interacts with product market behavior (Kovenock and Phillips, 1997) because the interactions among different stakeholders are affected by capital structure, which in turn affect capital structure decisions. For instance, some firm-specific investments would not be conducted if they are considered insecure under current capital structure. In addition, firms that produce unique products will be more likely to lose customers than those that produces generalized products when they seem likely to be under distress (Titman, 1984). Furthermore, financial policy would affect firms' incentive to maintain its reputation in producing products of high quality, which in turn would affect its value and leverage (Maksimovic and Titman, 1991). For instance, debts may discourage a firm to produce products of good quality under some conditions and this would decrease its value. However, high-quality products would be a better choice for such firms whose assets carry high salvage value in case of liquidation, which helps to maintain a high firm value.

There have been very few papers that explore the roles of firm operation characteristics in corporate financial conservatism in empirical studies so far. The interaction between firms' product and leverage explains part of the inter-industry variations for a particular capital structure (Harris and Raviv, 1991), even after controlling industrial organization considerations. Furthermore, corporate entities can also adjust business risk in addition to financial risk when facing shocks (Strebulaev and Yang, 2013). For example, operating and asset characteristics such as firm size and profitability have more power than financial risk in equity price risk (Bartram et al., 2013). Finally, firms with flexible output price are likely to have higher leverages than those that do not adjust prices very often, especially after the staggered implementation of bank deregulation (D'Acunto et al., 2017).

Following this logic, production flexibility, an operation characteristic, is supposed to affect financial decisions including leverage decisions. However, the definition of production flexibility varies in different financial and operation studies. Real-options literature provides one important definition of production flexibility, defining it as switching between inputs and outputs (Kulatilaka, 1988; Triantis and Hodder, 1990), varying production volumes(He and Pindyck, 1992), and initiating or terminating operations (Brennan and Schwartz, 1984). In finance literature, the definition of production flexibility also varies based on the data used. For instance, MacKay (MacKay, 2003) utilized plant-level data from the U.S. Census Bureau in 18 manufacturing industries and used elasticity of substitution to measure a firm's ability to adjust its factor intensity and product structure, along with the firm's ability to vary production levels defined as the elasticity of cash flow to output volume. However, when using proprietary data about utility companies, a different measurement of production flexibility that are based on asset-level characteristics is adopted (Reinartz and Schmid, 2016). Finally, in theoretical research, production flexibility is defined as the ratio of the cost of post-production activity to that of pre-production activities (Aivazian and Berkowitz, 1998) where cost of pre-production is the production cost before market price known while post-production activity is the production cost after market price is observed. Another set of production flexibility definitions comes from operation research. Jack and Raturi (Jack and Raturi, 2009) study production flexibility in terms of volume flexibility, which is defined as the product of return of sales and the log of the ratio of the variance in sales to that of inventory buffers and cost-of-goods. Intuitively, it measures the extent of sales fluctuations relative to cost fluctuations. For public companies, this measure does not require access to proprietary data. The higher is this ratio, the more volume flexibility exists. For our study, we do not have access to the proprietary data needed to compute the measures in the finance literature, so we use this measure from the operation research literature and follow Jack and Raturi's definition of volume flexibility (Jack and Raturi, 2009) in our study. More details will be covered in our empirical chapters.

Besides the varied definition of production flexibility, how production flexibility or related concept[4], affects leverage is still an open empirical question. On the one hand, production flexibility allows a firm to adjust its product structure, production level and factor intensity to accommodate different market situations, thus lowering default risk and increasing debt capacity (Mandelker and Rhee, 1984; Mauer and Triantis, 1994; Aivazian and Berkowitz, 1998; Kahl et al., 2014; Reinartz and Schmid, 2016). On the other hand, however, if lenders are concerned that shareholders would abuse these real options to engage in asset substitution or risk shifting, higher production flexibility would lead to lower leverage. The situation would get worse if risk shifting is intractable. To deal with such a situation, firms would rely more on short-term debt and bank loans (MacKay, 2003). In this dissertation, we explore how production flexibility affects leverage by focusing on volume flexibility, a factor that reflects a firm's ability to operate profitably at different output levels as it is an increasingly important source of competitiveness in both manufacturing and service firms (Sethi and Sethi, 1990).

We next comment in more detail on the three key papers mentioned above. In MacKay's paper (MacKay, 2003), financial leverage is found to be negatively related

[4] One example of "a related concept" is real flexibility, which is usually measured as the sensitivity of marginal production and investment decisions to variations in the economic environment, and we will refer all of them as "production flexibility" in our dissertation

with production flexibility. The production flexibility is measured using plant-level data from the U.S. Census Bureau in 18 manufacturing industries. Three dimensions of financial structure are explored in the paper: financial leverage, maturity structure of debt, and credit structure. The relationship between financial structure and production flexibility is explored by regressing financial-structure ratios on production flexibility, following the standard methods of production economics in the real-options literature. After controlling possible factors that could affect financial leverage, production flexibility is found to undermine financial policy, and the effect cannot be ignored. Additionally, firms with high production flexibility tend to use less leverage and prefer short-term debt and more monitored bank loans. These results indicate that lenders tend to tighten lending terms based on the anticipation of the ex-post opportunism of shareholders for firms with high production flexibility.

However, this conclusion has been challenged by other studies. Aivazian and Berkowitz (Aivazian and Berkowitz, 1998) set up a theoretical framework to explore how production decisions interact with financing outcome. The production decisions of the paper focus on asset specificity as well as flexibility of its production technology with asset specificity defined as the capability of firms to utilize an asset in multiple situations. As a theoretical paper, the basic assumptions made in the paper are as follows:

- Firms' production activities consist of ex-ante production and post production. Ex-ante production is the production conducted before firms fully know the actual market demand while post production is the production after the market production is fully known. Ex-ante production has a lower cost than post production. Production flexibility is reflected via the post production cost after controlling the ex-ante production cost. Higher post production costs imply lower levels of production flexibility. The production cost per unit and volume in the ex-ante production stage are c_y and y, respectively. After market price is observed, the production cost per unit and volume in the ex-post production stage are c_z and

z, respectively. With this being said, the total production volume Q is $Q = y + z$.

- The firm possesses some market power. Market price follows a stochastic process where price is negatively related with quantity. A firm faces a stochastic price p for units of output Q that it supplies to the market, where $\tilde{p} = p(Q) + \tilde{\theta}$ where $\tilde{\theta}$ is a random variable; and $\theta \in [\underline{\theta}, \overline{\theta}]$, with cumulative distribution function (i.e., CDF) being $\Theta(\theta)$, and probability distribution function (i.e., PDF) being $f(\theta)$. Moreover, $\frac{\partial p}{\partial \theta} > 0$ and $\frac{\partial p}{\partial Q} < 0$. Thus, the firm possesses some market power.

- Market participants are risk-neutral. The tax system is asymmetric such that profits are taxed but losses are not subsidized.

- Both equity and debt are issued to support ex-ante production activities. Furthermore, bankruptcy costs are less than the principal and interest of debt. Both debt and equity can be utilized to finance ex-ante production activities while ex-post capacity is financed using trade credit. With that being said,$c_y y \geq D$ is assumed.

- The firm aims to maximize the expected net value of the aggregate cash flows.

Working from these assumptions, the model reveals that production flexibility increases potential tax shields, thus lowering potential bankruptcy cost. In this case, production flexibility is positively related with leverage. The authors also found that operating leverage and financial leverage are complementary to each other when asset specificity is low. The complementarity would get stronger with increasing ex-post capacity adjustment costs. Also important in financial leverage decisions is tax rate. Increasing corporate tax rates would induce an increasing capacity investment and financial leverage under low asset specificity, especially when the ex-post capacity adjustment costs are higher.

This theoretical conclusion is also supported in empirical studies with a different definition of production flexibility (Reinartz and Schmid, 2016). A direct and forward-looking measurement of production flexibility as run-up time and ramp-up costs using proprietary utility company data confirms the positive relationship between production flexibility and financial structure. Production flexibility works on financial leverage via reductions in the cost of financial distress and increased present value brought by tax shields, even though which one is more important relies on firms' profitability.

The key difference in the arguments of the above-mentioned two empirical papers (MacKay, 2003; Aivazian and Berkowitz, 1998), which have opposite conclusions, is whether lenders anticipate shareholders to abuse high production flexibility for the ex-post opportunism. We will consider this in both our theoretical research and empirical research. Furthermore, besides the agency problem consideration, we also incorporate the issue of limited market access that firms face in our research as whether firms are financially constrained may affect how firm characteristics affect firm leverage (Ellul and Pagano, 2019). Previous research pointed out that corporate leverage responds differently to employee rights depending on whether it is driven by credit constraints. For financially unconstrained firms, leverage increases less to increasing property values or profitability when employees possess strong rights in restructuring but weak seniority in liquidation. However, leverage responds to leverage in the opposite direction in financially constrained firms. Our research also illustrates different mechanisms of the role of the two aforementioned factors on leverage when firms have different access to financial markets. When there is no constraint to financial market, the two factors have no effect on leverage. However, when the access to financial market (Faulkender and Pertersen, 2006) is constrained, production flexibility helps lower leverage in general and also increase (decrease) the tendency of firms going to a low-leverage state (high-leverage). The existence of agency problem concerns would change the relationship between production flexibility and leverage since such concerns would decrease firm value and thus increase debt ratio.

Standing on these studies, we will explore how production flexibility interacts with leverage when agency problems and financial constraints are considered in the following chapters. The study consists of a theoretical study and an empirical exploration. In theoretical study, the framework is an extension of the 2-stage production model (Aivazian and Berkowitz, 1998) where we consider more factors (such as agency problem and financial constraints) under different objective functions (such as firm value maximization versus equity value maximization). In the empirical study, we utilize a panel data analysis method to further explore how production flexibility affects leverage as well as tail leverage phenomenon under varying conditions framed by financial constraint and agency problem conditions.

Chapter 4

THEORETICAL EXPLORATION

In this section, we will explore the role of production flexibility on leverage decisions in settings that include agency problems and financial constraints. Our model extends a theoretical paper that uses a 2-stage production model where a pure operation-related production flexibility measurement is adopted (Aivazian and Berkowitz, 1998). By borrowing this framework, we can go beyond the traditional considerations in the finance literature to include operational factors of firms' production processes to explore how production considerations interact with other issues to affect leverage. This model setup is especially appropriate for retail firms, but may fit less well with service firms since it is difficult to differentiate production stages for service firms. Nonetheless, we will show in a subsequent chapter that the model can also help to understand patterns in leverage for the population of firms overall.

As previously discussed in the literature review chapter, Aivazian and Berkowitz, 1998 (AB for short afterwards) study a case in which firms must decide on production before demand conditions are fully known. They can make cost adjustments to production once more information comes out. In this setting, prior leverage choices are conditioned on the extent of uncertainty and on the cost of adjustment (i.e., production flexibility). AB do not include financial constraints and agency problems in the model. A point of our research is to understand how leverage debt has changed over long spans of time. Production flexibility, financial constraints, and the severity of uncontrolled agency problem have all changed over the decades. We suggest that bringing these into the model can provide insights and a basis for empirical investigation. Furthermore, the motivation of decision makers is also important, which is also needed to be considered in the discussion.

Existing studies prove that these three factors as well as the management's goal to achieve are important in leverage decisions. Whether firms are financially constraint gives firms significantly different leverage (Faulkender and Pertersen, 2006). Furthermore, it also affect other firm characteristics play roles in leverage ratios (Ellul and Pagano, 2019). For production flexibility, even though there are different research findings in how production flexibility affects leverage, it is found to play an important role in leverage (Aivazian and Berkowitz, 1998; MacKay, 2003). Finally, agency problem concerns and whether managers aim to maximize equity value instead of firm value have been proven to affect leverage decisions (Jensen, 1986b; Myers, 1977; Bertus et al., 2008).

For our purpose, "Financial constraint" is defined as the situation where firms face a limit in financing before production activities and need to reserve funds for their multiple-stage production activities, as they are not free to raise funds when needed afterwards. In detail, the limit in financing refers to the fact that some firms are not able to use new equity to finance their production activities. The "agency problem" defined in our research is that there exist conflicts of interest between existing shareholders/managers and debtholders. One typical example is risk shifting where shareholders pursue risky investments to gain potential extra profits but shift downside risk to debtholders. The funds needed for investment comes from production gains, which is the channel how production flexibility is exploited. We will monetize the gains/losses without loss of generality by setting up a two-state return condition for the risky investment, which will be discussed in detail in later section. Another modification of the model is that we explore our topic under different objects as to equity value maximization or firm value maximization since company behavior may differ under different goals. The optimization model aims to find optimal debt raised before production activity as well as the production level in the ex-ante production activity and ex-post production activity. However, to analyze the relationship between production activity and leverage, we do not need to get analytical solutions but rather the sign of the derivative of leverage with regard to production flexibility proxy. The cases we are going to explore are summarized in the table below:

Table 4.1: Summary of model cases to explore

The following table illustrates all the model cases we are going to explore in our research. In general, we have two panels with Panel A showing model cases under equity value optimization while Panel B showing model cases under firm value optimization. For each panel, there are four different situations classified by two dimensions: whether a firm is financially constrained and whether it faces agency problems.

Panel A: Managers to maximize equity value of existing shareholders		
	Agency problem situation	
Financial constraint situation	No agency problem and no financial constraint	Agency problem and no financial constraint
	No agency problem and financial constraint	Agency problem and financial constraint
Panel B: Managers to maximize firm value		
	Agency problem situation	
Financial constraint situation	No agency problem and no financial constraint	Agency problem and no financial constraint
	No agency problem and financial constraint	Agency problem and financial constraint

Making these modifications is important for us to understand possible mechanisms for how production flexibility works upon leverage. All the three key factors, — production flexibility, agency problem concerns, financial constraints — have their own direct impacts on firm activities. A direct influence of production flexibility on firm is production cost, and high production flexibility indicates low ex-post production cost and implies low production cost in general. The agency problem concerns affect the willingness of debtholders to lend while Financial constraint directly determines the structure of new raised funds as well as financing cost. These three effects combined would influence leverage by the following channels: production cost, possibility of managers seeking risky investment, likelihood of bankruptcy, amount of new equity. Our findings also demonstrate that different optimization goals determine effects via which channels are important for firms to consider in their decisions, further resulting in different relationship between production flexibility and leverage.

In the rest of this section, we introduce the assumptions of our theoretical study, and then exhibit our model and derive results.

4.1 Notation

The table below lays out the key notations that will be used in our model. Optimized values solved from our model will be expressed by adding $*$. For example, optimized value of y is y^*.

Table 4.2: Summary of annotation used in model

The following table summarizes all the annotations used within our theoretical model.

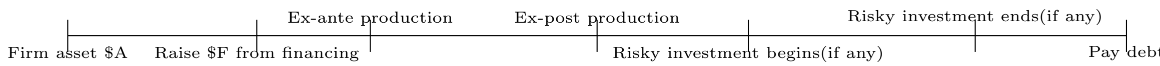
Notations	Meaning
θ	Product market price
A	Asset before production activity
B	Bankruptcy Cost
C	Total Cost
c_y	Ex-ante production cost per unit
c_z	EX-Post production cost per unit
D	Debt raised before production activities
F	Money raised before production activities
k_1	A constant that illustrates the degree of constraint
L	Money needed for risky investments
m	Return rate of risky investment when investment succeeds
n	To indicate the relative size of risky investment funds requirement with respect to debt D, i.e., $n = \frac{L}{D}$
p	Probability at which the risky investment fails
P_0	Profit threshold for triggering risky investments
Q	Total production volume
r	Debt interest rate
R	Total Revenue
$\mathrm{V}_E(y, z, D, \theta)$	Value of equity
$\mathrm{V}_E(y, z, D, \theta)$	Value of firm
y	Ex-ante production volume
z	Post production volume

4.2 Timeline and Assumptions

Timeline:

Firms need to raise $\$F$ from financing which could include debt and equity (if possible) to finance their production activities before any production activities begin. After financing is done, firms begin the production process, which consists of two phases: ex-ante production and ex-post production. Ex-ante production is conducted before firms gauge the actual demand since they do not know the market price. After observing the market price and demand, firms can decide to produce more if they have underestimated the demand. Therefore, the production after the realization of market demand is called ex-post production.

After the production activities are concluded, firms make decisions regarding whether to make risky investments. In our setup, risky investments can only be made if firms gain enough profits from production activities, and firms further gain profits or losses from the risky investment. After the investment stage, firms then need to pay back debts they borrow. This process of financing, production and investments is summarized in the general timeline below:



This timeline is operationalized in the model via the following assumptions.

Assumptions:

Assumption 1: Production Flexibility and Production Decision

Firms produce goods under uncertainty. Firms are involved in muti-stage production activities. We simplify the production activities to two stages: ex-ante production and ex-post production. The ex-ante production volume y is produced at the cost of c_y per unit while the ex-post production volume z is produced at the cost of c_z per unit. The post-production requires a speedier process and thus requires a higher production cost per unit, i.e., $c_z > c_y$. For firms having similar c_y, higher c_z indicates a lower volume flexibility. Based on these production assumptions, the total output of the firm Q satisfies $Q = y + z$. And the cost, C, of the two production activities is:

$$C = c_y y + c_z z.. \tag{4.1}$$

Firms need to make decisions on how to arrange production volume between ex-ante and ex-post activities, that is, y^* and z^* with the goal of equity/firm value maximization under different constraints. With the optimal value of y and z determined, the optimal total production volume would be $Q^* = y^* + z^*$.

Assumption 2: Market Demand and Market Price

The market demand can not be observed directly but can be inferred from the market price. In the ex-ante production process, market price θ is not fully known but has the following properties:

The range of θ is $[\underline{\theta}, \overline{\theta}]$, with cumulative distribution function being $F(\theta)$, and probability distribution function being $f(\theta)$. We assume θ follows a uniform distribution, i.e., $\theta \sim U[\underline{\theta}, \overline{\theta}]$, to not favor any price. Under this assumption, we have $f(\theta) = \frac{1}{\overline{\theta}-\underline{\theta}}$. The revenue R of the two production activities mentioned above is:

$$R(y, z, \theta) = \theta Q = \theta(y + z).. \tag{4.2}$$

The market price θ cannot be affected by a single firm's output, meaning that firms have no market power. After the price is observed, the optimal production Q^* is determined by the condition that marginal revenue meets marginal cost. The optimal values of y and z are obtained by solving our optimization problem backward sequentially. Firstly, we solve the optimization problem to obtain optimal z, i.e., z^*, given y and θ, then we substitute z^*, which is an expression of y and θ, back to the objective function and get the optimized value of y, i.e., y^*. We assume $\frac{\partial Q^*}{\partial \theta} > 0$ to illustrate a positive relationship between price and production from suppliers' perspective. We have:

$$Q^*(\underline{\theta}) \leq Q^* \leq Q^*(\overline{\theta}).. \tag{4.3}$$

In this case, the optimal post-production volume would be the net volume of optimal production net ex-ante production volume. Thus, given any y, when firms overproduce in the ex-ante production stage, firms should not produce any products in the ex-post production stage. Combining the two scenarios above, the optimum volume of the post production would be:

$$z^* = max(Q^* - y, 0).. \tag{4.4}$$

To simplify our calculation, we can transform the function of z^* above to the form below to simplify calculation without loss of generality since the transformation would not change function properties except for differentiability[5] :

$$z^* = \frac{Q^* - y}{2} + \frac{\sqrt{(Q^* - y)^2 + \gamma}}{2},. \tag{4.5}$$

where γ is infinitely small. we can see that:

1) When $Q^* \geq y$:

$$z^* = \tfrac{Q^*-y}{2} + \tfrac{\sqrt{(Q^*-y)^2+\gamma}}{2} \approx \tfrac{Q^*-y}{2} + \tfrac{Q^*-y}{2} = Q^* - y$$

.

2) When $Q^* \leq y$:

$$z^* = \tfrac{Q^*-y}{2} + \tfrac{\sqrt{(Q^*-y)^2+\gamma}}{2} \approx \tfrac{Q^*-y}{2} + [-\tfrac{Q^*-y}{2}] = 0 \ .$$

Thus, the function (4.5) aligns with function (4.4). The transformation would not change function properties except to transform the non-differentiable function to a differentiable one.

Assumption 3: Financing

Before the whole financing and production cycle, the firm has existing asset A, which is assumed to be equity for simplicity.

Firms need to raise money from debtholders or equity to support their production activities. For firms that face financial constraints, they are not able to raise money with equity but need to raise funds F from debt only to support pre-production

[5] There are multiple choices in perturbing the function form to gain differentiability and we pick the following one for its simplicity.

activity, that is, $F = D$. Any funds shortage in post-production activity can be supported by trade credits. Since the funds consumed in the ex-ante production is only a proportion of the funds newly raised, i.e.,

$$c_y y \leq D. \tag{4.6}$$

For firms that do not face such strict financial constraints mentioned above, meaning that firms do not have any financing problems in fund-raising to support ex-ante activities, the money F raised is a mix of debt D and equity, i.e., $F = k_1 D$, where $k_1 > 1$ since firms do not have financial constraints in this case and can raise more than D. Furthermore, money raised more than D is from external equity investors. Thus, pre-production activities can consume more money than the debt raised, i.e.,

$$c_y y \leq k_1 D. \tag{4.7}$$

Finally, the debt interest rate is r. Bankruptcy cost is B and it is assumed to satisfy: $B < D(1 + r)$. Such an assumption is made to guarantee the smooth pay of debtholders, meaning that debtholders can get fully paid to partially paid to zero in different scenarios. Otherwise, if $B > D(1 + r)$ is assumed, then debtholders would either get fully paid or zero, which does not really match the practice reality.

Assumption 4: Scenarios debtholders face[6]

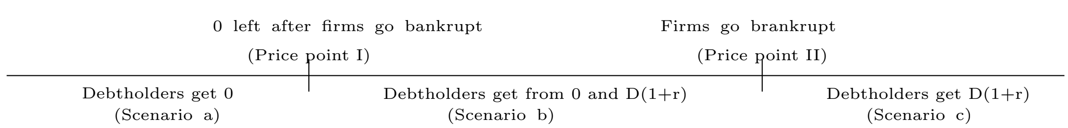
We will discuss the scenarios debtholders face under two different situations: "without agency problem" and "with agency problem", and discuss related thresholds for these scenarios one by one. To reiterate, the agency problem defined in our research is the conflict of interest between existing shareholders/managers and debtholders. One typical example is risk shifting where shareholders pursue risky investments to gain potential extra profits and shift downside risk to debtholders. Investment decisions

[6] Scenarios debtholders face and related thresholds are summarized in Appendix Table C.1.

need to be made right after production activity is done and gains are achieved but before debt is paid back to debtholders. Risk investment can only be possible when gains from production activity are bigger than L, the amount required for the investment. L is assumed to be greater than $D(1+r)$ and the relative size of the investment to debt is $n = \frac{L}{D}$. This assumption is made to illustrate the agency problem concern since debtholders would get fully paid if firms are able to make risky investments but do not do it. However, once firms make risky investments, there exists a possibility that debtholders cannot get fully paid. Every firm is assumed to face such a chance with n to be exogenous in our case.

For the risky investment, there is a probability of p where there is no gains and a probability of $1-p$ where shareholders/managers get $(m+1)L$. The riskier the investment is , the bigger p is. When the investment fails, it would be harder for debtholders to get their principal back than the scenarios where the investment does not happen. However, if the investment succeeds, the extra profit would go to shareholders/managers. Thus with this risky investment, shareholders/managers reap upside profits but shift downside failure to debtholders. m is assume to be big even when p is small to make this risky investment attractive and be a good motivation for shareholders in this risk-shifting setup.

When firms do not conduct the risky investment, debtholders in both two situations would face three common scenarios, i.e., debtholders get 0, debtholders get from 0 to $D(1+r)$, and debtholders get $D(1+r)$ as illustrated bellow:



In detail, the three scenarios are:

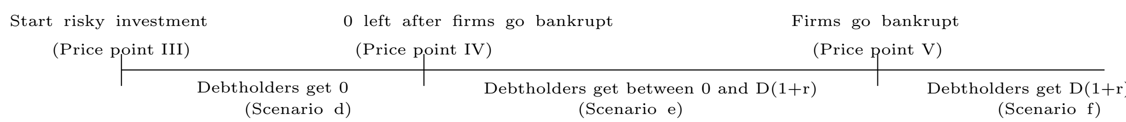
(Scenario a) Firms go bankrupt and debtholders get 0 after bankruptcy costs get paid.

(Scenario b) Firms go bankrupt and debtholders get from 0 to $D(1+r)$ after bankruptcy costs get paid.

(Scenario c) Firms do not go bankrupt and debtholders get $D(1+r)$.

These three scenarios can be separated by two price points: price point I and price point II, which we will discuss in more details later.

Furthermore, there are three additional scenarios debtholders face in a "with agency problem" situation where debtholders will not face in a "without agency problem" situation if the investment fails. The reason is that when the risky investment succeeds, debtholders will still get back their principal and interest since $(m+1)L \geq D(1+r)$ regardless of price situation after L has been achieved. However when the investment fails, in addition to the cases discussed in the "No agency problem" scenarios, debtholders would face the following three more cases as shown below:



Firms need to reach some profits threshold first to enable the risky investment, which is the price point III as shown in the figure above. Exceeding this price point, the details of the three extra scenarios are:

(Scenario d) The risky investment fails, firms go bankrupt and debtholders get 0 after bankruptcy costs get paid.

(Scenario e) The risky investment fails, firms go bankrupt and debtholders get between 0 to $D(1+r)$ after bankruptcy costs get paid.

(Scenario f) The risky investment succeeds, firms do not go bankrupt and debtholders get $D(1+r)$.

We will discuss the situation where there is no agency problem concern first in the following section.

1) Without agency problem:

The three scenarios listed above, i.e., scenario (a) to (c), can be separated by the two price points. One is the point to separate scenario (a) and scenario (b), the point at which firms go bankrupt with no money left to debt, and the other point is to separate scenario (b) and scenario (c), the point where firms are on the verge of bankruptcy with debtholders getting fully paid. These two thresholds can be described in detail below:

(Price point I) Firms go bankrupt, and the asset left is equal to the bankruptcy costs. In this case, the threshold of price θ_1 satisfies the following equation:

$$R(y, z, \theta_1) - c_y y - c_z z + F + A = B. \tag{4.8}$$

where F is money raised before ex-ante Production. In financial-constraint case, $F = D$, while in No financial-constraint case, $F = k_1 D$. This logic would be applied to later cases and would not be repeated again.

(Price point II) Firms are on the verge of going bankrupt. In this case, the threshold of price θ_2 satisfies the following equation:

$$R(y, z, \theta_2) - c_y y - c_z z + F + A = D(1 + r). \tag{4.9}$$

Based on the setup above, we can further calculate the expected value of existing shareholders and firm. We use $V_E(y, z, D, \theta)$ and $V_F(y, z, D, \theta)$ to represent equity value and firm value, respectively. Then we have the expected value of $V_E(y, z, D, \theta)$ as:

$$E(V_E(y, z, D, \theta)) = \frac{A}{A + F - D} \int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y, z,} \theta) - \mathrm{c_y y - c_z z - D(1 + r) + A + F}] f(\theta) \, \mathrm{d}\theta. \tag{4.10}$$

and the expected value of $V_F(y, z, D, \theta)$ as:

$$E(V_F(y, z, D, \theta)) = \int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y, z,} \theta) - \mathrm{c_y y - c_z z + F + A}] f(\theta) \, \mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B} f(\theta) \, \mathrm{d}\theta. \tag{4.11}$$

Since $f(\theta) = \frac{1}{\overline{\theta} - \underline{\theta}}$, we have:

$$E(V_E(y,z,D,\theta)) = \frac{A}{(A+F-D)(\overline{\theta}-\underline{\theta})}\int_{\theta_2}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_y y-c_z z-D(1+r)+A+F}]\,\mathrm{d}\theta. \tag{4.12}$$

$$E(V_F(y,z,D,\theta)) = \frac{1}{\overline{\theta}-\underline{\theta}}[\int_{\theta_1}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_y y-c_z z+F+A}]\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2}\mathrm{B}\,\mathrm{d}\theta]. \tag{4.13}$$

2) With agency problem:

The price point where the investment can be considered is where:

(Price point III) Firms gain enough from production activities to initiate risky investments:

$$R(y,z,\theta_3) - c_y y - c_z z + F + A = L. \tag{4.14}$$

Furthermore, the three cases mentioned above can be separated by the following two thresholds. One is the point at which firms go bankrupt with no money left to debt after risky investment, and the other point is where firms are on the verge of bankruptcy with debtholders getting fully paid after risky investment. These two thresholds are described in detail below :

(Price point IV) Firms go bankrupt due to risky investment, and the asset equals the bankruptcy costs. In this case, the threshold of price θ_4 satisfies:

$$R(y,z,\theta_4) - c_y y - c_z z + F - L + A = B, . \tag{4.15}$$

(Price point V) After the risky investment, firms are on the verge of bankruptcy. Letting θ_5[7] be the threshold, we have:

$$R(y,z,\theta_5) - c_y y - c_z z + F - L + A = D(1+r). \tag{4.16}$$

[7] The price points have been summarized in Appendix Table C.1.

The expected equity value is:

$$
\begin{aligned}
E(V_E(y,z,D,\theta)) = \frac{A}{A+F-D}\Big[\int_{\theta_2}^{\theta_3} &[\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{D(1+r)} + \mathrm{A} + \mathrm{F}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
+p\int_{\theta_5}^{\overline{\theta}} &[\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{D(1+r)} + \mathrm{A} + \mathrm{F} - \mathrm{L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
+(1-p)\int_{\theta_3}^{\overline{\theta}} &[\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{D(1+r)} + \mathrm{A} + \mathrm{F} + \mathrm{mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta\Big],
\end{aligned}
\tag{4.17}
$$

and the expected firm value is:

$$
\begin{aligned}
E(V_F(y,z,D,\theta)) = &\int_{\theta_1}^{\theta_3} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{F} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
&+p\Big(\int_{\theta_4}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{F} + \mathrm{A} - \mathrm{L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_4}^{\theta_5} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta\Big) \\
&+(1-p)\int_{\theta_3}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{F} + \mathrm{A} + \mathrm{mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta.
\end{aligned}
\tag{4.18}
$$

Summary: Optimization model framework

With all the assumptions made above, we summarize all the optimization models we are going to explore in this section since different optimization goals, i.e., equity value maximization versus firm value maximization, could lead to different production activities and financing behaviors (Dybvig and Zender, 1991b) as different optimization models are embedded with different implicit assumptions. We examine how the change in the objective function from equity value maximization to firm value maximization affects the relationship between production flexibility and firm value in our research. Four model cases discussed before under different optimization goals are summarized below:

Table 4.3: A summary of optimization models under equity value/ firm value maximization

The following table summarizes optimization models in different situations. Panel A illustrates models that optimize equity value while Panel B shows models that optimize firm value. Thus, the main difference between models in Panel A and those in Panel B is that models in Panel A only consider the equity value while Panel B also integrates interests of debtholders and new equity (if possible) into the optimization model. For each panel, there are four different situations classified by two dimensions: whether a firm is financially constrained and whether it faces agency problems. Since the optimal choice of z depends on y, θ, D, etc., models are solved sequentially backwards and optimized value of z follows equation 4.4. This optimized value of z is then put back into the expected function to solve the function with respect to ex-ante choice variables. However, in our essay, since we care more about the sign of the derivative of leverage with respect to c_z instead of model solutions, we will not cover details of how to utilize model equations to get solutions.

Panel A: Firms to maximize existing equity value	
Firm faces:	Optimization model
Financial constraint but no agency problem	$E(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{Dr} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq D$
No financial constraint but no agency problem	$E(V_E(y,z,D,\theta)) = \dfrac{A}{A+(k_1-1)D}\displaystyle\int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{A} + (k_1-1-r)D]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq k_1 D$
Financial constraint and agency problem	$E(V_E(y,z,D,\theta)) = \displaystyle\int_{\theta_2}^{\theta_3} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{Dr} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + p\displaystyle\int_{\theta_5}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{Dr} + \mathrm{A} - \mathrm{L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + (1-p)\displaystyle\int_{\theta_3}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{Dr} + \mathrm{A} + \mathrm{mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq D$
No financial constraint but agency problem	$E(V_E(y,z,D,\theta)) = \dfrac{A}{A+(k_1-1)D}[\displaystyle\int_{\theta_2}^{\theta_3} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{A} + (k_1-1-r)D]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + p\displaystyle\int_{\theta_5}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{A} + (\mathrm{k_1} - 1 - \mathrm{r})\mathrm{D} - \mathrm{L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + (1-p)\displaystyle\int_{\theta_3}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{A} + (\mathrm{k_1} - 1 - \mathrm{r})\mathrm{D} + \mathrm{mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta]$ s.t.: $c_y y \leq k_1 D$
Panel B: Firms to maximize firm value	
Firm faces:	Optimization model
Financial constraint but no agency problem	$E(V_E(y,z,D,\theta)) = \displaystyle\int_{\theta_2}^{\theta_3} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} - \mathrm{Dr} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq D$
No financial constraint but no agency problem	$E(V_E(y,z,D,\theta)) = \displaystyle\int_{\theta_2}^{\theta_3} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{A} + (\mathrm{k_1} - 1 - \mathrm{r})\mathrm{D}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq k_1 D$

Financial constraint and agency problem	$E(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\theta_3} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ $- \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ $+p(\int_{\theta_4}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A - L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_4}^{\theta_5} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta)$ $+(1-p)\int_{\theta_3}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A + mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq D$
No financial constraint but agency problem	$E(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\theta_3} [\mathrm{R(y,z,\theta) - c_y y - c_z z + A + k_1 D}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ $- \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ $+p(\int_{\theta_4}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + A + k_1 D - L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ $- \int_{\theta_4}^{\theta_5} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta)$ $+(1-p)\int_{\theta_3}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + A + k_1 D + mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$ s.t.: $c_y y \leq k_1 D$

We will discuss model frameworks under equity value maximization first in this section.

4.3 Comparative Statics

1) Optimization Model with the Goal of Equity Value Maximization

In this section, we will work on these optimization models listed in Table 4.3 to understand the role of production flexibility on leverage. Since in our study, we want to explore whether production flexibility is negatively or positively related with leverage in different situations, we aims to get the sign of the derivative of leverage with respect to production flexibility. To achieve that, we do not need to solve those optimization models to get analytical forms of solutions, but we utilize envelope theorem to calculate the derivative of the optimized objective functions (i.e., $E(V_E(y,z,D,\theta))$ or $E(V_F(y,z,D,\theta))$ in our study) with regard to parameter value changes (i.e., c_z in our study), which is further utilized to calculate the derivative of leverage or its transformation with respect to c_z. Finally, we discuss the sign of derivatives of leverage or its transformation with respect to c_z to understand whether the leverage is positively or negatively related with production flexibility.

For model case 1 (financial constraint and no agency problem):

For firms that do not face agency problems, the objective function of managers aiming to maximize equity value is:

$$E(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\bar{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A}] \frac{1}{\bar{\theta} - \underline{\theta}} \, \mathrm{d}\theta. \tag{4.19}$$

To understand the role of production flexibility on leverage, we do not need to solve this optimization problem to get an analytical form of solutions, but first need to analyze the impact of c_z on the value of equity $E(V_E(y,z,D,\theta))$. It is necessary to utilize envelope theorem to explore how the value changes of parameters, c_z in our case, affect the optimized objective functions, i.e., $E(V_E(y,z,D,\theta))$ in our study. Envelope theorem works at the differentiability properties of an optimized objective functions with regard to parameter value changes. Envelope Theorem proposes that when an initial choice set are convex and topological and the objective function is differentiable with respect to the parameters, the derivative of the objective function with respect to parameters equals to partial derivative of the value function to the parameters at optimal level (Takayama, 1985). In other words, changes in the optimizer of the objective function due to parameter value changes do not contribute to the changes of the objective function. Understanding that envelope theorem is an important tool for comparative statistics of optimization models in order to analyze the role of the parameters on the objective function (in our case, the role of c_z on $E(V_E(y,z,D,\theta))$), we apply the theorem by taking the first-order derivative of the objective function shown in equation (4.19) with respect to c_z:[8]

$$E^{'}(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - [R(y,z,\theta_2) - c_y y - c_z z - Dr + A]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}. \quad (4.20)$$

From equation (4.9), we get: $R(y,z,\theta_2) - c_y y - c_z z + F + A = D(1+r)$ where $F = D$, thus we have: $R(y,z,\theta_2) - c_y y - c_z z - Dr + A = 0$, which gives us:

$$E^{'}(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \quad (4.21)$$

Since $z \geq 0$, we would have the following:

$$E^{'}(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \leq 0. \quad (4.22)$$

Hence, the expected equity value is negatively related with c_z, i.e., equity value is positively related with production flexibility. The effect of production flexibility on equity value is via its effect on the production cost. Low production flexibility means high ex-post production cost, which further reduces equity value.

Based on this, we can further reason about the first-order derivative of debt ratio, i.e., $\frac{D}{E(V_E(y,z,D,\theta))+D}$, with respect to c_z. To facilitate the reasoning, we focus on the first-order derivative of $\frac{D}{E(V_E(y,z,D,\theta))}$ with respect to c_z since $\frac{D}{E(V_E(y,z,D,\theta))+D}$ are positively related with $\frac{D}{E(V_E(y,z,D,\theta))}$ and should generate the same signs regarding the derivatives.[9]

---

[8] In this section we use prime notation for first-order derivative with regard to c_z. For example, the first-order derivative of $E(V_E(y,z,D,\theta))$ with respect to c_z is $E^{'}(V_E(y,z,D,\theta))$. For second-order derivative, we use subscript notation instead to give clear expressions. For example, the second-order derivative of $E(V_E(y,z,D,\theta))$ with respect to c_z is $E_{c_z c_z}(V_E(y,z,D,\theta))$.

[9] To better understand the positive relationship between these two, we can do a simple transformation of $\frac{D}{E(V_E(y,z,D,\theta))+D}$ by dividing all terms with $E(V_E(y,z,D,\theta))$ and will get $\frac{\frac{D}{E(V_E(y,z,D,\theta))}}{1+\frac{D}{E(V_E(y,z,D,\theta))}}$, which can be further organized as $1-\frac{1}{1+\frac{D}{E(V_E(y,z,D,\theta))}}$. We can see that bigger value of $\frac{D}{E(V_E(y,z,D,\theta))}$ will generate a smaller value of $\frac{1}{1+\frac{D}{E(V_E(y,z,D,\theta))}}$ and a bigger value of $1-\frac{1}{1+\frac{D}{E(V_E(y,z,D,\theta))}}$, meaning bigger value of $\frac{D}{E(V_E(y,z,D,\theta))+D}$. Thus, $\frac{D}{E(V_E(y,z,D,\theta))}$ is positively related with $\frac{D}{E(V_E(y,z,D,\theta))+D}$.

Thus, we have the following:

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2}. \tag{4.23}$$

To get the sign of this expression, we need to know $D^{'}$. Since $D^{'} = \frac{E_{c_z c_z}(V_E(y,z,D,\theta))}{E_{c_z D}(V_E(y,z,D,\theta))}$, and we have:

$$\begin{aligned} E_{c_z c_z}(V_E(y,z,D,\theta)) &= \tfrac{\partial E^{'}(V_E(y,z,D,\theta))}{\partial c_z} = \tfrac{z^2}{y+z}\tfrac{1}{\overline{\theta}-\underline{\theta}} \geq 0, \\ E_{c_z D}(V_E(y,z,D,\theta)) &= \tfrac{E^{'}(V_E(y,z,D,\theta))}{\partial D} = \tfrac{zr}{y+z}\tfrac{1}{\overline{\theta}-\underline{\theta}} \geq 0\ . \end{aligned}$$

Thus, we have $D^{'} = \frac{z}{r} \geq 0$. This indicates that when firms aim to maximize equity value, firms need to borrow less from debtholders when c_z is smaller, i.e., when production flexibility is higher. Combined with previous results that $E^{'}(V_E(y,z,D,\theta)) \leq 0$, we have equation $D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) \geq 0$ which tells us that (4.23) is positive, i.e,

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} \geq 0\ .$$

Hence, c_z is positively related with leverage, which tells us that increased c_z, which indicates decreased production flexibility, results in higher leverage. This result makes sense since low production flexibility reduces equity value and increases the money needed to borrow from debtholders.

For model case 2 (no financial constraint and no agency problem):

For firms that do not face agency problems as well as financial constraints, the objective function of managers aiming to maximize equity value is:

$$\begin{aligned} &E(V_E(y,z,D,\theta)) \\ &\quad = \int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y - c_z z - Dr + A + (k_1 - 1)D}]\frac{\mathrm{A}}{\mathrm{A + (k_1-1)D}}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \end{aligned}$$

Similarly to what we do before in model case 1, we analyze the impact of c_z on the value of equity by applying envelope theorem and taking the first-order derivative of the objective function with respect to c_z and have:

$$
\begin{aligned}
E^{'}(V_E(y,z,D,\theta)) = \\
&\frac{-A}{A+(k_1-1)D}[R(y,z,\theta_2)-c_y y-c_z z-Dr+A+(k_1-1)D]\frac{\theta_2^{'}}{\overline{\theta}-\underline{\theta}} \\
&+\frac{A}{A+(k_1-1)D}\int_{\theta_2}^{\overline{\theta}}\frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta.
\end{aligned}
$$

From equation (4.9), we get: $R(y,z,\theta_2) - c_y y - c_z z + F + A = D(1+r)$ where $F = k_1 D$; thus, we have: $R(y,z,\theta_2) - c_y y - c_z z - Dr + A + D(k_1-1) = 0$, which gives us:

$$
E^{'}(V_E(y,z,D,\theta)) = \frac{A}{A+(k_1-1)D}\int_{\theta_2}^{\overline{\theta}}\frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta.
$$

The right hand side of the equation is the effect of production flexibility on equity value via production cost, analogous to what we have before. Similarly, we further focus on the first-order derivative of $\frac{D}{E(V_E(y,z,D,\theta))}$ with respect to c_z and we have the following:

$$
\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2}. \tag{4.24}
$$

To get the sign of this expression, we need to know $D^{'}$. Since $D^{'} = \frac{E_{c_z c_z}(V_E(y,z,D,\theta))}{E_{c_z D}(V_E(y,z,D,\theta))}$, and for each term, we have:

$$
E_{c_z c_z}(V_E(y,z,D,\theta)) = \frac{\partial E^{'}(V_E(y,z,D,\theta))}{\partial c_z} = \frac{z^2}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \geq 0,
$$

$$E_{c_zD}(V_E(y,z,D,\theta)) = \frac{\partial E^{'}(V_E(y,z,D,\theta))}{\partial D} = \frac{z(r-(k_1-1))}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}}+$$
$$\frac{-A(k_1-1)}{(y+z)[A+(k_1-1)D]^2}\int_{\theta_2}^{\overline{\theta}}\frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta.$$

For an extreme case where k_1 is big, we have $E_{c_zD}(V_E(y,z,D,\theta)) = \frac{z(r-(k_1-1))}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}}$. Thus, we have $D^{'} = \frac{z}{r-(k_1-1)} \leq 0$. This results shows that small c_z, i.e., high production flexibility, would lead to more debt D, which is consistent with the finding of AB (Aivazian and Berkowitz, 1998). Substituting all the results we have for D, $D^{'}$, $E(V_E(y,z,D,\theta))$, and $E^{'}(V_E(y,z,D,\theta))$, we would get:

$$D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) = -D[\frac{A}{A+(k_1-1)D}\int_{\theta_2}^{\overline{\theta}}\frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta]$$
$$+\frac{z}{r-(k_1-1)}\int_{\theta_2}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_yy-c_zz-Dr+A+(k_1-1)D}]\frac{\mathrm{A}}{\mathrm{A+(k_1-1)D}}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$$
$$=[\frac{z}{r-(k_1-1)}\int_{\theta_2}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_yy-c_zz-Dr+A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$$
$$(1-\frac{k_1-1}{(k_1-1)-r})D\int_{\theta_2}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta]\frac{A}{A+(k_1-1)D}.$$

For an extreme case where k_1 is big, we have $\frac{z}{r-(k_1-1)} \leq 0$, which would cause the first term in the square brackets to be negative. For the second term in the square brackets, $\frac{k_1-1}{(k_1-1)-r} \geq 1$, thus we have $(1-\frac{k_1-1}{(k_1-1)-r}) \leq 0$ and this further indicates that the second term is negative as well. Since both the two terms are negative, we would have the whole term in the square brackets to be negative. Since $\frac{A}{A+(k_1-1)D} \geq 0$, we would have the whole equation negative, i.e., $D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) \leq 0$, meaning that

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2} \leq 0.$$

Hence, c_z is negatively related with leverage, which tells us that increased c_z, which indicates decreased production flexibility, results in lower leverage. This makes sense since high production flexibility indicates low operation risk which gives firms more tolerance in financial risk due to high leverage.

For model case 3 (financial constraint and agency problem):

For firms that face agency problems concerns, the objective function is:

$$\begin{aligned} E(V_E(y,z,D,\theta)) = \int_{\theta_2}^{\theta_3} & [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\ +p\int_{\theta_5}^{\overline{\theta}} & [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A - L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\ +(1-p)\int_{\theta_3}^{\overline{\theta}} & [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A + mL}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \end{aligned} \tag{4.25}$$

We apply envelop theorem and take the first-order derivative of our objective function with respect to c_z to explore how c_z affects equity value and have:

$$\begin{aligned} E^{'}(V_E(y,z,D,\theta)) = -\int_{\theta_2}^{\theta_3}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [R(y,z,\theta_3) - c_y y - c_z z - Dr + A]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} \\ -[R(y,z,\theta_2) - c_y y - c_z z - Dr + A]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'} \\ -p\int_{\theta_5}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - p[R(y,z,\theta_5) - c_y y - c_z z - Dr + A - L]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_5^{'} \\ -(1-p)\int_{\theta_3}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - (1-p)[R(y,z,\theta_3) - c_y y - c_z z - Dr + A + mL]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}. \end{aligned} \tag{4.26}$$

From equation (4.9), (4.14) and (4.16), we have: $R(y,z,\theta_2) - c_y y - c_z z + D + A = D(1+r)$ and $R(y,z,\theta_3) - c_y y - c_z z + D + A = L$ and $R(y,z,\theta_5) - c_y y - c_z z + D + A - L = D(1+r)$. Thus, we have:

$$\begin{aligned} E^{'}(V_E(y,z,D,\theta)) = -\int_{\theta_2}^{\theta_3}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} - p\int_{\theta_5}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\ -(1-p)\int_{\theta_3}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - (1-p)[(m+1)L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}. \end{aligned} \tag{4.27}$$

When there is no agency problem concern by assuming $p = 0$, we have:

$$E^{'}(V_E(y,z,D,\theta)) = -\int_{\theta_2}^{\theta_3} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} - \int_{\theta_3}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta$$
$$-[(m+1)L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}, \quad (4.28)$$

i.e.,

$$E^{'}(V_E(y,z,D,\theta)) = -\int_{\theta_2}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - mL\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} \leq 0, \quad (4.29)$$

meaning that c_z is negatively related with equity value, which is is consistent with our results for model case 1. Otherwise, when considering extreme cases where there are severe concerns over agency problems, i.e., $p = 1$ and L is big, we would have:

$$E^{'}(V_E(y,z,D,\theta)) = -\int_{\theta_2}^{\theta_3} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} - \int_{\theta_5}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \quad (4.30)$$

In the equation above, the first term and the third term illustrates the effect of production flexibility on equity value via production cost, which is similar to the case we analyze before. Increased c_z means higher production cost and thus lower equity value. However, this effect is offset by the reduced probability in agency problem concerns illustrated in the second term. Similarly, we further focus on the first-order derivative of $\frac{D}{E(V_E(y,z,D,\theta))}$ with respect to c_z and we have the following:

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2}. \quad (4.31)$$

To get the sign of this expression, we need to know $D^{'}$. Since $D^{'} = \frac{E_{c_z c_z}(V_E(y,z,D,\theta))}{E_{c_z D}(V_E(y,z,D,\theta))}$, for each term:

$$E_{c_z c_z}(V_E(y,z,D,\theta)) = \frac{\partial E^{'}(V_E(y,z,D,\theta))}{\partial c_z} = \frac{z^2}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \geq 0,$$
$$E_{c_z D}(V_E(y,z,D,\theta)) = \frac{E^{'}(V_E(y,z,D,\theta))}{\partial D} = \frac{(r+n)z}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \geq 0\ .$$

Thus, we have $D^{'} = \frac{z}{(r+n)} \geq 0$. This indicates that when there is agency problem concern, debtholders would be less willing to lend money when c_z is small, i.e., when production flexibility is high. Substituting all the results we have for $D^{'}$, D, $E^{'}(V_E(y,z,\theta)$, and $E(V_E(y,z,D,\theta))$ renders:

$$
\begin{aligned}
D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) = \\
&\int_{\theta_5}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A - L}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta] \\
&+ \frac{z}{r+n}[\int_{\theta_2}^{\theta_3} [\mathrm{R(y,z,\theta) - c_y y - c_z z - Dr + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
&\quad - D[-\int_{\theta_2}^{\theta_3} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L - D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} - \int_{\theta_5}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta].
\end{aligned}
$$

Considering cases where L is large, meaning that n is big too, the dominant term in the equation above is $-DL\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}$. Since $\theta_3^{'} = z \geq 0$, thus $-DL\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} \leq 0$ and

$$D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) \leq 0,$$

which further renders:

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2} \leq 0.$$

Hence, c_z is negatively related with leverage, which tells us that increased c_z, which indicates decreased production flexibility, results in lower leverage. Thus, when there is agency problem concern, low production flexibility mitigates the agency problem concern that debtholders have for shareholders taking risky investment which deteriorates their interests.

For model case 4 (no financial constraint and agency problem):

For firms that face agency problem concerns but no financial constraints, the objective function is:

$$
\begin{aligned}
&E(V_E(y,z,D,\theta)) = \\
&\qquad \frac{A}{A+(k_1-1)D}\int_{\theta_2}^{\theta_3}[\mathrm{R(y,z,\theta)-c_y y-c_z z-Dr+A+(k_1-1)D}]\frac{\mathrm{d}\theta}{\overline{\theta}-\underline{\theta}} \\
&\quad + p\frac{A}{A+(k_1-1)D}\int_{\theta_5}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_y y-c_z z-Dr+A-L+(k_1-1)D}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
&+(1-p)\frac{A}{A+(k_1-1)D}\int_{\theta_3}^{\overline{\theta}}[\mathrm{R(y,z,\theta)-c_y y-c_z z-Dr+A+mL+(k_1-1)D}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta.
\end{aligned}
$$

Similarly to our previous analysis, we apply envelop theorem and take the first-order derivative of our objective function with respect to c_z to explore how c_z affects equity value:

$$
\begin{aligned}
&E^{'}(V_E(y,z,D,\theta)) = \\
&\qquad \frac{A}{A+(k_1-1)D}\{[R(y,z,\theta_3)-c_y y-c_z z-Dr+A+(k_1-1)D]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} \\
&\qquad -[R(y,z,\theta_2)-c_y y-c_z z-Dr+A+(k_1-1)D]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}-\int_{\theta_2}^{\theta_3}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
&\quad -p\int_{\theta_5}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta-p[R(y,z,\theta_5)-c_y y-c_z z-Dr+A-L+(k_1-1)D]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_5^{'} \\
&-(1-p)\int_{\theta_3}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta-(1-p)[R(y,z,\theta_3)-c_y y-c_z z-Dr+A+mL+(k_1-1)D]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}\}.
\end{aligned}
$$

From equation (4.9), (4.14) and (4.16), we have: $R(y,z,\theta_2)-c_y y-c_z z+k_1 D+A-D(1+r)=0$ and $R(y,z,\theta_3)-c_y y-c_z z+k_1 D+A=L$ and $R(y,z,\theta_5)-c_y y-c_z z+$

$k_1D + A - L - D(1+r) = 0$. Thus:

$$E^{'}(V_E(y,z,D,\theta)) = \frac{A}{A+(k_1-1)D}\{-\int_{\theta_2}^{\theta_3}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L-D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'} - p\int_{\theta_5}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - (1-p)\int_{\theta_3}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - (1-p)[(m+1)L-D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}\}.$$

Considering extreme cases where agency problem concern is high, i.e., $p = 1$, this would result in the following, with orders rearranged::

$$E^{'}(V_E(y,z,D,\theta)) = \frac{A}{A+(k_1-1)D}\{-\int_{\theta_2}^{\theta_3}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_5}^{\overline{\theta}}\frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + [L-D(1+r)]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_3^{'}\}. \tag{4.32}$$

In the equation above, the terms inside the brackets illustrate the effect of production flexibility on equity value via production cost, analogous to the case we analyzed before. Increased c_z means higher production cost and thus lower equity value. However, this effect is offset by the reduced probability in agency problem concerns illustrated in the third term. The negative effect via the production cost channel is further offset by the firms' capacity in borrowing equity as shown in the last two terms. Similarly, We further focus on the first-order derivative of $\frac{D}{E(V_E(y,z,D,\theta))}$ with respect to c_z and we have the following:

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2}.$$

To get the sign of this expression, we need to know $D^{'}$. Since $D^{'} = \frac{E_{c_zc_z}(V_E(y,z,D,\theta))}{E_{c_zD}(V_E(y,z,D,\theta))}$, for each term:

$$E_{c_z c_z}(V_E(y,z,D,\theta)) = \frac{\partial E^{'}(V_E(y,z,D,\theta))}{\partial c_z} = \frac{z^2}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \geq 0,$$

$$E_{c_z D}(V_E(y,z,D,\theta)) = \frac{E^{'}(V_E(y,z,D,\theta))}{\partial D} = \frac{z(r-(k_1-1))}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} + \frac{-A(k_1-1)}{(y+z)[A+(k_1-1)D]^2}\int_{\theta_2}^{\overline{\theta}}[\mathrm{L}-\mathrm{D}(1+\mathrm{r})]\frac{1}{\overline{\theta}-\underline{\theta}}\mathrm{d}\theta.$$

Considering extreme cases where k_1 and L are large, we would have $E_{c_z D}(V_E(y,z,D,\theta)) \leq 0$. Thus, we have $D^{'} \leq 0$. This indicates that when there is external equity funding, debtholders would be more willing to lend money when c_z is small, i.e., when production flexibility is high. Also when k_1 and L are large, we have (4.32) to be positive, i.e., $E^{'}(V_E(y,z,D,\theta)) \geq 0$. which results in $D^{'}E(V_E(y,z,D,\theta)) \leq 0$ and $-DE^{'}(V_E(y,z,D,\theta)) \leq 0$. Thus we have:

$$D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta)) \leq 0,$$

and:

$$\frac{\partial(\frac{D}{E(V_E(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_E(y,z,D,\theta)) - DE^{'}(V_E(y,z,D,\theta))}{(E(V_E(y,z,D,\theta)))^2} \leq 0.$$

Hence, c_z is negatively related with leverage, which tells us that increased c_z, which indicates decreased production flexibility, results in lower leverage. Therefore, when there is no financial constraint but agency problem, production flexibility is negatively related with both debt value and equity value. When there are severe concerns regarding agency problems, the effect of production flexibility on equity value is more than that on debt and thus renders a positive relationship between production flexibility and leverage.

Summarizing all the four cases we analyzed before, we see that production flexibility works on leverage via production cost and the probability to trigger firms conducting risky investment when firms aim to maximize the value of existing equity holders. We have the following conclusions summarized in the table below:

Table 4.4: Relationships between production flexibility and leverage under equity value maximization

The following table illustrates the relationship between production flexibility and leverage in different situations with the goal of maximizing equity value. There are four different situations classified by two dimensions: whether a firm is financially constrained and whether it faces agency problems.

		Agency problem condition	
		No agency problem	Agency problem
Financial constraint condition	No financial constraint	Positive	Positive
	Financial constraint	Negative	Positive

When there is no financial constraint, debt raised and leverage are positively related with production flexibility regardless of the existence of agency problem concern. When there is no agency problem concern, the financial intuition behind the positiveness between debt and production flexibility is that raising funds from external shareholders sends a good signal to debtholders that production activities are profitable since new equity would dilute the interests of existing shareholders and managers would not resort to equity if projects are not profitable. Furthermore, high production flexibility implies low production cost and low operation risk, which enables firm to take more financial risk and thus illustrates a positive relationship between production flexibility and leverage.

The positive relationship stays when there exists agency problem concerns but the underlying mechanism differs. Raising external equity sends a positive signal to debtholders regarding promising projects, but the positiveness is mitigated by the concern that the interests of debtholders may be sacrificed. Furthermore, high production flexibility leads to high probability of firms conducting risky investment, which lowers down equity value in the end since it is assumed to be more likely to fail as it is highly-risky investment and thus leading to high leverage.

When there is financial constraint, meaning debt is the the only financing source, production flexibility is then negatively related to debt and leverage, provided there are no agency problem concern. The reason is that when financial flexibility is high, meaning ex-post production cost is low, then the funds needed would be less when compared to low financial flexibility scenarios. In this case, we would see that debt is negatively related with production flexibility. Furthermore, high flexibility helps increasing equity value, which further helps lower down leverage, resulting a negative relationship between production flexibility and leverage.

However, the existence of agency problem turns the relationship between production flexibility and leverage to be positive. Even though debtholders are still unwilling to lend in this case, the risk of firms conducting risky investments when production flexibility is high reduces equity value, especially in situations where there are severe concerns over agency problems, which thus leads to a positive relationship between production flexibility and leverage.

2) Objective function: firm value maximization

In this section, we consider a different objective function, which is firm value maximization, but we use the same solution method and reasoning logic. The reasoning and results follow:

For model case 1_A (financial constraint and no agency problem)

For firms that do not face agency problems, the objective function of managers aiming to maximize firm value is:

$$E(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{D} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \tag{4.33}$$

Similarly to previous analyses, we apply envelop theorem by taking the first-order derivative of the objective function with respect to c_z to analyze the impact of c_z on the firm value:

$$E^{'}(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - [R(y,z,\theta_1) - c_y y - c_z z + D + A - B]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_1^{'} - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}. \tag{4.34}$$

From equation (4.8), we get: $R(y^*, \theta_1) - c_y y - c_z z + D + A = B$, thus getting:

$$E^{'}(V_F(y, z, D, \theta)) = \int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta} - \underline{\theta}} \, \mathrm{d}\theta - B \frac{1}{\overline{\theta} - \underline{\theta}} \theta_2^{'}. \tag{4.35}$$

Since $z \geq 0$ and $\theta_1^{'} \geq 0$ and $\theta_2^{'} \geq 0$, we would have the following:

$$E^{'}(V_F(y, z, D, \theta)) \leq 0. \tag{4.36}$$

Hence, the expected firm value is negatively related with c_z, i.e., firm value is positively related with production flexibility. We further focus on the first-order derivative of debt ratio, $\frac{D}{E(V_F(y,z,D,\theta))}$, with respect to c_z, we would have the following:

$$\frac{\partial(\frac{D}{E(V_F(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'} E(V_F(y, z, D, \theta)) - D E^{'}(V_F(y, z, D, \theta))}{(E(V_F(y, z, D, \theta)))^2}.$$

To get the sign of this expression, we need to know $D^{'}$. Since $D^{'} = \frac{E_{c_z c_z}(V_F(y,z,D,\theta))}{E_{c_z D}(V_F(y,z,D,\theta))}$, for each term:

$$\begin{aligned} E_{c_z c_z}(V_F(y, z, D, \theta)) &= \tfrac{\partial E^{'}(V_F(y,z,D,\theta))}{\partial c_z} = \tfrac{z^2}{y+z} \tfrac{1}{\overline{\theta} - \underline{\theta}} \geq 0, \\ E_{c_z D}(V_F(y, z, D, \theta)) &= \tfrac{\partial E^{'}(V_F(y,z,D,\theta))}{\partial D} = \tfrac{-z}{y+z} \tfrac{1}{\overline{\theta} - \underline{\theta}} \leq 0 \ . \end{aligned}$$

Thus, we have $D^{'} = -z \leq 0$. This indicates that debtholders would be more willing to lend money when c_z is small, i.e., when production flexibility is high. Combined with previous results we get for $V_F^{'}$, V_F and D, we would have:

$$\begin{aligned}
D^{'}E(V_F(y,z,D,\theta)) - DE^{'}(V_F(y,z,D,\theta)) &= -D[\int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}] \\
-z[\int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta] \\
= -z\int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + D\int_{\theta_1}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
+z\int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + DB\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'} \\
= -z\int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + D\int_{\theta_2}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
+DB\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}. \qquad (4.37)
\end{aligned}$$

For the first two terms in the equation, we have:

$$\begin{aligned}
-z\int_{\theta_2}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + D\int_{\theta_2}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \\
\leq -z\int_{\theta_2}^{\overline{\theta}} \mathrm{D(1+r)}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta + D\int_{\theta_2}^{\overline{\theta}} \frac{\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta \leq 0.
\end{aligned}$$

However, the third term of equation (4.37) is $DB\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'} \geq 0$, thus the sign of the equation is determined by the relative value of bankruptcy cost B versus asset and profits. Since it is often the case that asset and profits are usually bigger than bankruptcy cost, then it is safe to conclude that the negative terms are dominant and turns the equation negative, resulting in a positive relationship between leverage and production flexibility.

The economic intuition behind this is that when bankruptcy cost is low, meaning that firms have a higher probability of being paid, partially or fully, when price turns out not as good as expected, firms would raise more debt when production flexibility is high since financial risk can be mitigated by low operation risk. Furthermore, even when firms go bankrupt, debtholders still can get partially paid since the cost of bankruptcy is low. This would result in a positive relationship between production flexibility and leverage.

For model case 2_A (no financial constraint and no agency problem)

For firms that do not face agency problems as well as financial constraints, the objective function of managers aiming to maximize firm value is:

$$E(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y,z,\theta) - c_y y - c_z z + k_1 D + A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta. \tag{4.38}$$

We apply envelop theorem by taking the first-order derivative of the objective function with respect to c_z to analyze the impact of c_z on the firm value and have:

$$E^{'}(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - [R(y,z,\theta_1) - c_y y - c_z z + k_1 D + A - B]\frac{1}{\overline{\theta}-\underline{\theta}}\theta_1^{'} - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}. \tag{4.39}$$

From equation (4.9), we get: $R(y,z,\theta_1) - c_y y - c_z z + k_1 D + A = B$, thus, we get:

$$E^{'}(V_F(y,z,D,\theta)) = \int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2^{'}. \tag{4.40}$$

In the equation above, the first term illustrates the effect of production flexibility on equity value via production cost, which is similar to the case we analyzed before. Increased c_z means higher production cost and thus lower equity value. The second term shows the effect via the bankruptcy channel since higher c_z increase firms' probability of going bankrupt. Since both the two terms in the equation above are negative, meaning that $E^{'}(V_F(y,\theta))$ is negative, this means that the expected firm value is negatively related with c_z, i.e., firm value is positively related with production flexibility. If we further focus on the first-order derivative of debt ratio, $\frac{D}{E(V_F(y,z,D,\theta))}$, with respect to c_z, and have the following:

$$\frac{\partial(\frac{D}{E(V_F(y,z,D,\theta))})}{\partial c_z} = \frac{D^{'}E(V_F(y,z,D,\theta)) - DE^{'}(V_F(y,z,D,\theta))}{(E(V_F(y,z,D,\theta)))^2}.$$

To get the sign of this expression, we need to know D'. Since $D' = \frac{E_{c_z c_z}(V_F(y,z,D,\theta))}{E_{c_z D}(V_F(y,z,D,\theta))}$, for each term:

$$E_{c_z c_z}(V_F(y,z,D,\theta)) = \frac{\partial E'(V_F(y,z,D,\theta))}{\partial c_z} = \frac{z^2}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \geq 0$$
$$E_{c_z D}(V_F(y,z,D,\theta)) = \frac{\partial E'(V_F(y,z,D,\theta))}{\partial D} = \frac{-k_1 z}{y+z}\frac{1}{\overline{\theta}-\underline{\theta}} \leq 0 \ .$$

Thus, we have $D' \leq 0$. This indicates that debtholders would be more willing to lend money when c_z is small, i.e., when production flexibility is high. Combined with previous results that $E'(V_F(y,z,D,\theta)) \geq 0$, we have equation:

$$D'E(V_F(y,z,D,\theta)) - DE'(V_F(y,z,D,\theta)) = -D[\int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2']$$
$$-\frac{z}{k_1}[\int_{\theta_1}^{\overline{\theta}} [\mathrm{R(y,z,}\theta) - \mathrm{c_y y} - \mathrm{c_z z} + \mathrm{k_1 D} + \mathrm{A}]\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - \int_{\theta_1}^{\theta_2} \mathrm{B}\frac{1}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta]. \quad (4.41)$$

When k_1 is big, the last term would be dominant and turns the entire equation to be positive, thus we would have:

$$\int_{\theta_1}^{\overline{\theta}} \frac{-\mathrm{z}}{\overline{\theta}-\underline{\theta}}\,\mathrm{d}\theta - B\frac{1}{\overline{\theta}-\underline{\theta}}\theta_2' \geq 0,$$

i.e.,

$$\frac{\partial\left(\frac{D}{E(V_F(y,z,D,\theta))}\right)}{\partial c_z} \geq 0 \ .$$

Hence, c_z is positively related with leverage, which tells us that increased c_z, which indicates decreased production flexibility, results in higher leverage. Thus we can see that when financial constraint is relaxed and firms have more access to financing market, the relationship between production flexibility and debt ratio would change to a negative relationship.

For model case 3_A (Financial constraint and agency problem):

Under the goal of firm value maximization, if the the net present value (NPV for short henceforth) of a risky investment is negative, i.e., $p*0+(1-p)(m+1)L-L \leq 0$, firms would not decide to invest such a project since it would hurt firm value. In this situation, the case would go back to model case 1_A where there is financial constraint but no agency problem.

However, when the risky investment has positive NPV, i.e., $(1-p)(m+1)-1 \geq 0$, then firms would decide to conduct the investment. However in this case, there is no risk-shifting behavior, the agency problem concern we mentioned does not exist, which renders it trivial for discussion.

For model case 4_A (no financial constraint and agency problem):

This model case faces a similar situation in the model case 3_A where there is financial constraint and agency problem. If the the net present value of a risky investment is negative, i.e., $p*0+(1-p)(m+1)L-L \leq 0$, firms would not decide to invest in such projects since it would hurt firm value. In this situation, the case would go back to model case 2_A where there is financial constraint but no agency problem.

However, when the risky investment has positive NPV, i.e., $(1-p)(m+1)-1 \geq 0$, firms would decide to conduct the investment. However in this case, there is no risk-shifting behavior and the agency problem concern we mentioned does not exist, which make it trivial for discussion.

Summarizing all four cases analyzed before, we see that production flexibility works on leverage via production cost, the probability to trigger firms conducting risky investment, and the probability of firms going bankrupt when firms aim to maximize firm value. We have the following conclusions summarized in the table below:

Table 4.5: Relationships between production flexibility and leverage under firm value maximization

The following table illustrates the relationship between leverage and production flexibility in different situations with the goal of maximizing firm value. There are four different situations classified by two dimensions: whether a firm is financially constrained and whether it faces agency problems.

Panel A: Relationship between production flexibility and leverage

		Agency problem	
		No agency problem	Agency problem
Financial constraint	No Financial constraint	Negative	Negative
	Financial constraint	Positive	Positive

When there is no financial constraint, leverage is negatively related with production flexibility regardless of agency problem. The logic is that when firms aim to maximize firm value, debtholders do not worry about firms exploiting their interests for the benefits of shareholders and thus are willing to lend money to support production activities. However, raising funds from equity helps increase whole equity value and thus decrease leverage.

The relationship between production flexibility and leverage changes when there exists financial constraint. When there is no agency problem concern, high production flexibility indicates low operation risk, thereby enabling firms to bear high financial risk and thus more debt. Since debt is the only funding source in this case, raised debt would increase firm leverage. The relationship would not change even if there exists agency problem concern under firm value maximization since conducting risky investment with negative NPV goes against the goal.

4.4 Model Conclusion

We explore how production flexibility affects leverage within a theoretical framework in this section. The relationship is affected by the following three factors: whether firms face agency problem concerns, whether there exist financial constraints, and the maximization goal of managers in conducting financing and operation activities. In general, production flexibility, which is measured as the ex-post cost, affects equity (firm) value via production cost and the possibility to trigger managers seeking risky investment when firms aim to maximize the benefits of existing shareholders. One more channel where production flexibility works on leverage is via probability of bankruptcy when firms aims to maximize firm value. Different working channels lead to different results regarding the relationship between production flexibility and leverage when firms have different maximization goals.

When firms aim to maximize equity value of existing shareholders, whether or not firms are capable of raising funds from external equity investors plays an important role in leverage. When there is no agency problem concern and there exists no financial constraint, high production flexibility indicates low operation risk and enables firms to take high financial risk, meaning that firms are able to take on high leverage. Thus, we observe a positive relationship between production flexibility and leverage when there is no financial constraint. When there exists financial constraint but still no agency problem concern, the relationship turns negative since high production flexibility implies low funds request which can be only raised from debtholders.

Whether there exists agency problem concern also influences how production flexibility works on leverage. Under similar settings in previous cases mentioned, but differing in the presence of concerns over agency problems, a different relationship between production flexibility and leverage would be found. For example, when there is no financial constraint but agency problem concerns exist, conducting risky investments would decrease equity value, which would further strengthen the positive relationship between production flexibility and leverage. For another situation where there is financial constraint as well as agency problem concern, the relationship turns from negative to positive due to decreasing equity value brought by risky investments compared to the situation where firms face financial constraints but no agency problem concerns.

Different maximization goals would affect the relationship between production flexibility and leverage as well. When firms aim to maximize firm value, the relationship between production flexibility and debt as well as the relationship between production flexibility and leverage changes. Within this goal, agency problem cases would become trivial since projects with negative NPV would not be conducted since it reduces firm value, while projects with positive NPV do not fit in the "agency problem" setup. Thus, the "agency problem" cases would degenerate to related "non agency problem" cases under this optimization goal. Following this logic, model cases with same financial constraint situations would generate same conclusions regardless of agency problem concerns.

Whether there exist financial constraints affects how production flexibility works on firm value. When there are financial constraints, an increase in debt would help increase firm value at same scale and thus leads to a higher leverage. However, when there is no financial constraint, financing sources from external equity investors would also help to improve firm value, meaning that firm value increases more than debt and results in a low leverage. This makes sense since when firms also consider new equity in the optimization goal, they would prefer new equity over debt until it is fully utilized since new equity does not require interest and thus is less costly than debt.

Chapter 5

EMPIRICAL RESEARCH

In this section, we will conduct empirical studies to provide evidences regarding the theoretical results developed in the theoretical section. In the previous theoretical section, we explore how production affects leverage under different contexts that are specified by three factors, i.e., financial constraints, agency problem concerns, and optimization goals. We find that all these factors play important roles on leverage decisions and the relationship between production flexibility and leverage differs under different specification. Since optimization goals cannot be observed, we would develop hypotheses based on factors that can be proxied by real data. Thus, the empirical work bears the task of testing the relationship between production flexibility and leverage under four different contexts defined by two dimensions: financial constraint versus no financial constraint, agency problem versus no agency problem. The four hypotheses as well as their alternative hypotheses are summarized as below:

H1A:Production flexibility is negatively related with leverage when there is no financial constraint and no agency problem concern.

H1B:Production flexibility is positively related with leverage when there is no financial constraint and no agency problem concern.

H2A:Production flexibility is negatively related with leverage when there is financial constraint but no agency problem concern.

H2B:Production flexibility is positively related with leverage when there is financial constraint but no agency problem concern.

H3A:Production flexibility is negatively related with leverage when there is no financial constraint but agency problem concern.

H3B:Production flexibility is positively related with leverage when there is no financial constraint but agency problem concern.

H4A:Production flexibility is negatively related with leverage when there are financial constraint and agency problem concern.

H4B:Production flexibility is positively related with leverage when there are financial constraint and agency problem concern.

Our initial expectation regarding the hypotheses is that empirical results may support our theoretical conclusions within the framework of equity value maximization goal. In other words, in real data we may observe a positive relationship between leverage and production flexibility when there is no financial or when there is financial constraint but no agency problem. However, when there is financial constraint but no agency problem, we observe a negative relationship between production flexibility and leverage. The reason why we favor equity value maximization framework is that agency problem concerns are not uncommon in practice and a large population of firms make decisions for equity holders' interests instead of whole stakeholders that includes debtholders, employees, etc.

Panel data analysis method is adopted to test these hypotheses. Model data covers non-financial firms that are publicly listed in the United States from 1962 to 2017 and fixed firm effects are considered in all model regressions. Our empirical tests start with a general situation where no specification is required to understand how leverage reacts to production flexibility over decades[10]. Then we further investigate the relationship in different industries, under different financial constraint and agency problem concerns situations to test the four hypotheses mentioned before. In this process, we adopt different measures of financial constraints and agency problem concerns to ensure the

[10] We understand that the empirical models and results indicate a strong correlation between production flexibility and leverage instead of causality between these two factors. However, the evidence we found here is sufficiently strong for possible causality. Furthermore, since endogeneity is a big concern in making causal statements here since both production flexibility and leverage are measured from the perspective of firms' activity solely and they could be related with same other endogenous factors which could be the true factor that drivers leverage changes instead, thus we alter the measurement of production flexibility a little bit to test the relationship in the robustness test.

robustness of our conclusions. Finally we explore the relationship by market segments to investigate the slight difference in leverage behavior observed between NYSE market and Nasdaq market. We further extend our analysis to leverage tails tests, hoping to understand the role of production flexibility on the fatter tails of leverage distribution over time.

Our study finds that high production flexibility goes with low leverage in general, which fits in the logic that less debt is needed under high production flexibility condition when firms that face financial constraint try to maximize equity value. This logic is further confirmed by the negative relationship between production flexibility and leverage on selected firms facing financial constraint situation only, and this favors hypothesis H2A. What's more, the result that existence of agency problem concerns mitigates the negative relationship under financial constraint situation supports hypothesis H3B. We also find that the relationship between production flexibility and leverage has been strengthened over time. Similar results are found for left tail and right tail exploration, which helps explain growing left tail and right tail observed over decades.

The empirical study is organized as follows. We introduce data in Section 1, followed by model and variables in Section 2. In Section 3 empirical exploration results are exhibited, which covers results on leverage and leverage tails. Section 4 shows robustness test results and Section 5 concludes.

5.1 Discussion about Identifying Assumptions

The fact that important factors could be missed in regression models brings up the identifiability issue in empirical studies. When there are no unmeasured variables, all causal effects can be identified and the effect can be estimated from model coefficients. However, this assumption cannot be satisfactorily guaranteed. Especially in a social science study, it is impossible for us to know all of the factors that contribute to an effect. We also rarely have a full picture of how these factors work together to make the effect occur. Thus, in most cases, what we get is correlations instead of causality

(Pearl, 2000). Facing such limitations in finding complete factors, causal relationships are not deterministic but are thought to increase the probability that increase the odds that an effect will happen (Eells, 1991). This also explains why certain causality found in certain studies could not be found in other studies where data differs in time, population et al..

One approach to conduct cause-probing studies is to design a counterfactual model (Pearl, 2019). "Counterfactual" means something that goes against the fact. We cannot observe counterfactual in the real world, but can create some reasonable approximations. For instance, we can observe what happens if a certain group of people received certain treatments. A counterfactual model tells us what would happen if the same group of people had not received these treatments in an imaginary alternative world. The difference between what happened when there are treatments and what would have happened without treatments is the effect we are interested in (Lewis, 1973). Based on this process, causal inference from a counterfactual model in experiments approach is more qualitative than quantitative. However, statisticians have developed a model called Rubin's Causal Model to formalize the analysis of causal inference in experiments, which have been widely used in many areas, including psychology, medicine, and sociology (Sekhon, 2007).

To explore possible causal relationship in an experiment, we need to test the following three criteria in experiments: (1) the cause factor should occur before the effect factor, not the other way around, (2) these two factors are correlated with each other and vary together to ensure there are effects between the two factors, and (3) there is not a third factor can be found to bring up the effect to rule out other possible explanations. These three features are mirrored in the following way: (1) the presumed cause is manipulated to see whether the expected outcome occur after it; (2) whether variations in the cause result in variations in the outcome effect; (3) multiple methods are utilized to help reduce the effects of other plausible causes. Furthermore, to identify the effects that are indirect in nature, some experiment designs allow imperfect manipulation of an intermediate factor while others require a perfect manipulation of these factors. Even though experimentation is considered as a powerful method to explore causal relationships in empirical research, it bears some criticism in that it is not able to provide a clear mechanism of how effects come about. This implies a serious limitation of utilizing experimentation in social science studies since social science studies strive to identify the underlying process that drive the effects.

Admitting the limitations experimentation bears, researchers turn to utilize the methodology of quasi-experiments for causality exploration (Liu et al., 2021). Quasi-experimental design is useful in such cases where true experiments are not feasible for either ethical or practical reasons. A quasi-experiment differs from a true experiment in that it is not subjected to random assignments (Harris et al., 2006). Instead, the subjects under study are grouped by some non-random criteria. The basic process of designing quasi-experiments is to list all possible alternative explanations, which are further assessed using logic or other approaches to decide which could be the most plausible reason that can be used to explain all observed effect. However, listing all possible reasons completely seems infeasible. Furthermore, some of the reasons could be context-based and thus approaches to remove them from consideration could vary across different studies. In this way, quasi-experimental design features could provide less compelling support than true experiments and counterfactual inferences.

Besides experimentation method, one alternative way to identify causal relationships is to make adequate identifying assumption. Identification tries to form a unique mapping between data features to quantities that researchers are interested in. In statistics or econometrics, Identification is to find a parameter that minimize an objective function that is related to certain economic problems. In social science studies, some researchers believe that making sufficient identification assumption is good enough to detect causalities. However, it is usually not easy to identify an economically interesting parameter (Kahn and Whited, 2018).

Due to that difficulty, one method synonymous with identification is to establish causality by finding exogenous factors in causality-probing studies. An exogenous factor is a factor whose value does not rely on other variables of the system we are interested in (Angrist and Pischke, 2008). Thus an exogenous factor cannot be determined or explained by any factors within the causal model or system, and it is usually taken as fixed and given. Hence exogenous factors fit in the concept of being a ceteris paribus comparison in causality exploration where all other factors within the system can be kept constant when exogenous factor is altered to study its role on the effect factor. In this way, by adopting exogenous factors, we are hoping to rule out the partial or full correlation between this factor and other explanatory factors in our model or the concern that these factors are affected by an underlying alternative factor that we are not aware of. Once we find a statistically significant parameter between this exogenous factor and the effect we are interested in, we are often confident that there is a causal link between what this exogenous factor represents and the effect. In our paper, we understand the empirical panel model we utilized is a good way to explore correlations between independent variables and dependent variable since there could be some other factors that are not included in the model, but can potentially influence both independent variables and dependent variables. We cannot conduct experiments or quasi-experiments to build a stronger support for the causality since it is not possible to do so.

5.2 Data

Our data comes from the the CRSP (Center for Research in Security Prices)/ Compustat Merged Database over the period 1962–2017. We start in the year 1962 because the Compustat data before 1962 are known to be biased towards large firms (Strebulaev and Yang, 2013). The data contains comprehensive information for U.S. companies listed in the New York Stock Exchange (NYSE for short), the American Stock Exchange, and Nasdaq. Accounting data are from Compustat annual files and stock-price-related data come from Center for Research in Security Prices (CRSP) monthly files. In the sample, we exclude financial companies (Standard Industrial Classification (SIC) codes from 6000 to 6999), non-U.S. companies (entries in Compustat with International Standards Organization country code of in corporation (FIC) not equal to

USA), and non-publicly-traded firms and subsidiaries (entries in Compustat with stock ownership variable, STKO, equal to one or two). Firms with leverage bigger than 1 (in either one of the three leverage measures) are also excluded. Finally, we have 261,400 firm-year observations left.

Table 5.1 shows the evolution of debt and leverage calculated in different measurements (Strebulaev and Yang, 2013). The table also illustrates low-leverage and high-leverage phenomena in detail from different perspectives. The first two columns display calendar years and number of companies listed each year. From Columns 3 to 5, we see that the leverages - the ratio of debt to total asset in book value - with debts, measured in all three different ways, are quite stable, i.e., the leverage stays around 0.23 for the one measured using total debt (Lever_TDA), around 0.17 for the one measured using long-term debt (Lever_LDA), and 0.48 for the one measured using total liability (Lever_TLA). Columns 6 to 9 demonstrate statistics of low-leverage firms that are defined in different measurements: almost all zero-leverage (AZTD) firms that have book leverage not exceeding 5% in a given year, ZL firms that have zero book debt (DLTT + DLC = 0, where DLTT and DLC refer to the Compustat long-term debt and debt in current liabilities respectively), ZLTD firms that have zero long-term debt (DLTT = 0), and NPND firms that have non-positive net debt in a given year. Shown in Column 6 of Table 5.1, from 1962 to 2017, 13.83% of firm-years exhibit zero leverage (measured in total debt) on average, ranging from a minimum of 6.01% in 1974 to 23.82% in 2013. Furthermore, the proportion of zero-leveraged firms has been growing since late 1970's. When only long-term debt is considered in Column 8, around 19.97% of the sample carries no long-term debt, meaning that 30.75% of firms with zero long-term debts finance their operations using short-term debts. Similar to the "ZL" definition, the growing trend of ZLTD firms has been stable since late 1970's. Relaxing the cut-off of absolute "0" to "5%", we see a bigger proportion of AZTD firms at 25%. Finally, the fraction reaches at an astonishing 37% if cash is considered as negative debt, varying from 18.49% to 52.96%. The analyses of high-leveraged firms are illustrated from Columns 10 to 12 in Table 5.1. High-leveraged firms are the firms whose leverage exceeds 75% of all firms in the sample, i.e., 0.367 in our case. The leverages measured in these three columns are the ratios of the book value of "total debt", "long-term debt", and "total liability" to the book value of "total asset". Though the results are not as impressive as they are for the low-leveraged firms, we still observe a fluctuating growth, especially after 2011.[11]

Furthermore, since firms listed in different markets demonstrate slightly different patterns, we also explore leverage (tail) patterns on a subsample of firms based on the markets they are listed in within this chapter.

[11] The evolution of debt and leverage calculated in different measurements for NYSE firms and Nasdaq firms are illustrated in Appendix Table B.2 and Appendix Table B.3.

Table 5.1: Leverage and frequency of zero-leverage firms

Column N presents the number of firms in the sample. Column Lever_TLA, Lever_TDA, and Lever_LDA report the average leverage measured using total liability, total debt, and total long-term debts respectively. ZL firms have zero book debt (DLTT + DLC = 0, where DLTT and DLC refer to the Compustat long-term debt and debt in current liabilities respectively). ZLTD firms have zero long-term debt (DLTT = 0). Almost all zero-leverage (AZTD) firms have book leverage not exceeding 5% in a given year. NPND firms have non-positive net debt in a given year, i.e., $DLTT + DLC - CHE \leq 0$, where CHE refers to the Compustat cash holdings. Columns ZL, ZLTD, AZTD, and NPND report the corresponding fractions of firms relative to the total sample in each year. On the other hand, HL, HLTD, and HLL report the percentage of firms that have leverages higher than 75% of the full sample, measured using total debt, long-term debt, and total liability exceeding 0.9 each year.

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1962	1778	0.20	0.16	0.40	25.82%	15.97%	21.99%	44.94%	19.18%	19.57%	0.45%
1963	1948	0.20	0.16	0.41	24.59%	15.86%	21.36%	43.53%	19.35%	20.43%	0.51%
1964	2102	0.21	0.16	0.41	24.02%	14.94%	21.17%	39.96%	19.36%	21.22%	0.52%
1965	2249	0.22	0.17	0.43	22.19%	13.21%	19.70%	37.79%	21.30%	21.65%	0.58%
1966	2404	0.23	0.18	0.45	19.13%	11.86%	17.64%	32.32%	22.75%	22.46%	1.00%
1967	2532	0.25	0.18	0.45	17.38%	10.07%	15.64%	29.66%	24.33%	24.29%	0.83%
1968	3025	0.25	0.19	0.46	17.36%	9.36%	14.51%	29.52%	24.36%	24.17%	1.26%
1969	3286	0.26	0.18	0.47	15.31%	8.28%	12.72%	27.12%	26.75%	23.83%	0.79%
1970	3300	0.28	0.19	0.48	12.79%	6.97%	10.97%	22.33%	30.91%	26.18%	0.91%
1971	3355	0.27	0.20	0.47	13.80%	6.89%	11.03%	23.99%	30.04%	26.65%	0.77%
1972	3445	0.26	0.19	0.47	14.69%	6.94%	11.12%	25.22%	27.87%	26.79%	0.78%
1973	3477	0.27	0.19	0.49	12.63%	6.30%	10.67%	19.44%	27.90%	26.86%	0.63%
1974	4872	0.29	0.19	0.52	12.34%	6.01%	11.58%	18.49%	32.29%	26.91%	1.23%
1975	4871	0.28	0.20	0.50	14.43%	7.10%	11.85%	21.54%	30.18%	27.02%	1.48%
1976	4851	0.27	0.19	0.50	15.01%	7.34%	11.85%	24.28%	27.81%	25.29%	1.24%
1977	4808	0.27	0.19	0.51	14.75%	7.07%	11.48%	22.86%	28.45%	26.41%	1.44%
1978	4669	0.27	0.20	0.52	14.11%	7.03%	10.84%	22.10%	29.30%	28.08%	1.46%
1979	4569	0.28	0.20	0.53	13.42%	6.52%	10.70%	21.14%	30.97%	28.50%	1.07%
1980	4621	0.27	0.20	0.53	14.26%	6.77%	10.89%	22.07%	30.27%	28.89%	1.28%
1981	4678	0.26	0.19	0.50	16.16%	7.48%	11.80%	26.89%	27.79%	26.53%	0.98%

Continuation of Table 5.1:

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1982	4800	0.26	0.19	0.50	16.98%	7.71%	11.81%	27.58%	28.85%	27.58%	1.25%
1983	4994	0.25	0.18	0.49	19.76%	8.31%	12.58%	32.96%	26.43%	24.69%	1.42%
1984	5063	0.25	0.18	0.49	19.81%	8.28%	12.70%	32.25%	25.83%	24.41%	1.20%
1985	5192	0.26	0.18	0.50	18.86%	8.36%	13.04%	30.76%	28.56%	25.67%	1.71%
1986	5362	0.27	0.18	0.51	19.54%	9.34%	15.03%	32.75%	30.81%	26.78%	2.11%
1987	5412	0.27	0.19	0.51	19.88%	9.74%	15.24%	32.37%	30.17%	27.83%	2.44%
1988	5234	0.27	0.19	0.51	19.72%	9.95%	15.84%	31.58%	31.10%	28.28%	2.31%
1989	5061	0.27	0.19	0.52	20.31%	10.02%	16.46%	31.08%	31.61%	29.01%	2.98%
1990	5008	0.27	0.19	0.52	20.91%	10.56%	16.65%	31.35%	31.35%	28.37%	2.72%
1991	5107	0.26	0.18	0.51	22.50%	11.10%	17.21%	33.74%	29.41%	27.63%	2.33%
1992	5393	0.24	0.17	0.49	25.37%	12.18%	18.97%	35.86%	26.92%	25.85%	2.08%
1993	5773	0.22	0.16	0.48	26.94%	13.32%	20.54%	38.11%	23.77%	24.56%	1.65%
1994	6060	0.22	0.16	0.49	27.33%	13.33%	19.88%	38.00%	23.55%	23.71%	1.68%
1995	6640	0.23	0.16	0.49	27.94%	14.02%	20.78%	39.08%	24.04%	24.53%	1.96%
1996	6927	0.22	0.16	0.47	30.35%	14.85%	21.41%	41.94%	23.42%	23.83%	1.95%
1997	6717	0.22	0.17	0.47	30.97%	15.13%	21.32%	41.83%	23.89%	25.52%	2.35%
1998	6663	0.22	0.17	0.48	31.07%	15.71%	22.48%	42.88%	25.39%	26.53%	2.73%
1999	6569	0.22	0.17	0.47	32.04%	15.74%	23.44%	43.17%	25.99%	26.29%	2.60%
2000	6153	0.22	0.16	0.47	34.19%	16.79%	25.08%	44.82%	25.32%	24.65%	2.68%
2001	5442	0.22	0.16	0.47	34.05%	17.73%	26.07%	45.44%	24.49%	24.55%	2.50%
2002	5085	0.21	0.16	0.48	34.32%	18.94%	26.92%	46.31%	22.65%	23.48%	2.48%
2003	4895	0.20	0.15	0.47	35.65%	20.59%	28.42%	48.80%	20.45%	22.90%	2.00%
2004	4803	0.19	0.15	0.45	37.62%	21.67%	29.25%	50.82%	18.36%	21.40%	1.92%
2005	4635	0.18	0.14	0.46	38.30%	22.20%	29.45%	51.46%	17.67%	20.32%	1.77%
2006	4488	0.18	0.14	0.45	37.54%	22.33%	29.57%	50.80%	17.76%	20.08%	1.49%
2007	4331	0.18	0.14	0.45	38.37%	22.67%	30.22%	51.10%	18.52%	20.99%	2.08%
2008	4018	0.20	0.15	0.47	35.99%	22.20%	30.99%	47.78%	21.23%	23.79%	2.14%
2009	3782	0.19	0.15	0.46	37.47%	22.53%	31.31%	50.98%	19.20%	22.29%	1.72%
2010	3716	0.18	0.14	0.46	38.94%	23.01%	31.97%	52.96%	16.90%	20.69%	1.61%

Continuation of Table 5.1:

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
2011	3660	0.18	0.15	0.46	37.57%	23.42%	31.78%	50.63%	17.38%	22.16%	1.56%
2012	3662	0.19	0.15	0.47	36.05%	22.91%	31.65%	49.89%	19.31%	23.29%	1.77%
2013	3677	0.19	0.16	0.46	36.44%	23.82%	31.25%	50.80%	19.50%	24.07%	2.07%
2014	3619	0.20	0.17	0.47	34.43%	23.02%	30.23%	48.94%	21.80%	26.25%	1.88%
2015	3401	0.21	0.18	0.48	33.25%	22.49%	29.23%	48.02%	23.64%	28.52%	2.50%
2016	3181	0.22	0.19	0.49	31.34%	21.19%	27.98%	45.83%	24.11%	30.84%	2.29%
2017	2203	0.24	0.21	0.51	26.15%	17.43%	22.29%	39.81%	26.60%	35.54%	2.41%
Mean	4421	0.23	0.17	0.48	25.04%	13.83%	19.97%	36.92%	24.95%	25.15%	1.67%

End of Table

5.3 Empirical Model and Variables

5.3.1 Empirical Model

We use the traditional panel data analysis method to explore the role of production flexibility on leverage with the empirical form presented below:

$$Y_{it} = \alpha + X'_{it}\beta + F_i + \epsilon_i(i = 1, 2, ..., N; t = 1, 2, ..., T),$$

where Y_{it} is the dependent variable of the model and X_{it} represents independent variables. Y_{it} differs according to the tests we are working on. In the leverage test, the dependent variable is leverage and is the static or shifting behavior of left/right tails in the exploration of tail phenomenon. X_{it} includes key variables that we want to explore as well as control variables. The key variables are mainly production flexibility proxies and agency problem proxies while control variables are the factors that have been proven important in determining leverage based on our literature review. F_i refers to the firm fixed effects designed to capture the specific effects engendered by firms. ϵ_i is the model residual.

5.3.2 Dependent Variables

The dependent variable differs according to different tests we are working on. It is leverage in leverage exploration, while in leverage tail exploration, it is either the measurement of left (right) tails state or the indicator of state transition behavior.

(1) Leverage Measurements

In the leverage test, the dependent variable is the leverage $Lever\_TDA$ which is measured as the ratio of total debt to total asset, that is:

$$Lever\_TDA = \tfrac{TotalDebt}{TotalAsset},$$

where $TotalDebt$ includes long-term debt ("DLTT" field from Compustat) and debt in current liabilities ("LCT" field from Compustat), and $Totalasset$ comes from "AT" from Compustat.

(2) Left Tail Measurements

We study the impact of product flexibility on left tail state and left tail transition behavior with measurements described below:

1) Left Tail State: $AZTD$

To understand how production flexibility and agency problems affect the state of low-leverage phenomenon, we set Y_{it} to be $AZTD$ with $AZTD$ defined as:

$$AZTD = \begin{cases} 1, & \text{if} \quad Lever\_TDA <= 0.05, \\ 0, & \text{if} \quad Lever\_TDA > 0.05. \end{cases}$$

where $AZTD$ is a dummy variable indicating whether the leverage is in the low-leverage region. In our case, we use 0.05 as a threshold following Strebulaev and Yang's research (Strebulaev and Yang, 2013). As is implied by the name, $AZTD$ is set to 1 if the firm holds a leverage lower than 0.05 and 0 otherwise.

2) Left Tail Transition: $AZTD\_C$

Besides the exploration of low-leverage state, we further explore the state transition between Low-leverage firms and Not-low-leverage firms. To understand the transition, we set Y_{it} to be $AZTD\_C$ with $AZTD\_C$ defined as:

$$AZTD\_C = \begin{cases} 1, & \text{if} \quad Lever\_TDA\_Lag > 0.05 \quad \text{and} \quad Lever\_TDA <= 0.05 \\ 0, & \text{if} \quad Lever\_TDA\_Lag > 0.05 \quad \text{and} \quad Lever\_TDA > 0.05. \end{cases}$$

where $AZTD\_C$ is a dummy variable indicating whether a firm transits from a non-AZTD firm to a AZTD firm. $AZTD\_C$ is set to 1 if the firm lowers its leverage from the high level at above 0.05 to the lower level at less than 0.05, and is set to 0 otherwise.

(3) Right Tail Measurements

We also study the impact of product flexibility on right tail state and right tail transition behavior with measurements described below:

1) Right Tail State: HTD

To understand how production flexibility and agency problems affect the state of High-leverage phenomenon, we set Y_{it} to be HTD with HTD defined as:

$$HTD = \begin{cases} 1, & \text{if} \quad Lever\_TDA >= 0.33595 \\ 0, & \text{if} \quad Lever\_TDA < 0.33595 \end{cases}$$

where HTD is a dummy variable indicating whether the leverage is at a high level, which is defined as the top 75% of the overall sample, that is 0.33595 in our case. We have HTD equal to 1 if the leverage of a firm is higher than 0.33595, and 0 otherwise.

2) Right Tail Transition: HTD_C

To understand the transition of Not-high-leverage firm to High-leverage firm, we set Y_{it} to be HTD_C with HTD_C defined as:

$$HTD\_C = \begin{cases} 1, & \text{if} \quad Lever\_TDA\_Lag < 0.33595 \quad \text{and} \quad Lever\_TDA >= 0.33595 \\ 0, & \text{if} \quad Lever\_TDA\_Lag < 0.33595 \quad \text{and} \quad Lever\_TDA < 0.33595 \end{cases}$$

Where HTD_C is a dummy variable indicating whether a firm transits from a Not-high-Leverage firm to a High-Leverage firm. HTD_C is set to 1 if the firm increases its leverage from the level lower than 0.33595 to the lever higher than 0.33595, and it it set to 0 otherwise.

(3) Key Variables

We adopt one or multiple measurements for key variables including production flexibility, agency problem concerns and financial constraint to ensure robust model results.

1) Production Flexibility

There are multiple ways to measure production flexibility. As discussed in the literature review, we adopt a measure that does not rely on proprietary data. The variable used as a proxy of production flexibility, $ProductionFlexibility$, is the volume flexibility measure borrowed from Jack and Raturi, 2009. It is calculated as:

$$ProductionFlexibility = ROS * Log(\frac{SD_{sales}}{\sqrt{(SD^2_{inventory}+SD^2_{CGS})}}).$$

The ratio measures the relative efficiency of firms in utilizing inventory buffers and cost control to support sale fluctuations. The higher is the ratio, the more production flexibility exists. In other words, firms with higher flexibility is able to resist sale fluctuations with less variations in product costs. In the formula above, ROS refers to the return of sales while SD_{sales}, $SD_{inventory}$, and SD_{CGS} refer to the standard deviation of sales, the standard deviation of inventories, and the standard deviation of total costs of goods in the previous three years respectively.

2) Financial Constraint

Four measurements are adopted as proxies of financial constraints. These four measurements are all dummy variables but they measure financial constraints from different perspectives and are discussed in detail:

(I) $FinancialConstraintDummy$

The first proxy is a dummy variable $FinancialConstraintDummy$ that is measured from the perspective of debt rating (Titman and Roberto, 1988) with the following formula:

$$FinancialConstraintDummy = \begin{cases} 0, & \text{if} \quad Debt\_rating < 13 \\ 1, & \text{if} \quad Debt\_rating >= 13. \end{cases}$$

where $Debt\_rating$ is the annual average debt rating of senior debt or subordinated debt. $FinancialConstraintDummy$ takes a value of 1 if the $Debt\_rating$ is higher than 13 and 0 otherwise. The fact that $Debt\_rating$ is higher than 13 means that

S&P rates the firm debt as investment grade and that these firms will encounter less constraints in financing. In this case, $FinancialConstraintDummy$ equals 1. When the $Debt\_rating$ is less than 13, firm debts are rated as insecure investment and thus leading to constrained access to financing market. In this case, we have $FinancialConstraintDummy$ as 0. Unfortunately, Compustat database does not report bond ratings before 1985. Thus, two options of missing treatment are tested. One is to assign 0 to $Debt\_rating$ for missing values. The other is to remove these observations that have no bond rating. The two treatments are tested in the regression explorations later on.

(II) $Z - Dummy$

The second measure used to represent financial constraints is a dummy variable $Z - Dummy$ that is derived from the Altman Z Score (Altman, 1968; Beaver, 1966, 1968a,b) :

$$\text{Z-Dummy}=\begin{cases} 0, & \text{if} \quad Z - score < 1.81, \\ 1, & \text{if} \quad Z - score > 2.99. \end{cases}$$

The $Z - score$ in the formula above is Altman Z Score, which is used to predict the likelihood of a business going bankruptcy within the coming two years. Information needed to calculate Altman Z Score comes from public financial statements. The calculation of Z Score is as follows:

$$ZScore = 0.012A + 0.014B + 0.033C + 0.006D + 0.01E$$

where

A = working capital / total assets;

B = retained earnings / total assets;

C = earnings before interest and tax / total assets;

D = market value of equity / total liabilities;

E = sales / total assets.

For a company that has an Altman Z Score below 1.85, it would be thought on the verge of bankruptcy. The companies that have scores above 3 are thought to not likely to go bankrupt.

(III) $KZ - Dummy$

The third access constraint indicator $KZ - Dummy$ is related to $KZ - Index$. This dummy variable is derived using the following formula below. The threshold is the median $KZ - Index$ of the data sample. $KZ - Dummy$ is set to 1 when it is above the median and 0 otherwise, that is,

$$KZ - Dummy = \begin{cases} 1, & \text{if } KZ - Index \geq Median\ of\ the\ sample \\ 0, & Otherwise. \end{cases}$$

where $KZ - Index$ indicates the difficulty of firms financing its ongoing operations. It is a widely used financial constraint measure and is calculated as below (Lamont et al., 2015):

$$\begin{aligned} KZ - Index = {} & -1.001909 * CashFlows/K + 0.2826389 * Q + 3.139193 * Debt/TotalCapital \\ & - 39.3678 * Dividends/K - 1.314759 * Cash/K. \end{aligned}$$

where,

Cash Flows = Income Before Extraordinary Items + Total Depreciation and Amortization;

K = PP&E in previous year;

Q = (Market Capitalization + Total Shareholder's Equity- Book Value of Common Equity-Deferred Tax Assets) / Total Shareholder's Equity;

Debt = Total Long Term Debt + Notes Payable+ Current Portion of Long Term Debt;

Dividends = Total Cash Dividends Paid (common and preferred);

Cash = Cash and Short-Term Investment.

(IV) $SA - Dummy$

The last proxy of financial constraints is $SA - Dummy$ which is related to the SA index. The SA index is utilized because the KZ-Index is shown to be invalid in certain circumstances but firm size and age are found to be particularly useful predictors of financial constraint levels (Hadlock and Pierce, 2010). The index is calculated as:

$$SA - Dummy = \begin{cases} 1, & \text{if } SA \geq Median\ of\ the\ sample, \\ 0, & Otherwise. \end{cases}$$

with $SA = (-0.737 * Size) + (0.043 * Size^2) - (0.040 * Age)$, where $Size$ equals the log of inflation-adjusted book assets and Age is the number of years that the firm has been listed with a non-missing stock price on Compustat. To calculate this index, $Size$ is winsorized (i.e., capped) at (the log of) \$4.5 billion, and Age is winsorized at 37 years (Hadlock and Pierce, 2010).

3) Agency Problem Concern

Agency problem concerns illustrate the conflict of interest between two parties, i.e., debt holders versus shareholders/managers. We adopt two opposite measures to illustrate the severity of the agency problem quantitatively.

(I) $NonAgencyDummy$

One agency problem concerns measurement is related to the ownership of the top five executives and it is formulated as:

$$NonAgencyDummy = \begin{cases} 1, & \text{if the percentage proportion of top five executives is 0}, \\ 0, & \text{if the percentage proportion is higher than or equal to 0.592}. \end{cases}$$

$NonAgencyDummy$ is a dummy variable indicating whether the benefits of the top five executives are aligned with shareholders. The variable illustrates whether managers make decisions that would benefit the firm as a whole. It equals to 1 if the percentage of top five executives is less than or equal to one-third of the total observations (0 in our case), meaning that agency problems are low and managers are more likely to make decisions for the whole firm's benefit. Otherwise, the dummy variable equals to 0 if the percentage of top five executives is higher than or equal to two-thirds of the total observations (0.592 in our case), meaning that the agency problem is high and managers are more likely to act on the behalf of shareholders.

(II) Tangibility

Another measure of agency problem proxy is tangibility. The logic is that high tangibility represents a high collateral to debt holders and thus a high tangibility would proxy a low level of agency problems.

Finally, the interaction between production flexibility and agency problem is also introduced in our regression as $ProFlex\_Agen$. The regression is formulated as below:

$$\text{ProFlex\_Agen}=\text{ProductionFlexibility}*\text{NonAgencyDummy} .$$

(4) Control Variables

The control variables are borrowed from existing research studies (Frank and Goyal, 2009) and are shown in Table 5.2.[12] In our research here, we only cover the four most important variables in our specific context and leave all the macro variables for the robustness test. These four most important variables are picked from previous studies (Graham et al., 2015) and they are firm size, market value to book value ratio, profitability, and tangibility.

We show the empirical results based on these methods and variable proxies in the next section.

[12] Detailed definition of these variables, independent variables, and other key variables mentioned earlier are demonstrated in Appendix Table B.4

Table 5.2: Control variables

This table lists the control variables included in our research from 1962 to 2017. They can be referenced from (Frank and Goyal, 2009). The first column shows the variable name while the second column describes the meaning of each variable in detail.

$Variable$	Variable Description
$Capex$	Growth: capital expenditure to assets ratio.
$ChgAsset$	Growth: change in log assets.
$CPIExpec$	Macroeconomic conditions: expected change in the consumer price index over the coming year.
$Debt_Rating$	The rating of debt, is also called Supply-side factors. It is a debt rating dummy.
$FirmSize$	Firm size: Log of assets.
$InduChgAsset$	Industry growth: median of change in log assets.
$InduLeverTDA$	Industry growth: median of industry leverage (Total book value debt to assets).
$InvTaxCr$	Taxes: investment tax credit-balance sheet to assets.
$Leverage\_Lag$	Leverage in previous year.
$MacroGr$	Macroeconomic conditions: difference in log of real gross domestic product in 1996 dollars.
$MacroProf$	Macroeconomic conditions: difference in log of aggregate annual corporate profits after tax for non-financial firms.
$MktBk$	Growth: market-to-book ratio.
$Profitability$	Profitability: ratio of operating income before depreciation to assets.
$ReturnVar$	Risk: variance of stock returns.
RnD	Nature of assets: research and development to sales.
SGA	Nature of assets: ratio of general, and administration expenses to sales.
$StockMarketReturn$	Stock market condition: index return.
$StockMeanRet$	Stock market condition: average raw return.
$Tangibility$	Nature of assets: ratio of Compustat item 8, net property, plant, and equipment to assets.
$TaxRate$	Taxes: top statutory tax rate.
$TermSprd$	Debt market conditions: term spread between 10-year interest rate and 1-year interest rate.

5.4 Empirical Results

This section reports model results regarding the role of production flexibility on leverage and leverage tails. We first demonstrate our findings on leverage in various conditions, then we further extend our analysis to left tail and right tail behaviors, respectively.

5.4.1 Role of Production Flexibility on Leverage

The empirical research starts with a univariate analysis of the variables involved in our analysis and then go through regression models to explore the role of production flexibility on leverage under various conditions. We do missing treatments before the analysis. The basic principle of missing value treatments is that any missing variables that represent firm characteristics are replaced with 0 and data with any missing macro variable is removed. Besides that, leverages that are out of range (0,1) are removed out of the sample.

(1) Univariate Analysis Results

After missing value treatments, we have 119,073 firm-year observations in total that have valid non-missing values for all variables as illustrated in the Table 5.3. We have

12,395 companies left in the data sample. The average leverage $Lever\_TDA$ is 0.23 with standard deviation at 0.187, minimum at 0, and maximum at 0.985. The leverage in the previous year $Lever\_Lag$ demonstrates a similar distribution to $Lever\_TDA$, which is consistent with the sense that leverage is sticky over years. Production flexibility varies a lot from -9.1 to 6.5. The average production flexibility over years is around 0.32.

Table 5.3: Descriptive statistics

This table reports the descriptive statistics of variables included in our research from 1962 to 2017. Column (1) describes the variable name and Columns (2) to (5) contain a report on the number of observations, the Mean, the Standard deviation, and the Minimum and Maximum of each individual variable.

Variable	N	Mean	Std Dev	Minimum	Maximum
Capex	119,073	0.068	0.072	-0.152	1.98
CPIExpec	119,073	-0.51	1.59	-5.00	4.00
Debt_Rating	119,073	0.14	0.34	0.00	1.00
FirmSize	119,073	5.00	2.12	-1.80	13.59
Grow_ChgAsset	119,073	0.10	0.30	-3.95	4.41
Indu_ChgAsset	119,073	0.08	0.10	-1.31	1.97
Indu_Lever_TDA	119073	0.21	0.13	0.00	0.94
IndusRegulated	119,073	0.03	0.16	0.00	1.00
InvTaxCr	119,073	0.001	0.005	-0.004	0.08
Lever_Lag	119,073	0.23	0.18	0.00	0.98
Lever_TDA	119,073	0.23	0.19	0.00	0.99
MacroGr	119,073	0.003	0.002	-0.003	0.01
MacroProf	119,073	0.01	0.18	-0.24	0.43
MktBk	119,073	1.59	35.56	-0.06	12253.26
NatureAsset_Unique	119,073	0.29	0.46	0.00	1.00
ProductionFlexibility	119,073	0.32	0.67	-9.06	6.51
Profitability	119,073	0.09	0.21	-7.37	3.25
ReturnVar	119,073	0.14	0.10	0.00	3.08
RnD	119,073	-0.01	1.55	-494.14	0.04
SGA	119,073	0.39	5.24	-20.60	806.00
StockMarketRet	119,073	0.10	0.18	-0.37	0.37
StockMeanRet	119,073	0.01	0.05	-0.36	1.12
Tangibility	119,073	0.31	0.23	0.00	1.00
Tax_Depr	118,993	0.31	0.23	0.00	1.00
Tax_NOLCF	119,073	0.22	1.25	-0.34	103.65
TaxRate	119,073	0.40	0.06	0.34	0.53
TermSprd	119,073	1.07	1.22	-1.43	3.38

To illustrate how production flexibility and leverage evolve over decades, a descriptive analysis of production flexibility and leverage by decades is shown in Table 5.4 below. We find that production flexibility has been growing over decades from 0.249 in the 1960's to 0.399 in the 2010's. Seen from the vantage of average leverage, we observe a decreasing trend from 0.24 in the 1960's to 0.21 in the 2010's after a slight increase from 0.26 in the 1970's. The minimum value has always been zero. In general, we observe a decreasing trend of leverage and an increasing trend of production flexibility over decades in general.

Table 5.4: Descriptive statistics of production Flexibility and leverage by decades

This table further shows some simple statistics of both production flexibility and leverage by decades, which includes the number of observations, the Mean, the Standard deviation, and the Minimum and Maximum of both variables.

Production Flexibility						
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	6183	24308	22513	23371	31085	11613
Mean	0.25	0.27	0.30	0.31	0.35	0.40
Std.	0.46	0.42	0.54	0.69	0.82	0.86
Min	-4.06	-4.59	-8.43	-9.06	-8.36	-7.49
Max	5.06	4.57	5.49	6.51	5.20	5.44
Lever_TDA						
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	6183	24308	22513	23371	31085	11613
Mean	0.24	0.26	0.26	0.23	0.19	0.21
Std.	0.16	0.17	0.19	0.20	0.19	0.19
Min	0	0	0	0	0	0
Max	0.91	0.94	0.94	0.99	0.94	0.91
End of Table						

(2) Regressions of Leverage on Production Flexibility

Table 5.5 reports the regression of leverage on production flexibility with other firm characteristics controlled. Panel regressions with fixed firm effects are estimated for each decade, and all dependent variables are lagged behind by one year. The results shown in the table mirror the previous work on leverage (Graham et al., 2015). Increases in the tangibility of assets is observed to raise firm debt ratio, even though the impact varies among decades. We also find that more profitable firms (EBITDA/Sales) tend to have lower leverages, which is consistent with the thought that profitable firms use their earnings to pay off debt and lower leverage levels. Leverage is also found to be positively correlated with size (Graham et al., 2015). This makes sense since larger firms are less risky and more diversified, and thus have less financial distress costs and lower probabilities of financial distress. Moreover, the economy of scale and lower issue costs further motivate firms to adopt a relative higher leverage than firms with smaller size. Though the magnitudes of the coefficients vary among decades, they are statistically significant and in their economic meanings. Besides these control variables, production flexibility is found to be significantly and negatively related with leverage since 1980's but not significant from 1962 to 1980. On average, a 1% increase in production flexibility has decreased leverage by 0.38% from 1981 to 2017.

We further test the results based on market where companies are listed. The companies in the data sample are listed in NYSE or NASDAQ market. However, stocks traded on NASDAQ are known to be growth oriented and volatile while firms listed in NYSE are more mature. Furthermore, investors of stocks in these two markets are different. Based on these facts, firms listed in NYSE may behave differently with firms listed in Nasdaq with respect to the relationship between production flexibility and leverage. We find similar results in NYSE-only firms shown in Appendix Table B.5. The effect of production flexibility has been significant since the 1990's, and the impact is bigger than it is for the full sample. On average, a 1% increase in production flexibility helps decreasing leverage by 0.62% from 1991 to 2017. For Nasdaq-only firms, no significant roles of production flexibility are found across decades (shown in Table B.6). However, based on the t-statistics, the impact has been found to be increasing in magnitude and in statistics since the 1990's even though it has not reached significant levels statistically.

The negativity found between production flexibility and leverage is consistent with the findings in previous studies (MacKay, 2003), which state that firms with higher production flexibility tend to use less leverage. The regression result is also consistent with the simple statistical results before (refer to Table 5.4). In addition, comparing Column "(1981-1990)" with Column "(1986-1990)" to understand the impact of removing observations before 1985 on our regression results, we find that the relationship between production flexibility and leverage remains unchanged, except for the disappearance of the significance levels. Realizing that replacing the missing debt rating with 0 would overestimate the role of production flexibility on leverage, we exclude these before-1985 observations in later explorations to be conservative.

Table 5.5: Leverage on production flexibility

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The dependent variable is leverage measured using total debt divided by total asset. All dependent variables are lagged behind by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

	Leverage on Production Flexibility by Decades						
	(1962-1970)	(1971-1980)	(1981-1990)	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.08***	0.04***	0.05***	0.06***	0.04***	0.02***	0.07***
	(29.24)	(22.10)	(29.71)	(15.16)	(26.96)	(18.32)	(25.56)
MktBk	0.002	0.01***	-0.002**	0.001	-0.002***	-0.002***	0.001
	(1.22)	(8.68)	(-2.82)	(0.91)	(-4.95)	(-5.72)	(0.70)
Tangibility	0.12***	0.23***	0.15***	0.08***	0.20***	0.15***	0.13***
	(6.04)	(23.38)	(15.12)	(4.64)	(18.50)	(16.16)	(7.16)
Profitability	-0.23***	-0.27***	-0.10***	-0.10***	-0.06***	-0.05***	-0.06***
	(-10.19)	(-31.59)	(-13.68)	(-7.92)	(-9.35)	(-11.02)	(-7.54)
ProductionFlexibility	0.003	0.002	-0.005**	-0.003	-0.003*	-0.002*	-0.005**
	(0.98)	(1.39)	(-2.70)	(-1.02)	(-2.28)	(-2.51)	(-3.00)
Intercept	-0.12***	0.07***	-0.01	0.01	-0.01	0.02*	-0.26***
	(-8.48)	(8.48)	(-1.02)	(0.41)	(-1.22)	(2.03)	(-13.89)
N	6183	24308	22513	10245	23371	31085	11613
Adjusted R^2	0.19	0.09	0.08	0.05	0.06	0.027	0.079
Firm Fixed Effect	yes	yes	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(3) Regressions of Leverage on Production Flexibility by Industry

We further test whether the role of production flexibility on leverage would differ in manufacturing industries and in service industries. Realizing that the two components of production flexibility, - cost of goods and cost of inventory -, differ between manufacturing firms and service firms, we separate our firms based on whether they are in the service industry. Companies whose SIC codes fall into the range of 7000 to 8999 are classified as service industry companies and others are simply classified as manufacturing industry companies. We report the results in Table 5.6. For the service industry shown in Panel A, the relationship between production flexibility and leverage fluctuates and does not show a clear pattern, which could be partly due to the small data size. For non-service manufacturing industry shown in Panel B, a consistent pattern with previous regressions is found, i.e., production flexibility is getting more and more negatively related with leverage and the significance level has been growing over decades. In general, around 1% increase in production flexibility would lead to around 0.004% decrease in leverage. Although realizing the noise of including service industry in the regression regarding the role of production flexibility on leverage, we still include them in our analysis to ensure that any pattern we capture can be applied to the real economy as a whole.

Table 5.6: Leverage on production flexibility by industry

This table reports the result of panel data analysis on the sample over 1986 to 2017, as per whether firms are in the service industry and by decades. The sample includes all firms in the Compustat and CRSP merged database from 1986 to 2017, with financial firms excluded. The dependent variable is leverage measured using total debt divided by total asset. All dependent variables are lagged behind by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. Regressions for service firms are reported in Panel A while regressions for non-service firms are reported in Panel B. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Leverage on Production Flexibility by Decades and by Whether Firms are in Service Industry				
Panel A: Service Firm				
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.06***	0.03***	0.03***	0.07***
	(7.03)	(10.04)	(12.67)	(11.93)
MktBk	0.002	-0.001*	-0.001	0.01*
	(0.89)	(-2.04)	(-0.56)	(2.45)
Tangibility	0.10**	0.23***	0.11***	0.10*
	(2.65)	(9.57)	(5.07)	(2.06)
Profitability	-0.06*	-0.03**	-0.04***	-0.11***
	(-2.25)	(-2.63)	(-5.26)	(-4.98)
ProductionFlexibility	-0.02*	-0.001	0.0005	-0.00002
	(-2.57)	(-0.27)	(0.24)	(-0.01)
Intercept	0.02	0.01	-0.05**	-0.25***
	(0.53)	(0.93)	(-3.07)	(-6.55)
N	1565	4294	6753	2294
Adjusted R^2	0.07	0.07	0.03	0.09
Firm Fixed Effect	yes	yes	yes	yes
Panel B: Non-Service Firm				
FirmSize	0.05***	0.04***	0.02***	0.07***
	(13.31)	(25.07)	(13.70)	(22.76)
MktBk	0.001	-0.002***	-0.003***	-0.0002
	(0.39)	(-4.72)	(-6.22)	(-0.15)
Tangibility	0.0746***	0.19***	0.16***	0.13***
	(3.76)	(15.81)	(15.55)	(6.77)
Profitability	-0.11***	-0.06***	-0.05***	-0.05***
	(-7.84)	(-9.24)	(-9.72)	(-6.37)
ProductionFlexibility	0.0003	-0.004*	-0.003**	-0.01**

Continuation of Table 5.6:				
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
	(0.11)	(-2.36)	(-3.06)	(-3.17)
Intercept	0.01	-0.02	0.04***	-0.27***
	(0.49)	(-1.70)	(4.03)	(-12.42)
N	8680	19077	24332	9319
Adjusted R^2	0.05	0.06	0.03	0.08
Firm Fixed Effect	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(4) Regressions of Production Flexibility on Leverage by Financial Constraint Situation

We further explore the role of financial constraints on the relationship between production flexibility and leverage by bringing in financial constraints factor into regressions with results reported in Table 5.7. The dependent variable is leverage measured as total debt divided by total asset. All independent variables are lagged behind by one year. Independent variables include $ProductionFlexibility$ and other control variables. Panel regressions with fixed firm effects are estimated for each decade. We separate our data to two subgroups based on whether firms encounter financial constraints from Panel A to Panel D to explore how financial constraints affect the relationship between production flexibility and leverage. In Panel A, we test the robustness of the debt-rating measurement by removing the observations of years prior to 1986 and report both financial constraint and non-financial constraint situations since we found the replacing missing debt-rating with 0 would overestimate the role of production flexibility. Other measurements of financial constraints are tested in Panel B to Panel D. In detail, financial constraint is measured by Z-Dummy in Panel B, by KZ-Index in Panel C and by SA-Dummy in Panel D.

Table 5.7: Leverage on production flexibility segmented by financial constraints

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The dependent variable is leverage measured using total debt divided by total asset. All dependent variables are lagged behind by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. In Panel A, we test the robustness of debt-rating measurement by removing the observations with year prior to 1986 and report and financial constraint situations and non-financial constraint situations. In Panel B, we further use the Z-Dummy as a financial constraint situation and rerun the regressions. We utilize KZ-Index in Panel C and SA-Dummy in Panel D, respectively. Regression coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Panel A: financial constraint by debt_rating with year prior to 1986 deleted

	Constrained				Not Constrained			
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.05***	0.04***	0.03***	0.06***	0.06***	0.02***	-0.02***	0.06***
	(13.26)	(23.20)	(17.59)	(18.25)	(5.42)	(5.79)	(-6.39)	(13.51)
MktBk	0.001**	-0.001***	-0.002***	-0.001	0.06***	0.02***	-0.02***	0.06***
	(0.39)	(-3.72)	(-5.28)	(-0.96)	(3.70)	(-1.95)	(-2.56)	(8.17)
Tangibility	0.09***	0.22***	0.17***	0.15***	0.026	0.04	0.09***	0.13***
	(4.7)	(18.60)	(15.26)	(6.07)	(0.57)	(1.59)	(5.19)	(4.42)
Profitability	-0.10***	-0.05***	-0.04***	-0.05***	-0.15	-0.17***	-0.16***	-0.10**
	(-7.28)	(-7.96)	(-9.89)	(-6.29)	(-1.79)	(-3.99)	(-8.69)	(-3.20)
ProductionFlexibility	-0.002	-0.003*	-0.002*	-0.006**	-0.005	0.003	-0.001	-0.0003
	(-0.8)	(-1.99)	(-2.08)	(-2.98)	(-0.81)	(0.96)	(-0.56)	(-0.10)
Intercept	0.04*	-0.01	-0.03**	-0.22***	-0.17	0.16***	0.46***	-0.28***
	(2.14)	(-1.12)	(-3.14)	(-10.81)	(-1.83)	(4.20)	(17.63)	(-6.47)
N	8839	19641	22599	7666	1406	3730	8486	3947
Adjusted R^2	0.05	0.07	0.03	0.07	0.05	0.03	0.02	0.08
Firm Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes

Panel B: financial constraint by Z-Score with year Prior to 1986 Deleted

	Constrained				Not Constrained			
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.05***	0.04***	0.02***	0.07***	0.05***	0.02**	0.03***	0.04***
	(10.95)	(24.77)	(15.85)	(23.78)	(4.83)	(3.28)	(5.52)	(4.96)
MktBk	0.002	-0.001**	-0.002***	0.001	0.01	-0.01*	-0.003	0.02***
	(1.59)	(-3.22)	(-5.20)	(0.57)	(1.77)	(-2.27)	(-0.82)	(3.87)
Tangibility	0.07***	0.20***	0.15***	0.13***	0.14**	0.13**	0.06	0.10
	(3.58)	(16.56)	(13.96)	(6.29)	(3.10)	(3.06)	(1.74)	(1.87)
Profitability	-0.09***	-0.05***	-0.04***	-0.06***	-0.18***	-0.05	-0.15***	-0.04

Continuation of Table 5.7:								
	(-6.26)	(-7.16)	(-9.33)	(-6.99)	(-3.34)	(-1.46)	(-6.35)	(-1.43)
ProductionFlexibility	-0.004	-0.004*	-0.003**	-0.004*	0.01	0.001	0.01	-0.03***
	(-1.39)	(-2.18)	(-3.23)	(-2.17)	(0.73)	(0.11)	(1.10)	(-4.46)
Intercept	0.04	-0.03**	0.02*	-0.28***	0.0	0.17***	0.08*	-0.34**
	(1.83)	(-3.01)	(2.13)	(-13.60)	(0.32)	(4.58)	(2.13)	(-2.98)
N	7670	17838	25207	9474	1048	2083	2201	952
Adjusted R^2	0.04	0.07	0.02	0.08	0.08	0.02	0.05	0.08
Firm Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes

Panel C: financial constraint by KZ-Index with year Prior to 1986 Deleted								
	Constrained				Not Constrained			
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.06***	0.04***	0.02***	0.07***	0.06***	0.03***	0.02***	0.07***
	-12.96	-22.71	-12.05	(17.25)	-7.47	-9.52	-11.22	-17.45
MktBk	-0.002	-0.004***	-0.01***	0.00002	0.004	-0.0002	-0.0004	-0.0001
	(-0.90)	(-6.33)	(-6.47)	(0.01)	-1.65	(-0.29)	(-0.78)	(-0.10)
Tangibility	0.07***	0.18***	0.15***	0.10***	0.12*	0.11***	0.09***	0.13**
	-3.49	-14.01	-12.85	(4.68)	-2.11	-3.37	-4.02	-3
Profitability	-0.10***	-0.07***	-0.06***	-0.09***	-0.09**	-0.04**	-0.04***	-0.05***
	(-6.61)	(-8.12)	(-7.13)	(-6.07)	(-2.79)	(-2.93)	(-6.09)	(-3.94)
ProductionFlexibility	-0.002	-0.002	-0.003*	-0.01*	-0.004	-0.005	-0.002	-0.003
	(-0.55)	(-1.12)	(-1.99)	(-2.29)	(-0.77)	(-1.76)	(-1.63)	(-1.32)
Intercept	0.02	0.01	0.07***	-0.21***	-0.11**	-0.01	-0.03**	-0.32***
	-1.07	-1.01	-5.33	(-7.52)	(-3.03)	(-0.92)	(-2.64)	(-11.57)
N	7915	17210	19917	6580	2330	6161	11168	5033
Adjusted R^2	0.05	0.06	0.02	0.07	0.06	0.03	0.02	0.09
Firm Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes

Panel D: financial constraint by SA-Dummy by Decades and year Prior to 1986 Deleted								
	Constrained				Not Constrained			
	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)	(1986-1990)	(1991-2000)	(2001-2010)	(2011-2017)
FirmSize	0.06***	0.04***	0.01***	0.07***	0.0455***	0.0392***	0.03***	0.05***
	-10.74	-15.75	-4.84	(15.87)	-8.07	-14.36	-14.33	-13.12
MktBk	0.01**	-0.01***	-0.005***	0.01**	0.001	-0.001	-0.002**	0.0003
	-3.11	(-5.15)	(-4.82)	(3.09)	-0.41	(-1.72)	(-3.08)	-0.21
Tangibility	0.04	0.18***	0.15***	0.11***	0.11***	0.20***	0.16***	0.15***

Continuation of Table 5.7:								
	-1.29	-11.77	-10.51	(4.53)	-4.56	-12.36	-12.06	-5.52
Profitability	-0.20***	-0.21***	-0.13***	-0.19***	-0.08***	-0.03***	-0.04***	-0.02**
	(-5.98)	(-10.72)	(-10.49)	(-8.42)	(-5.12)	(-4.94)	(-7.88)	(-2.71)
ProductionFlexibility	-0.001	-0.0008	-0.003	-0.005	-0.003	-0.003	-0.0004	-0.001
	(-0.30)	(-0.35)	(-1.63)	(-1.87)	(-0.91)	(-1.33)	(-0.34)	(-0.65)
Intercept	-0.08*	0.01	0.14***	-0.28***	0.09***	0.02*	-0.02	-0.14***
(-2.16)	-0.61	-8.07	(-8.10)	-5.25	-1.99	(-1.78)	(-5.96)	
N	4646	11686	15542	5807	5599	11685	15543	5806
Adjusted R^2	0.05	0.06	0.02	0.08	0.04	0.05	0.03	0.05
Firm Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Panel A, production flexibility is found to play a significant role on firms that have financial constraints since the 1980's, which is consistent with our theoretical exploration. However, production flexibility plays no significant role on leverage for firms that have free access to debt markets, partly because firms could raise more extra funds than needed in reality when market condition is good according to market timing theory. In Panel B, consistent results regarding the relationship between production flexibility and leverage are found for firms that are segmented by Altman Z score dummy for years before 2011. A possible reason for the insignificant results from 2011 to 2017 could be the small sample size.

In Panel C, when firms are grouped based on financial constraints measured in KZ-Index, production flexibility is found to be negatively related with leverage and this negativity has been growing over decades. The negativity turns important in the recent two decades. This result is consistent with what we have found before. However, we do not observe any significance of production flexibility on leverage when it is measured by SA-Dummy in Panel D, even though the significance has been growing in the last decade, reaching a marginally significant level.

These results are consistent with our logic in theoretical section about the underlying assumption of debt access. The results make sense since firms can produce any amount of goods beforehand without any shortage of funding. Additionally, the results fit in the phenomenon (Faulkender and Pertersen, 2006) that leverage is also affected by supply since firms could be rationed by their lenders. The same regressions are conducted for NYSE-only firms (shown in Table B.7) and Nasdaq-only firms (reported in Table B.8). Still, we find some significant role of production flexibility for NYSE firms when facing financial constraints, and no role when the access is free. For Nasdaq firms, similar to previous regressions, no significant roles are found, with statistical significance increasing over decades from the 1980's.

(5) Regressions of Leverage on Production Flexibility and agency problem concerns Within a Situation of financial Constraints Situation

Understanding the role of financial constraints on leverage decisions, we further bring in agency problem concerns to understand how this factor would affect the relationship between production flexibility and leverage. Regression results are reported in Table 5.8. Two measurements for agency problem are tested respectively in Panel A and Panel B. In Panel A, the agency dummy is related with whether the top five executive shares are above two-thirds of the population, and the agency problem is related to tangibility in Panel B.

Table 5.8: Leverage on production flexibility, financial constraints and agency dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The dependent variable is leverage measured using total debt divided by total asset. All dependent variables are lagged behind by one year. In Panel A, independent variables include $ProductionFlexibility$, Agency Problem Dummy $NonAgencyDummy$, the interaction between the two $ProFlex\_Agen$, and other control variables. Panel regressions with fixed firm effects are estimated on both financial constrained firms and not constrained firms for all firms (Column 1 and 2), NYSE firms (Column 3 and 4), and Nasdaq firms (Column 5 and 6). In Panel B, we test the robustness of agency problem logic by using tangibility as a proxy. All the 12 regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
Panel A: Top 5 executives' share as agency problem proxy						
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.03***	0.001	0.03***	-0.0005	0.02***	0.01
	(17.18)	(0.57)	(9.92)	(-0.22)	(11.73)	(1.71)
MktBk	-0.001**	-0.001	-0.01***	-0.001	-0.001	-0.0004
	(-2.86)	(-0.64)	(-3.71)	(-0.54)	(-1.32)	(-0.16)
Tangibility	0.14***	0.05***	0.11***	0.05**	0.15***	0.07
	(9.76)	(3.79)	(4.68)	(2.94)	(7.11)	(1.77)
Profitability	-0.08***	-0.18***	-0.12***	-0.21***	-0.06***	-0.10*
	(-8.96)	(-9.50)	(-5.14)	(-10.06)	(-4.91)	(-1.98)
ProductionFlexibility	0.0001	0.001	-0.01*	-0.001	0.002	0.01
	(0.05)	(0.28)	(-2.14)	(-0.52)	(1.00)	(0.91)
NonAgencyDummy	0.01***	0.002	0.01***	0.002	0.009***	0.001
	(5.37)	(1.46)	(5.11)	(1.53)	(4.29)	(0.30)
ProFlex_Agen	-0.01***	-0.001	-0.006	0.001	-0.01***	-0.01
	(-4.71)	(-0.68)	(-1.79)	(0.36)	(-4.25)	(-1.18)
Intercept	-0.05***	0.30***	-0.01	0.32***	-0.06***	0.20***
	(-4.48)	(14.63)	(-0.59)	(14.22)	(-4.14)	(3.68)
N	13585	10414	5285	8182	7207	1800
Adjusted R^2	0.04	0.01	0.05	0.016	0.03	0.005
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
Panel B: Tangibility as Agency Problem proxy						
FirmSize	0.02***	0.01***	0.01***	0.005***	0.01***	0.02***
	(21.31)	(5.79)	(9.52)	(3.41)	(9.86)	(4.17)

Continuation of Table 5.8:						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
MktBk	-0.002***	-0.0002	-0.004***	0.002	-0.003***	-0.001
	(-9.41)	(-0.21)	(-4.76)	(1.36)	(-7.45)	(-0.56)
Tangibility	0.22***	0.04***	0.18***	0.03*	0.26***	0.08**
	(35.92)	(4.05)	(14.40)	(2.33)	(28.35)	(2.74)
Profitability	-0.03***	-0.16***	-0.08***	-0.19***	-0.03***	-0.09*
	(-11.03)	(-10.36)	(-8.61)	(-11.09)	(-6.09)	(-2.35)
ProductionFlexibility	-0.001	-0.002	-0.01**	-0.01	0.001	0.004
	(-1.18)	(-0.59)	(-3.00)	(-1.77)	(0.52)	(0.65)
ProFlex_Tang	-0.01*	-0.001	0.004	0.006	-0.02**	-0.01
	(-2.51)	(-0.24)	(0.63)	(1.00)	(-3.05)	(-1.11)
Intercept	0.06***	0.27***	0.10***	0.29***	0.05***	0.20***
	(14.72)	(20.87)	(10.10)	(20.34)	(8.72)	(5.95)
N	58745	17569	15030	13113	28251	3354
Adjusted R^2	0.04	0.01	0.03	0.01	0.04	0.01
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

In Panel A, we find that *NonAgencyDummy* is positively related with leverage, meaning that the higher the shares held by the top five executives, the lower the leverage, which is consistent with existing research findings (Strebulaev and Yang, 2013). Furthermore, the interaction term *ProFlex_Agen* is significantly and negatively related with leverage, meaning that high executive ownership would mitigate the negativity between production flexibility and leverage for firms facing financial constraints, a result that is consistent with our previous findings and support our hypotheses H3B and H4B. The interaction between production flexibility and agency problem flips the role that production flexibility plays on leverage. For instance, in the first column of the regression on all firms, the production flexibility is negatively related with leverage with impact at around -0.00741 for low-manager-owned firms, but it is positively related with leverage with impact at around 0.00008 for high-manager-owned firms. For NYSE-only firms, the role of the interaction term is marginally significant with same negative sign. For Nasdaq-only firms, a slightly bigger impact is found with magnitude being -0.00795. So far, our findings help reconcile the opposite findings in Reinartz and Schmid, 2016 and MacKay, 2003. Compared to manufacturing firms, managers in utility firms usually have less ownership due to the special nature of such kinds of firms. Thus, it is more likely for us to find negative relationship between production flexibility and leverage in utility firms than in manufacturing firms.

A consistent result is found in Panel B with agency problem measured by tangibility. Since high tangibility represents a high collateral to debt holders, and thus less agency problem concern, it is legitimate to find that tangibility is positively related with debt ratio. The interaction term $ProFlex\_Tang$ is found to be significantly negative for firms under financial constraint situation, meaning that high production flexibility under high agency problem concerns would increase debt ratio.

Since we find more significant role of production flexibility on leverage in previous regression, we further split our observations into two time periods with the split point set at 2001 and report the results in Table 5.9 to observe the dynamic roles of production flexibility in different time periods. We split in half the span of the data period for which we have debt ratings available, which is from 1986 to 2017. Results for all firms, NYSE firms, and Nasdaq firms are reported in Panel A, Panel B, and Panel C respectively. We find that only firms that fall in financial constraint condition and in "years after 2001" time window show a significant relationship between production flexibility and leverage for all related firms and Nasdaq firms. This finding implies that the role of production flexibility has been increasing in recent years, which is consistent with our previous findings. Furthermore, after splitting observations by 2001, we find a higher effect of the interaction term than what we found before. For instance, in the all firms regression shown in Panel A, the coefficient between interaction term and leverage is -0.0115 for financial constraint firms after year 2001, while the coefficient between production flexibility and leverage found previous is around -0.007 (Table 5.8).

Table 5.9: Leverage on production flexibility, financial constraints and agency dummy with firms split by financial constraints and year splitting point

To observe the dynamic roles of production flexibility in different time periods, we further split our observations into two time periods with splitting point set at 2001. Thus, all firms would be categories into four classes shown in the four columns of each panel. Results for all firms, NYSE firms and Nasdaq firms are reported in Panel A, Panel B, and Panel C respectively. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

	Panel A: All Firms			
	(Years before 2001) (Financially Constrained)	(Years before 2001) (Not Constrained)	(Years after 2001) (Financially Constrained)	(Years after 2001) (Not Constrained)
FirmSize	0.0310***	0.03***	0.04***	0.01*
	(10.97)	(7.23)	(15.32)	(2.42)
MktBk	-0.000708	-0.004*	-0.002*	-0.001
	(-1.07)	(-2.13)	(-1.97)	(-0.51)
Tangibility	0.201***	-0.01	0.05*	0.12***
	(7.56)	(-0.55)	(2.53)	(6.39)
Profitability	-0.07***	-0.08*	-0.09***	-0.19***
	(-4.54)	(-2.07)	(-7.61)	(-9.01)
ProductionFlexibility	0.01	0.007	0.004	-0.002
	(0.63)	(0.19)	(1.56)	(-0.47)
NonAgencyDummy	0.0125	0.002	0.01*	-0.01
	(0.77)	(0.09)	(2.27)	(-1.76)
ProFlex_Agen	-0.02	-0.01	-0.0115**	0.002
	(-0.78)	(-0.14)	(-3.26)	(0.55)
Intercept	-0.06*	0.09*	-0.12***	0.22***
	(-2.54)	(2.06)	(-6.89)	(7.71)
N	4788	3273	8819	7189
Adjusted R^2	0.05	0.03	0.04	0.02
Firm Fixed Effect	yes	yes	yes	yes

	Panel B: NYSE Firms			
	(Years before 2001) (Financially Constrained)	(Years before 2001) (Not Constrained)	(Years after 2001) (Financially Constrained)	(Years after 2001) (Not Constrained)
FirmSize	0.04***	0.03***	0.05***	0.004
	(8.24)	(6.87)	(10.32)	(1.18)
MktBk	-0.004	-0.004	-0.004	-0.003
	(-1.63)	(-1.50)	(-1.91)	(-1.03)

Tangibility	0.19***	-0.05	0.05	0.12***
	(4.77)	(-1.68)	(1.73)	(5.51)
Profitability	-0.21***	-0.08	-0.12***	-0.206***
	(-4.83)	(-1.71)	(-4.32)	(-9.02)
ProductionFlexibility	0.01	0.01	-0.01	-0.01
	(0.18)	(0.27)	(-1.18)	(-1.57)
NonAgencyDummy	0.03	0.02	0.01*	-0.01
	(1.26)	(0.88)	(2.09)	(-1.59)
ProFlex_Agen	-0.02	-0.01	-0.005	0.01
	(-0.39)	(-0.22)	(-0.69)	(1.12)
Intercept	-0.10*	0.07	-0.16***	0.26***
	(-2.19)	(1.45)	(-4.55)	(8.01)
N	2179	2693	3114	5527
Adjusted R^2	0.08	0.04	0.06	0.03
Firm Fixed Effect	yes	yes	yes	yes

Panel C: Nasdaq Firms

	(Years before 2001) (Financially Constrained)	(Years before 2001) (Not Constrained)	(Years after 2001) (Financially Constrained)	(Years after 2001) (Not Constrained)
Firmsize	0.02***	0.02	0.04***	0.02**
	(4.54)	(1.46)	(11.98)	(2.87)
MktBk	0.0002	-0.004	-0.001	0.003
	(0.25)	(-1.01)	(-0.83)	(0.83)
Tangibility	0.23***	0.21*	0.06*	0.14**
	(5.56)	(2.25)	(2.07)	(2.83)
Profitability	-0.04	-0.26*	-0.07***	-0.12*
	(-1.77)	(-2.38)	(-4.80)	(-2.05)
ProductionFlexibility	0.01	0.60	0.01*	0.01
	(0.27)	(0.14)	(2.38)	(0.81)
NonAgencyDummy	0.001	0.38	0.01**	-0.01
	(0.06)	(0.11)	(3.26)	(-1.28)
ProFlex_Agen	-0.003	-0.59	-0.01***	0.001
	(-0.11)	(-0.13)	(-3.34)	(0.07)
Intercept	-0.02	-0.26	-0.14***	0.08
	(-0.56)	(-0.08)	(-6.51)	(1.15)
N	1986	364	5235	1446
Adjusted R^2	0.03	0.05	0.04	0.02

Firm Fixed Effect	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.4.2 Model Conclusion

We test the hypotheses developed in the theoretical section with empirical data in this section. The data covers non-financial companies listed in NYSE and Nasdaq from 1962 to 2017. Over decades, we observe increasing production flexibility along with decreasing leverage. Further investigation confirms the negative relationship between production flexibility and leverage in general as shown in Panel A of Table 5.10. The negative relationship is further strengthened when firms face financial constraints. This finding confirms our previous hypothesis that debtholders are unwilling to lend when there is financial constraint under the equity value optimization goal. However, the negative relationship would be mitigated when there exist agency problem concerns, supporting our hypothesis of the relationship between production flexibility and leverage under financial constraint and agency problem situation. This finding stands when we run regressions on firms only listed in Nadaq, but does not get statistically significant support for NYSE firms. The different findings in these two markets support that the roles of production flexibility, agency problem and financial constraint are more evident in Nadaq firms than NYSE firms.

The relationship between production flexibility and leverage under financial constraint situation satisfies our conclusion in the theoretical section as shown in Panel B of Table 5.10; however, in “No financial constraints” situation, we do not observe clear positive pattern between production flexibility and leverage as suggested by our theoretical studies. One possible reason is that when firms face no financial constraint, they can raise debt or equity based on market conditions even though there is no financial need from production activities and this does not satisfy our assumptions in the theoretical discussion about the relationship between production flexibility and leverage.

Table 5.10: relationship between production flexibility and leverage

The following table illustrates the relationship between production flexibility and leverage in empirical study in Panel A and the relationship between flexibility and leverage in Panel B in the theoretical section with the goal of equity value maximization. There are four different situations classified by two dimensions: whether a firm is financially constrained and whether it faces agency problems.

Panel A: Relationship between production flexibility and leverage in empirical study			
		Agency problem condition	
		No agency problem	Agency problem
financial constraint condition	No financial constraint	NULL	NULL
	financial constraint	Negative	Negativity mitigated/ positive
Panel B: Relationship between production flexibility and leverage in theoretical section with equity value maximization			
		Agency problem condition	
		No agency problem	Agency problem
financial constraint condition	No financial constraint	Positive	Positive
	financial constraint	Negative	Positive

5.5 Empirical Research: Test of theory to Understand Tail Firms

In this section, we extend the empirical study of investigating the role of production flexibility on leverage to leverage tails, including both left tails and right tails. Extreme low leverage and high leverage have been interesting topics in finance for decades; however, most of these studies evaluate possible impacts from financial factors while ignoring operation factors. In this section, we will examine the contribution of an operation factor, i.e., production flexibility, on the growing trend of left tails first.

5.5.1 Explanation of Left Tail Phenomenon

5.5.1.1 Univariate Analysis

Table 5.11 below shows how $AZTD$ and $AZTD\_C$ evolve over decades. We find that the $AZTD$ has been growing over decades in general from 0.14 in the 1960's to 0.3 in the 2010's, meaning the proportion of firms that have been adopting an almost zero-leverage policy has been growing from 14% to 30%. For $AZTD\_C$, the same increasing trend is observed for AZTD_C from the 1960's to the 2010's with an average value evolving from 0.162 to 0.321, meaning a growing percentage of firms, i.e., 26.2% to 32.1% are transitioning to almost 0 leverage firms.

Table 5.11: Descriptive statistics of low leverage by decades

This table describes the state of leverage for left tail variable $AZTD$ and the transition of not-left tail state to left tail state variable $AZTD\_C$ by decades. The descriptive statistics include number of observations, Mean, Standard deviation, Minimum and Maximum.

	AZTD					
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	6183	24308	22513	23371	31085	11613
Mean	0.14	0.12	0.16	0.25	0.35	0.3
Std.	0.35	0.32	0.37	0.44	0.48	0.46
Min	0	0	0	0	0	0
Max	1	1	1	1	1	1
	AZTD_C					
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	6183	24308	22513	23371	31085	11613
Mean	0.16	0.12	0.16	0.26	0.35	0.32
Std.	0.37	0.32	0.37	0.44	0.48	0.47
Min	0	0	0	0	0	0

End of Table

5.5.1.2 Regression Model Results

To explore the role of production flexibility and agency problem concerns on $AZTD$ and $AZTD$ transition decisions under different financial constraints conditions and agency problem concerns scenarios, we run similar regressions for all firms, NYSE firms, and Nasdaq firms in Table 5.12. Two measurements of financial constraints are tested. One is debt rating with results shown in Panel A, and the other is SA-Dummy with results shown in Panel B. The results are consistent with what we had in previous regressions. $ProFlex\_Agen$ is found to be positively related with $AZTD$ for firms facing financial constraints only In Panel A, meaning that the interaction between agency problem and production flexibility is found to flip the sign of impact that production flexibility plays on the almost zero-leverage decision. Thus, high production flexibility with high agency problem concerns leads to a relatively lower probability of almost zero-leverage decision, while high production flexibility with low agency problem concerns leads to a relatively higher probability of having an almost zero-leverage decision.

To further elaborate, for firms that are financially constrained, the relationship between production flexibility and the log odds of being in low-leverage state is around 0.10. In Panel B, a similar result is found for the interactions among production flexibility, financial constraint and agency problem concern. The results here are consistent with what we found in previous research.

Table 5.12: Regressions of low leverage on production flexibility, agency problem dummy by whether facing financial constraints

This table reports the result of panel data analysis on the sample from 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged databases from 1962 to 2017, with financial firms excluded. The table reports the regression AZTD firms on independent variables with different financial constraints measures. Panel A and Panel B exhibit regression results with financial constraints measured as debt rating and SA-Dummy, respectively. Independent variables include *ProductionFlexibility*, Agency Problem Dummy *NonAgencyDummy*, the interaction between the two *ProFlex_Agen*, and other control variables. All dependent variables are lagged behind by one year in the regressions. For Panel A, AZTD firms are the firms with the ratio of total debt to total asset not bigger than 5%. Regressions are run on both debt access constrained firms and debt access unconstrained firms for all firms (Column 1 and 2), NYSE firms (Column 3 and 4), and Nasdaq firms (Column 5 and 6). All 6 regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
Panel A: debt rating to separate firms						
FirmSize	-0.55***	-0.35*	-0.42***	-0.53**	-0.64***	-0.35
	(-11.01)	(-2.44)	(-4.51)	(-2.66)	(-9.82)	(-1.52)
MktBk	0.13***	0.07	0.25***	0.06	0.10***	0.03
	(5.87)	(1.37)	(4.69)	(1.01)	(3.85)	(0.39)
Tangibility	-3.84***	-5.30***	-2.32**	-4.26**	-4.42***	-6.50**
	(-8.70)	(-4.47)	(-3.25)	(-2.99)	(-7.22)	(-2.93)
Profitability	1.57***	0.90	1.75*	2.00	1.05**	0.91
	(5.51)	(0.86)	(2.43)	(1.37)	(3.08)	(0.55)
ProductionFlexibility	0.10	-0.26	0.43***	-0.04	0.03	-0.38
	(1.79)	(-1.62)	(3.74)	(-0.16)	(0.50)	(-1.68)
NonAgencyDummy	-0.1*	-0.20	-0.09	-0.34*	-0.17**	-0.04
	(-2.22)	(-1.82)	(-1.14)	(-2.42)	(-2.91)	(-0.22)
ProFlex_Agen	0.151**	0.18	0.06	0.28	0.19***	0.10
	(3.26)	(1.39)	(0.57)	(1.52)	(3.47)	(0.58)
N	7809	1538	2784	939	4455	558
ll	-3117.1	-533.5	-1090.8	-311.6	-1782.2	-204.1
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
Panel B: SA-Dummy to separate firms						
FirmSize	-0.56***	-0.35*	-0.42***	-0.53**	-0.64***	-0.34
	(-11.08)	(-2.44)	(-4.52)	(-2.68)	(-9.91)	(-1.49)
MktBk	0.13***	0.07	0.25***	0.06	0.10***	0.04
	(5.94)	(1.42)	(4.71)	(0.99)	(3.91)	(0.48)

Continuation of Table 5.12:						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
Tangibility	-3.86***	-5.27***	-2.35***	-4.22**	-4.44***	-6.49**
	(-8.75)	(-4.46)	(-3.29)	(-2.97)	(-7.25)	(-2.92)
Profitability	1.57***	0.81	1.78*	1.95	1.06**	0.75
	(5.53)	(0.77)	(2.46)	(1.35)	(3.10)	(0.46)
ProductionFlexibility	-0.06	-0.44	0.34*	-0.36	-0.15	-0.44
	(-0.80)	(-1.85)	(2.11)	(-1.09)	(-1.85)	(-1.33)
NonAgencyDummy	-0.20*	-0.39	-0.20	-0.72*	-0.34**	-0.03
	(-2.26)	(-1.83)	(-1.26)	(-2.54)	(-2.88)	(-0.08)
ProFlex_Agen	0.31***	0.36	0.14	0.59	0.38***	0.17
	(3.33)	(1.43)	(0.69)	(1.64)	(3.49)	(0.50)
N	7850	1545	2788	943	4492	561
ll	-3127.2	-536.1	-1093.5	-313.1	-1790.0	-205.2

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

The transition between different states of leverage level, i.e., from higher than 5% to almost zero, is also studied in Table 5.13. Similarly, two measurements for financial constraints are tested: one is debt rating with results shown in Panel A and the other is SA-Dummy with results shown in Panel B. Results of control variables gained in both Panel A and Panel B are consistent with what we have seen in previous regressions. In Panel A, $ProFlex\_Agen$ is found to be positively related with the probability of firms transiting from Non-AZTD firms to AZTD firms for both the all firm sample and Nasdaq firms. Furthermore, the interaction between agency problem and production flexibility is found to flip the sign of the impact that production flexibility play on the low leverage transition decision. High production flexibility with high agency problem concerns leads to a low probability of transition to a low-leverage states, while high production flexibility with low agency problem concerns leads to a relatively higher probability of transition to a low-leverage state. However, no significant effects are found in Panel B when financial constraint is measured using SA-Dummy.

Table 5.13: Low leverage change on production flexibility, financial constraints and agency dummy

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The table reports the regression of transition from Non-AZTD firms to AZTD firms on independent variables. Panel A and Panel B exhibit regression results with financial constraints measured as debt rating and SA-Dummy, respectively Independent variables for both panels include $ProductionFlexibility$, Agency Problem Dummy $NonAgencyDummy$, the interaction between the two $ProFlex\_Agen$ and other control variables. All dependent variables are lagged behind by one year in the regressions. The dependent variable $AZTD\_C$ is set to 1 for firms lowering down their leverage from above 5% to lower than 5% and is set to 0 otherwise. Regressions are run on both financially constrained firms and not financially constrained firms for all firms (Column 1 and 2), NYSE firms (Column 3 and 4), and Nasdaq firms (Column 5 and 6). All the 6 regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Panel A: debt rating to separate firms						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	-0.68***	-0.42*	-0.60***	-0.53*	-0.69***	-0.27
	(-7.17)	(-2.14)	(-3.44)	(-2.06)	(-5.61)	(-0.83)
MktBk	0.14**	0.07	0.23*	0.06	0.10	-0.07
	(2.78)	(1.00)	(2.20)	(0.77)	(1.55)	(-0.52)
Tangibility	-1.78*	-2.81	0.17	-1.09	-3.23**	-5.67*
	(-2.37)	(-1.92)	(0.14)	(-0.59)	(-2.98)	(-2.16)
Profitability	1.99**	0.62	2.81	1.41	1.26	2.24
	(3.10)	(0.43)	(1.69)	(0.72)	(1.59)	(0.85)
ProductionFlexibility	0.173	-0.16	0.45*	0.003	0.09	-0.38
	(1.62)	(-0.79)	(2.09)	(0.01)	(0.63)	(-1.23)
NonAgencyDummy	-0.33***	-0.13	-0.21	-0.25	-0.42***	0.07
	(-3.75)	(-0.89)	(-1.38)	(-1.37)	(-3.61)	(0.25)
ProFlex_Agen	0.21*	-0.08	-0.02	0.09	0.32**	-0.25
	(2.23)	(-0.51)	(-0.08)	(0.39)	(2.74)	(-1.04)
N	2925	1076	1195	692	1551	360
ll	-857.4	-280.2	-322.4	-174.2	-471.9	-94.11
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
Panel B: SA-Dummy to separate firms						
FirmSize	-0.57***	-0.68***	-0.48*	-0.66***	-0.56***	-0.81***
	(-4.90)	(-5.60)	(-2.31)	(-3.53)	(-3.65)	(-4.48)
MktBk	-0.01	0.27***	-0.02	0.40***	-0.04	0.14

Continuation of Table 5.13:						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(-0.15)	(3.90)	(-0.17)	(3.44)	(-0.48)	(1.19)
Tangibility	-0.03	-3.01**	0.95	-1.28	-0.47	-5.28***
	(-0.02)	(-3.28)	(0.47)	(-0.95)	(-0.22)	(-3.48)
Profitability	1.07	1.21	4.44	-1.49	0.71	2.96*
	(1.19)	(1.36)	(1.93)	(-0.92)	(0.67)	(2.06)
ProductionFlexibility	-0.08	0.33	0.36	0.63	-0.29	0.07
	(-0.51)	(1.46)	(1.02)	(1.60)	(-1.51)	(0.24)
NonAgencyDummy	-0.35	-0.35	-0.05	-0.34	-0.63*	-0.51
	(-1.65)	(-1.54)	(-0.14)	(-0.94)	(-2.15)	(-1.54)
ProFlex_Agen	0.17	-0.36	0.11	-0.64	0.28	-0.04
	(0.86)	(-1.18)	(0.26)	(-1.32)	(1.11)	(-0.10)
N	1819	1711	861	814	888	796
ll	-535.5	-462.1	-227.5	-211.4	-277.1	-212.6
End of Table						

To summarize, the mechanism of production flexibility, financial constraints and agency problem concerns functions similarly in both a state of low-leverage decision and in the transition of not-low leverage to low leverage decision. Whether there are financial constraints determines whether production flexibility plays a significant role. When there are no financial constraints, production flexibility plays no significant roles in both Low-leverage decision and Transition from non Low-leverage to Low-leverage decision regardless whether there is an agency problem concern. However, when there are financial constraints, the existence of agency problem concerns determines how production flexibility affects low-leverage (transition). Agency problem concerns prevent firms from adopting low-leverage decisions or transiting from Not-low-leverage firms to Low-leverage firms.

5.5.2 Explanation of Right Tail Phenomenon

5.5.2.1 Univariate Analysis

We first show how the right tail measurements HTD and HTD_C evolves over decades in Table 5.14 (see below). We find that the HTD has been slightly growing over decades in general from 0.199 in 1960s to 0.219 in 2010s, meaning that a slightly increasing proportion, i.e., 2%, of all firms, has been adopting a high leverage policy. For HTD_C, fluctuating trends are found.

Table 5.14: Descriptive statistics of high leverage by decades

This table illustrates descriptive statistics of the revolution of the right tail state HTD and right tail transition HTD_C by decades. These statistics cover number of observations, Mean, Standard deviation, Minimum and Maximum.

HTD						
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	6183	24308	22513	23371	31085	11613
Mean	0.20	0.26	0.28	0.25	0.18	0.22
Std.	0.40	0.44	0.45	0.43	0.39	0.41
Min	0	0	0	0	0	0
Max	1	1	1	1	1	1
HTD_C						
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Obs	5095	17937	16611	17819	25259	9340
Mean	0.07	0.06	0.09	0.08	0.05	0.06
Std.	0.25	0.24	0.29	0.28	0.22	0.24
Min	0	0	0	0	0	0
Max	1	1	1	1	1	1

End of Table

5.5.2.2 Regression Model Results

The exploration of production flexibility on High-leverage decision is reported in Table 5.15. Two measurements of financial constraints are tested. One is debt rating with results shown in Panel A, and the other is SA-Dummy with results shown in Panel B. Results of control variables gained in both Panel A and Panel B are consistent with what we have in previous regressions. In Panel A, $ProFlex\_Agen$ is found to be negatively related with the probability of firms being in High-Leverage firms state for both all firm sample and Nasdaq firms. Furthermore, the interaction between agency problem and production flexibility is found to flip the sign of the impact that production flexibility plays on the High-Leverage decision. High production flexibility with high agency problem concerns leads to a lower probability of being in High-Leverage state, while high production flexibility with low agency problem concerns would lead to a relatively higher probability of being in High-Leverage state. Noticing that for NYSE-only firms, no statistical significance is found for constrained sub-market and the magnitude is quite small. At the same time, for unconstrained ones, the interaction term is found to be positively significant with regards to leverage.

One possible explanation to the disappearing significance is that the debt raising cost of firms listed in NYSE is less costly that that of firms listed in Nasdaq market while the bankruptcy cost is higher in NYSE than it is in Nasdaq. In Panel B, when financial constraint is measured using SA-Dummy, no significant effects for Nasdaq firms are found.

Table 5.15: High leverage on production flexibility, financial constraints and agency dummy

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The table reports the regression of HTD firms on independent variables. Panel A exhibits regression results with financial constraints measured by debt rating and Panel B show results with financial constraints measured by SA-Dummy. Independent variables for both panels include $ProductionFlexibility$, Agency Problem Dummy $NonAgencyDummy$, the interaction between the two $ProFlex\_Agen$ and other control variables. All independent variables are lagged behind by one year in the regressions. High-leveraged firms are the firms with leverage higher than the 75% of the firms in the sample, with the threshold being 0.367 in our case. Thus, the dependent variable HTD is 1 with firms holding leverage higher than 0.33595. For each panel, regressions are run on both financially constrained firms and not financially constrained firms for all firms (Column 1 and 2), NYSE firms (Column 3 and 4), and Nasdaq firms (Column 5 and 6). All 6 regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Panel A: financial constraint by debt rating						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	0.57***	-0.09	0.56***	-0.13	0.50***	0.20
	(7.98)	(-1.31)	(4.52)	(-1.64)	(4.89)	(1.05)
MktBk	-0.03	0.03	-0.13	0.01	0.02	0.10
	(-1.04)	(0.53)	(-1.61)	(0.22)	(0.41)	(1.05)
Tangibility	2.90***	-0.09	3.11***	-0.04	3.29***	-1.34
	(5.11)	(-0.20)	(3.43)	(-0.09)	(3.63)	(-1.11)
Profitability	-1.87***	-4.05***	-1.05	-4.83***	-2.17***	-1.60
	(-4.22)	(-6.20)	(-0.97)	(-6.34)	(-3.31)	(-1.09)
ProductionFlexibility	0.13	-0.04	0.07	-0.05	0.18	-0.06
	(1.61)	(-0.61)	(0.48)	(-0.58)	(1.72)	(-0.37)
NonAgencyDummy	0.22***	-0.02	0.25**	0.02	0.30**	-0.095
	(3.46)	(-0.37)	(2.59)	(0.37)	(3.19)	(-0.77)
ProFlex_Agen	-0.25***	0.12	-0.11	0.16*	-0.34***	0.01
	(-3.49)	(1.85)	(-0.82)	(2.14)	(-3.46)	(0.07)
N	4398	6368	1955	5194	1958	964
ll	-1576.6	-2533.6	-725.0	-2081.8	-668.0	-366.5
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
Panel B: financial constraint by SA-Dummy						
FirmSize	0.31***	0.25***	0.07	0.13	0.29*	0.46***
	(3.79)	(4.40)	(0.59)	(1.88)	(2.44)	(3.82)
MktBk	0.09*	-0.07*	0.04	-0.10	0.11*	0.07

Continuation of Table B.12:						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(2.35)	(-2.00)	(0.56)	(-1.21)	(2.09)	(0.96)
Tangibility	-2.16*	1.26**	-2.19	0.99*	-3.29	1.62
	(-2.22)	(3.26)	(-1.76)	(2.18)	(-1.81)	(1.79)
Profitability	-2.04***	-2.***	-5.14***	-2.45**	-0.45	-3.36**
	(-3.41)	(-4.98)	(-4.32)	(-3.22)	(-0.58)	(-3.16)
ProductionFlexibility	0.11	-0.08	-0.006	-0.21	0.09	0.19
	(0.93)	(-0.71)	(-0.02)	(-1.43)	(0.61)	(0.93)
NonAgencyDummy	0.03	0.04	0.05	0.08	-0.03	0.24
	(0.19)	(0.44)	(0.27)	(0.71)	(-0.12)	(1.17)
ProFlex_Agen	-0.38*	0.20	-0.31	0.37*	-0.42	-0.22
	(-2.50)	(1.50)	(-1.27)	(2.24)	(-1.89)	(-0.78)
N	3298	6411	1962	4456	1156	1459
ll	-1147.4	-2536.9	-706.2	-1808.4	-379.3	-525.0

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

In Table 5.16, the role of production flexibility as well as agency problem concerns on the transition from Not-high-Leverage firms to High-Leverage firms is reported. Two measurements for financial constraints are tested. One is debt rating with results shown in Panel A, and the other is SA-Dummy with results shown in Panel B. Results of control variables gained in both Panel A and Panel B are consistent with what we have in previous regressions. In Panel A, $ProFlex\_Agen$ is found to be negatively related with the probability of firms transitioning from Not-high-Leverage firms to High-Leverage firms for both the full sample and Nasdaq firms. Furthermore, the interaction between agency problem and production flexibility is found to flip the sign of the impact that production flexibility plays on the transition. High production flexibility with high agency problem concerns tends to lower down the probability of High-Leverage state transition, but high production flexibility with low agency problem concerns would lead to relatively higher probability of High-Leverage state transition. However, In

Panel B, when financial constraint is measured by SA-Dummy, no significant effects for Nasdaq firms are found.

Table 5.16: High leverage change on production flexibility, financial constraints and agency dummy

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The table reports the regression of transition from Not-HTD firms to HTD firms on independent variables with financial constraints measured by debt rating, while Panel B reports the regression of transition from Not-high-leveraged firms to High-leveraged firms on the same set of independent variables with financial constraints measured by SA-Dummy. Independent variables for both panels include $ProductionFlexibility$, Agency Problem Dummy $NonAgencyDummy$, the interaction between the two $ProFlex\_Agen$ and other control variables. All dependent variables are lagged behind by one year in the regressions. The dependent variable HTD_C is 1 if firms increasing their leverage from lower than 0.367 to higher than 0.367. Regressions are run on both debt access constrained firms and debt access not constrained firms for all firms (Column 1 and 2), NYSE firms (Column 3 and 4), and Nasdaq firms (Column 5 and 6). All the 6 regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Panel A: financial constraint by Debt Rating

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	0.39***	-0.05	0.41*	-0.04	0.28*	-0.13
	(3.95)	(-0.42)	(2.32)	(-0.34)	(2.03)	(-0.40)
MktBk	0.04	0.07	-0.09	0.06	0.11*	0.12
	(1.34)	(1.07)	(-0.86)	(0.76)	(2.40)	(1.01)
Tangibility	2.07*	-0.32	1.32	-0.21	2.64*	-1.33
	(2.54)	(-0.39)	(0.99)	(-0.23)	(2.09)	(-0.61)
Profitability	-0.42	-0.25	1.89	-0.04	-1.00	-1.12
	(-0.76)	(-0.23)	(1.20)	(-0.03)	(-1.22)	(-0.48)
ProductionFlexibility	0.17	0.07	-0.05	-0.004	0.34*	0.28
	(1.46)	(0.53)	(-0.23)	(-0.03)	(2.18)	(0.82)
NonAgencyDummy	0.09	-0.101	0.0148	-0.121	0.25	-0.07
	(0.99)	(-1.19)	(0.11)	(-1.25)	(1.75)	(-0.33)
ProFlex_Agen	-0.33**	0.08	-0.16	0.14	-0.52***	0.12
	(-3.17)	(0.70)	(-0.91)	(1.00)	(-3.67)	(0.38)
N	2817	2936	1231	2378	1297	463
ll	-793.7	-891.2	-356.7	-715.6	-344.5	-142.7
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: financial constraint by SA-Dummy

FirmSize	0.14	0.20	0.04	0.19	0.04	0.29

Continuation of Table 5.16:						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(1.19)	(1.65)	(0.18)	(1.23)	(0.23)	(1.13)
MktBk	0.20***	-0.03	0.05	-0.08	0.28***	0.09
	(3.99)	(-0.44)	(0.50)	(-0.60)	(3.87)	(0.97)
Tangibility	0.53	1.26	2.14	0.76	-5.57	4.18*
	(0.36)	(1.36)	(1.03)	(0.67)	(-1.96)	(2.12)
Profitability	-0.76	0.21	0.24	1.22	-0.31	-1.37
	(-0.86)	(0.25)	(0.12)	(0.94)	(-0.26)	(-0.88)
ProductionFlexibility	0.18	0.51*	-0.10	0.14	0.25	1.39*
	(0.81)	(2.10)	(-0.23)	(0.47)	(0.93)	(2.34)
NonAgencyDummy	-0.176	-0.01	-0.41	-0.22	0.06	0.79
	(-0.70)	(-0.05)	(-1.17)	(-0.85)	(0.13)	(1.74)
ProFlex_Agen	-0.45	-0.39	-0.19	0.07	-0.64	-1.45*
	(-1.79)	(-1.36)	(-0.42)	(0.19)	(-1.91)	(-2.08)
N	1856	1914	1007	1244	722	531
ll	-491.4	-548.3	-269.4	-361.6	-182.5	-137.5

End of Table

5.6 Robustness Test

5.6.1 New Measure of Production Flexibility

To further test our findings, we alter the measure of production flexibility from the perspectives of outside investors. The new production flexibility is measured as the ratio of investors' surprise in profitability to that in sales. The surprise here is the difference between actual data reported in the financial statements and estimated data from outsiders. The identifying assumption of this new method is that investor‘'s surprise is exogenous and could be a good proxy. The surprise data comes from Institutional Brokers' Estimate System (I/B/E/S database), which is a database from Thomson Reuters. The database contains analysts' estimates of 20 forecast measures. The new measurement is calculated below:

$$\text{PF\_Supply}=\frac{SurpriseofGrossMarginRate}{\frac{SurpriseofSales}{TotalAssets}}$$

where “Surprise of Gross Margin Rate”, “Surprise of Sales” can be directly gained from ”Surprise Mean” field (raw name is “surpmean”) from the database. Since the magnitude of the sales surprise is affected by the firm size, we modify the sales surprise by total assets to make comparisons among companies consistent. This new measure identifies the effect of unexpected sale fluctuations on profitability surprise from the perspective of investors, In other words, this measurement is similar to the previous measurement, but standing on a different point, i.e., from investors instead of companies. Corresponding to this change of production flexibility measurement, the interaction between production flexibility and non-agency problem concerns are also changed to:

$$ProFlex\_Agen{=}PF\_Supply{*}NonAgencyDummy$$

Explorations of how financial constraints affect the relationship between production flexibility and leverage is explored in Table 5.17. Though no significance is found, the new production flexibility proxy is marginally significant for financial constraint segment. For not financial constraint section, the production flexibility measured in surprise is not significant.

Table 5.17: Leverage on production flexibility by financial constraints

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The table reports the regression of leverage on independent variables. Independent variables for both panels include PF_Supply and other control variables. All independent variables are lagged behind by one year in the regressions. The dependent variable is *Leverage* iRegressions are run on both financially constrained firms and not financially constrained firms for all firms. The regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

	(1)	(2)
	Constrained	Not Constrained
FirmSize	0.03***	0.01***
	(17.29)	(4.47)
MktBk	0.0001	-0.002
	(1.43)	(-1.31)
Tangibility	0.14***	0.05**
	(9.15)	(2.93)
Profitability	-0.07***	-0.14***
	(-8.37)	(-6.41)
PF_Supply	-0.00002	-0.00005
	(-1.69)	(-0.20)
constant	-0.04***	0.23***
	(-4.25)	(10.05)
N	15100	7302
df_b	5	5
r2_w	0.0310	0.0139

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We also explore the Low-Leverage state, Low-Leverage shifting, High-Leverage, and High-Leverage shifting in Table 5.18 with results shown in Panel A to Panel D respectively. In Low-Leverage regression for Nasdaq firms in Panel A, significant roles are found for the interaction of production flexibility and agency problem concerns on the dependent variables, which is consistent with our findings in previous section. No significance is found in Panel B to Panel D, even though the signs of key variables are consistent.

Table 5.18: Low (High)-leverage state (shift) on production flexibility and agency problem concerns by financial constraints

This table reports the result of panel data analysis on the sample over 1962 to 2017 by decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms excluded. The table reports the regression of leverage on independent variables. Independent variables for both panels include PF_Supply, $ProFlex\_Agen$ and other control variables. All dependent variables are lagged behind by one year in the regressions. The dependent variable is $Low\_Leverage$ in Panel A, $Low\_LeverageShift$ in Panel B, $High\_Leverage$ in Panel C, $High\_LeverageState$ in Panel D. Regressions are run on both debt access constrained firms and debt access unconstrained firms for all firms. The regressions are estimated with firm effects fixed. Coefficients marked with ***,**, and * are significant at 0.1%, 1%, and 5% respectively.

Panel A: Low-Leverage State						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	-0.73***	-0.04	-0.66***	-0.07	-0.74***	-0.33
	(-8.44)	(-0.17)	(-3.70)	(-0.21)	(-7.13)	(-0.78)
MktBk	0.14***	0.11	0.21*	0.14	0.15***	-0.002
	(4.32)	(1.74)	(2.55)	(1.68)	(3.78)	(-0.02)
Tangibility	-3.51***	-10.67***	-1.67	-9.17**	-3.62***	-16.25***
	(-5.15)	(-4.51)	(-1.33)	(-3.00)	(-4.29)	(-3.79)
Profitability	2.13***	1.56	3.16*	2.38	1.45*	1.59
	(4.39)	(1.00)	(2.52)	(1.02)	(2.49)	(0.72)
PF_Supply	0.02	20.70	0.69	36.13*	0.01	-4.85
	(1.50)	(1.89)	(0.68)	(2.16)	(1.11)	(-0.28)
NonAgencyDummy	-0.13	0.55	-0.41*	0.31	-0.10	0.81
	(-1.15)	(1.76)	(-1.97)	(0.77)	(-0.68)	(1.47)
ProFlex_Agen	-0.43	-35.64*	-0.13	-36.06	-0.75*	-31.56
	(-1.22)	(-2.35)	(-0.11)	(-1.86)	(-2.13)	(-0.69)
N	3928	720	1252	392	2542	328
ll	-1508.9	-249.8	-468.0	-140.7	-984.6	-103.3

Panel B: Low-Leverage Shift						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	-0.86***	-0.06	-1.24***	0.05	-0.***	-0.34
	(-5.14)	(-0.18)	(-3.45)	(0.11)	(-3.55)	(-0.51)
MktBk	0.14*	0.05	0.07	0.11	0.17*	-0.18
	(2.03)	(0.62)	(0.50)	(1.17)	(2.05)	(-1.07)
Tangibility	-2.17	-5.48*	-0.41	-1.71	-3.10*	-11.70*

	(-1.72)	(-1.97)	(-0.16)	(-0.48)	(-2.04)	(-2.33)
Profitability	2.25*	1.65	3.40	2.51	1.48	0.19
	(2.29)	(0.74)	(1.24)	(0.84)	(1.32)	(0.05)
profle_sup_rs_training	-0.03	18.33	1.58	15.87	-0.03	19.83
	(-0.38)	(1.38)	(0.83)	(1.09)	(-0.45)	(0.74)
PF_Supply	-0.65**	0.44	-0.96*	0.43	-0.45	0.47
	(-2.81)	(1.08)	(-2.38)	(0.87)	(-1.50)	(0.61)
ProFlex_Agen	0.40	-35.42	-0.82	-31.19	-1.85	-34.51
	(1.02)	(-1.16)	(-0.42)	(-0.98)	(-0.88)	(-0.51)
N	1341	467	510	262	806	205
ll	-392.5	-130.8	-132.7	-74.94	-247.5	-52.94

Panel C: High-leverage						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	0.44***	-0.0004	0.74**	-0.01	0.22	0.26
	(3.33)	(-0.00)	(2.92)	(-0.06)	(1.32)	(0.75)
MktBk	0.004	0.07	-0.03	0.04	-0.04	0.35
	(0.07)	(0.94)	(-0.21)	(0.44)	(-0.60)	(1.70)
Tangibility	1.59	1.00	1.20	2.60**	1.89	-5.40*
	(1.45)	(1.16)	(0.65)	(2.59)	(1.32)	(-2.39)
Profitability	-2.07**	-2.15*	-0.70	-2.34*	-1.70	-1.55
	(-2.66)	(-2.10)	(-0.36)	(-1.99)	(-1.84)	(-0.68)
PF_Supply	-0.01	-5.40	4.04	1.54	-0.01	-9.05
	(-1.31)	(-1.34)	(1.55)	(0.21)	(-1.25)	(-1.57)
NonAgencyDummy	0.08	-0.15	0.38	-0.04	0.01	-0.29
	(0.46)	(-1.00)	(1.35)	(-0.20)	(0.05)	(-0.73)
ProFlex_Agen	-1.65	5.25	-3.94	-1.62	-2.35	13.62
	(-1.09)	(1.29)	(-1.01)	(-0.22)	(-1.13)	(1.26)
N	1589	2225	696	1818	806	378
ll	-556.3	-870.7	-251.5	-720.3	-276.8	-130.6

Panel D: High Leverage Shift						
	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	0.03	-0.32	0.79	-0.21	-0.25	-1.00
	(0.18)	(-1.12)	(1.95)	(-0.63)	(-1.05)	(-1.45)
MktBk	0.13	-0.09	0.15	-0.003	0.07	-0.64

	(1.80)	(-0.50)	(1.05)	(-0.02)	(0.78)	(-1.18)
Tangibility	-0.76	3.60	-0.11	4.31	-1.75	1.75
	(-0.46)	(1.52)	(-0.04)	(1.55)	(-0.74)	(0.32)
Profitability	0.15	3.70	2.15	2.98	0.95	9.13
	(0.12)	(1.56)	(0.78)	(1.22)	(0.61)	(1.42)
PF_Supply	-0.96	-8.24	-11.67	-1.46	-0.346	-30.15
	(-0.81)	(-0.82)	(-1.75)	(-0.09)	(-0.30)	(-0.91)
NonAgencyDummy	-0.48	-0.65*	-0.87	-0.78	-0.12	-0.93
	(-1.70)	(-1.96)	(-1.91)	(-1.91)	(-0.28)	(-1.15)
ProFlex_Agen	-16.85	8.06	-10.80	1.37	-15.49	12.06
	(-1.25)	(0.79)	(-0.50)	(0.08)	(-0.91)	(0.26)
N	926	665	417	531	474	128
ll	-245.7	-182.1	-107.5	-145.2	-122.6	-32.39

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.6.2 Robustness Tests: Panel Data Results with More Control Variables

To explore the robustness of the results regarding the role of production flexibility on leverage under different circumstances, more control variables are adopted in the regression to test the stability of our results. These variables include both firm characteristics variables and macro-economy variables that have been proven to play important roles on firm leverage decisions.

Before running the robustness regression, correlation analysis among variables are conducted first with the results shown in Appendix Table B.9. Variables that have correlations higher than 0.4 or lower than -0.4 are defined as highly correlated variables. $Capex$ is highly related with $StockMarketRet$ and $MacroProfit$, and $Capex$ is removed since it is less related with $Lever_{Tda}$ in both cases. Additionally, $MacroPro$ is highly correlated with $Growth$, and "TermSprd", and $Growth$ is kept since it has higher correlation with $Lever\_Tda$. After removing these highly correlated variables, remaining variables are brought into the robustness regression tests. Appendix Table B.10 to Appendix Table B.14 report the regression results of leverage on production flexibility and agency problem concerns with a full list of firm characteristics and macroeconomic condition variables controlled for all firms. Most of the results stay the same as they were in the previous regression results.

In Appendix Table B.10 , $ProFlex\_Agen$ is found to be significantly and negatively related with leverage, working to flip the sign of the relationship between production flexibility and leverage. The impact is positive for low agency problem concerns firms but is negative for high agency problem concerns firms. A similar scenario is found for Nasdaq firms. For Low-Leverage decision that is shown in Appendix Table B.11, a marginally significant role of production flexibility is found for all the firms while significant impact is found for Nasdaq-only firms. For firms in the sample, the impact is negative for high agency problem concerns firms, but positive for firms with low agency problem concerns. The same results are also applicable to Nasdaq firms. Appendix Table B.12 reports the financial decision for High-Leverage, where for financially constrained firms, the role of the interaction team is found to flip the relationship between production flexibility and high leverage decision, especially for Nasdaq-only firms. The impact is negative for high agency problem concerns firms but is positive for low agency problem concerns firms. The same positive relationship in the interaction term is found to hold for NYSE firms, which is consistent with what we found before. The same explanation regarding market differences in NYSE firms and Nasdaq firms might help to explain the results we have here.

Consistent results are found for the Low-leverage transition regression in Appendix Table B.13 and high-leverage transition regression in Appendix Table B.14. The interaction between production flexibility and agency problem works to flip the sign of the impact production flexibility exerts on both the low-leverage transition decision and high-leverage transition decision, especially for Nasdaq-only firms.

Thus, our findings still hold true even after introducing more control variables into regressions in the robustness test.

5.7 Conclusion

Motivated by the increasing proportion of zero/low-leverage firms over decades, we explore the evolution of leverage distribution of real industry within a time window of 60 years. The average leverage has been found to be stable over decades. Further exploration reveals that significant growing left tails over decades and the slightly increasing right tails in recent decades. This pattern is found in both NYSE firms and Nasdaq firms. The difference in the leverage evolution pattern between these two exchange markets is that NYSE firms show a clear increasing trend in right tails while Nasdaq firms show a clear increasing trend in left tails. A factor, namely production flexibility, is proposed to explain the growing left tail and right tail phenomenon based on literature research. Production flexibility has proven to play an important role on leverage but the role differs among researchers. Our study sets up a theoretical framework where the disagreement is unified.

The theoretical model illustrates that the role of flexibility on leverage differs under different conditions. The relationship between production flexibility and leverage varies with regard to whether there exist financial constraints, whether there are agency problem concerns and the goal of managers in whether to maximize firm value or equity value. When firms try to maximize equity value, production flexibility is positively related with leverage if firms do not face any financial constraints. The reason is that high production flexibility in general helps mitigate the distress cost brought by debt, leading to a positive relationship between production flexibility and leverage. However, when there exist financial constraints but no agency problem concerns, the relationship turns negative since when debt is only the funds source, higher production flexibility enables firms to borrow less from external sources. The story changes when agency problem concerns exist. In this scenario, high production flexibility exacerbates agency problem, results in low equity value and leads to high leverage.

The relationship between production flexibility and leverage changes when firms aim to maximize firm value. when firms do not face financial constraints, firms can borrow from external equity sources instead of debts to satisfy operation needs especially when production flexibility is high, leading to a negative relationship between production flexibility and leverage. However, when there exist financial constraints, firms can only resort to debt for funds raising and high production flexibility makes firms more attractive to be a good investment for debts lenders, leading to a positive relationship between production flexibility and leverage.

The logic in the theoretical model under equity value maximization is further confirmed in the empirical results. Using data coming from the merged annual Compustat and Center for Research in Security Prices (CRSP) data set over the period 1962–2017 with financial companies excluded, production flexibility is shown to help lower leverage in general. Further segmentation based on financial constraints proves that production flexibility plays an important role in leverage only when the access to financial market is constrained. This is partly inconsistent with our theoretical exploration regarding the roles of production flexibility and agency problem concerns when there is no financial constraint. This inconsistent empirical finding could be because firms could raise more funds than needed in a good market condition according to market timing theory when there is no financial constraint.

When there is financial constraint, the concern of agency problem helps to explain the different role of production flexibility on leverage. When that concern is higher, production flexibility is positively related with leverage, meaning that higher production flexibility leads to higher leverage. On the other hand, when the concern is low, production flexibility is negatively related with leverage, meaning that higher production flexibility leads to high firm value, thus lower leverage. These findings are more evident in Nasdaq firms and in recent 20 years. The same results are found for low-leverage decisions and low-leverage transition decisions. The role of production flexibility on high-leverage decisions and high-leverage transition decisions are reversed. Even after adopting a different proxy of production flexibility and bringing in more control variables, these findings still hold.

Part II

MACHINE LEARNING EXPLORATION IN PANEL DATA

Chapter 6

INTRODUCTION AND MOTIVATION

In Part 2, we will focus on model prediction performance instead of explanatory variables since predicting leverage appropriately is also important. In Part 1, we explored the role of production flexibility on leverage decisions via economic reasoning by theoretical exploration and empirical experiment, wherein we paid more attention to the explanatory contributions of independent variables instead of the prediction power of the model itself. However, predicting leverage correctly is important for both investors and regulators. Well-predicted leverage helps investors and regulatory institutions to estimate the default risk of firms more precisely, which is an important criterion for investors to decide about further investments and for regulatory institutions to make regulatory actions in controlling the risks of the financial system.

However, traditional panel data analysis techniques have been developed and leveraged for explanations and thus bear some limitations in prediction. The traditional assumption of linearity between independent variables and dependent variables limits the prediction power of traditional panel data analysis models. Furthermore, introducing a dummy variable for each individual section when handling fixed effects seems reasonable, but is not efficient or necessary when some sections have very tiny differences in their fixed effects, and thus higher model performance could be achieved with fewer dummy variables.

A modified algorithm called PanelGBM is designed to tackle non-linearity for better prediction power and to group individual sections efficiently. PanelGBM integrates dummy variables that have learned to group different sections based on their fixed effects with existing machine learning methods. The machine learning method utilized in this paper is Gradient Boosting Machine (GBM for short), which relaxes the strict linearity assumptions the traditional panel data analysis method makes. The dummy variables are derived from residuals of the machine learning process and introduced sequentially to learn fixed effects of different sections. Learning fixed effects by introducing new dummy variables sequentially bears the following advantages: 1) assumption-free in the distribution of fixed effects; 2) far fewer dummy variables are introduced; 3) being able to handle fixed effects caused by exogeneity other than known independent variables.

PanelGBM also avoids some shortcomings of existing machine learning methods applied in panel data studies. Current machine learning explorations in handling panel data can be categorized in three ways: 1) assuming that individual effects are driven by a latent factor that follows a certain distribution, and the estimation is done via expectation maximization (EM for short) (Hajjem et al., 2011, 2014); 2) looping between panel regression and machine learning tools until certain stopping criteria are met (Miller et al., 2014); 3) grouping observations using independent variables by clustering and running predictive machine learning tools separately for each group (Verma et al., 2018). All these three categories hold some assumptions in the learning process, which are relaxed in the PanelGBM.

Furthermore, this algorithm is flexible in the way that the underlying machine learning algorithm can be chosen based on specific problems and data. In our dissertation, we use GBM as the underlying machine learning to handle non-linearity. However, GBM could be replaced with other machine learning methods depending on the characteristics of data and research topics. For instance, for time-series data that illustrates temporal connections, recurrent neural networks can be chosen and integrated with the sequential dummy variables to explore both temporal dynamic behavior and specific individual effects.

PanelGBM has been proven to perform well in prediction and section grouping after being tested in simulation data and real empirical data. In simulation data, we find that PanelGBM outperforms existing methods in prediction performance, model stability, and grouping efficiency. Furthermore, PanelGBM works as plain GBM when no fixed effects exist but it works much better when the fixed effects are important or non-homogeneous among different subjects. The algorithm also works well in empirical data in model performance. What's more, the dummy variables generated from the algorithm bear high variable importance in prediction and help improve model performance when introduced into linear regressions.

Finally, PanelGBM has wide applications in many financial and business topics. One use case is in finance research when dealing with panel data of short length but broad sections. Moreover, it can be used in business practice. For instance, when analyzing or predicting a bunch of customers' behaviors when they are not on record for a long time but there is a large sample size, this algorithm can be applied. One detailed example can be analyzing customers' credit/fraud risk when they do not have a long record history, meaning that they have not shown a long traceable history to use. In this case, we need to draw from the patterns shared among different users to extract commonalities while admitting the difference among different users by introducing dummy variables to group them under different categories based on their heterogeneity .

The rest of this chapter is organized as follows. The literature review chapter summarizes the traditional panel data analysis mechanism, existing machine learning methods that have been proposed in panel data application and the logic underlying the modified algorithm. The next chapter then illustrates the model performance in a simulation data set compared to other machine learning methods in model performance as well as its efficiency in grouping. The simulation study explore various situations of panel data to test the applicability of PanelGBM. Following the performance test in simulation data, the algorithm is further tested in the real finance data, i.e., the leverage data we used in our finance study section. We conclude in Chapter 10.

Chapter 7

LITERATURE REVIEW

Parametric models, especially linear models, are leveraged heavily in social science studies for explanation purposes (Burnham and Anderson, 2003). The traditional parametric model most commonly used in economic and finance research is linear regression for its simplicity and interpretability. This is partly because finance research focuses more on outcomes and these outcomes can be explained by predictors (Tukey, 1969). The basic form of the linear regression model is:

$$Y_i = \alpha + X^{'}\beta + \mu_i (i = 1, 2, ..., N),$$

where μ is an error term that represents white noise. The traditional panel data methods utilize linear models as well. For panel data, extra terms that represent individual specific fixed effects or random effects are introduced to form the basic structure of panel data analysis models. In finance, observations are often categorized into different groups heuristically based on business understanding or some facts such as companies, industries, and locations in regions, etc. These clustering variables are either assumed to be random, meaning the effect is a random sample out of a full population, or to be fixed (non-random). Different assumptions of these clustering variables result in different structures of panel data models (Laird and Ware, 1982).

Panel data method can compensate for the lack of time depth by information inherited in the cross sections. When different sections in the sample are homogeneous, the panel data is pooled. Otherwise, a heterogeneous effect, either fixed or random effect, is introduced into the panel model. Simply put, for the equation shown below:

$$Y_{it} = \alpha + X_{it}^{'}\beta + \mu_{it} (i = 1, 2, ..., N; t = 1, 2, ..., T),$$

with $\mu_{it} = \mu_i + \lambda_t + \upsilon_{it}$. μ_i and λ_t represent sectional fix effect and time fix effect. υ_{it} is the model residual that follows an independent and identically distributed normal distribution, i.e., N(0,σ^2). When $\mu_i = \lambda_t = 0$, the model utilized the panel data as pooled panel data where there are no section-specific or time-specific differences. Further different non-zero assumptions about μ_i *or* λ_t lead to different approaches, i.e., fixed effects or random effects in estimating different coefficients. If either μ_i *or* λ_t is a non-zero constant that indicates heterogeneity in either cross-section level or time level, then the models will be called fixed effect models. Generally, if correlations between unobserved effects and observed explanatory variables are allowed, this approach is usually adopted. However, if either μ_i *or* λ_t is assumed to follow a certain distribution, the model becomes a random effect model. The random effect approach is generally used when drawing randomly from a large population. Finally, mixing both types in one model would generate a mixed effect model (Wooldridge, 2010).

For the fixed effect model, the coefficients of the model can be estimated either by using OLS with dummy variables for each cross-sectiontime period (so-called "least-squares dummy variables" method or LSDV hereafter) or by being demeaned to wipe out heterogeneous individual effects, presented by the "demean method". Generally, the estimates obtained from these two methods should be the same. For prediction purpose, a mixed-effect model is preferred as it provides shrunken estimates (or empirical Bayes) of the random effects (Skrondal and Rabe-Hesketh, 2003). When it comes to a random intercept model, these estimates of random effects go closer to the grand mean or within-group mean, which is dependent on the sample size and data variance of each group (Afshartous and de Leeuw, 2005; Gelma and Hill, 2007). When samples are finite, group means are found to be less efficient than the shrunken or empirical Bayes estimates in evaluating random effects (Afshartous and de Leeuw, 2005).

Even though panel data can utilize the underlying data structure well, it encounters some problems in model selection process. For models with large mixed effects, model selection carries some doubts as we cannot clearly detect or even tell nonlinear predictor effects from random predictor effects at the same time. To solve this problem, some authors have recommended to bring in a group-specific effect along with each predictor (Barr et al., 2013). However, such a strategy does not work well statistically and computationally for finite samples. Besides that, existing research also point out that missing nonlinear effects are usually wrongly approximated by a random slope in the model (Bauer and Cai, 2009). In this way, an exploratory regression analysis with a mixed effect model chosen from a model selection process could result in a wrong model which could further mislead researchers to give incorrect interpretations and wrong conclusions. Thus, although the panel data analysis techniques have already expanded the applicability of linear models to better fit data and research, the underlying assumption of the linearity still limits model prediction power. Moreover, the results can be spurious and even misleading when data gets larger and non-linearity is not considered (Bauer and J. Cal, 2009).

Realizing these shortcomings, many parametric, semi-parametric, and recursive partitioning approaches have been proposed and discussed to improve exploratory regression analysis in the form of mixed effect models. One approach is to adopt a double-penalized likelihood approach to select models and estimate model parameters (Ni et al., 2009; Schelldorfer et al., 2010; Rohart et al., 2014). Furthermore, parameter regularization can also be achieved by updating the model coefficients step-wise via likelihood-based boosting (Groll and Tutz, 2014). The primary criticism of this approach lies in the linear assumptions for all effects.

Thus, despite the great interpretability and rigorous theoretical foundations of the models mentioned above, the strict assumptions in model structure and data distribution largely limit the model performances. Firstly, these models believe that linear equations can adequately capture the underlying pattern of observed data, which implicitly rules out possible interactions among variables and any nonlinear effects. Even if some nonlinear or interaction effects can be included in the model, it is still difficult to specify all potential effects, including directly-related factors, interactions among different factors and nonlinear effects in a parametric model due to limited knowledge. In addition, a small sample makes it impossible to estimate these effects simultaneously. Furthermore, to fully explore observed data, lots of models are needed but still may not be able to capture all important features in structure of explanatory factors.

Some other approaches such as automatic procedures, including step-wise regression and best subset analysis have been proven to be unable to correctly identify predictors (Hocking and Leslie, 1967). Thirdly, inferences following parametric model selection may incur a Type I Error, which make it difficult to repeat the same results (Burnham and Anderson, 2003; Gelman and Loken, 2014). Finally, none of these approaches explicitly test the accountability of hierarchical clustering (Miller et al., 2014).

To handle nonlinear effects, researchers have proposed to utilize splines to capture nonlinear effects (Wood, 2006). Two examples of such semi-parametric models are the gamm function in the R package mgcv (Wood, 2006) built on nlme (José and Bates, 2000), and the gamm4 function (Scheipl et al., 2014) derived from the R package lme4. While these approaches are proved effective in catching nonlinear effects, the long runtime in parameter estimation when dealing with big predictors sets and large data samples has been criticized by studies .

Another proposed method that has been widely discussed is non-parametric models such as decision trees (Breiman et al., 1984). Decision trees are flexible and powerful in dealing with nonlinear effects and interactions. Furthermore, missing values in the predictors can be easily handled without making any extra statistical assumptions regarding data properties. Besides the flexibility and power in prediction, decision rules can be easily interpreted as the structure is formed by splitting variables into different buckets. Each splitting tree branch represents a conditional rule (e.g., if $x_1 < c$ and $x_2 < d$ then...), and observations in the same branch get same scores since they are considered in a same group. One problem with decision trees is that the estimated structure is dependent on data, meaning tree structure could greatly differ from one data set to another. To mitigate variances, bagging (bootstrap aggregation) is then designed and integrated in the training process to improve algorithm stability and accuracy (Breiman, 1996). Furthermore, bagging can also be adopted to help reduce possible over-fitting. Building on bagging algorithm, random forests (Breiman, 2001) are developed by fitting many decision trees to data samples and these decision trees form an ensemble model with different weighting rules. Random forests is more robust to sample variations than decision trees. These methods have been successfully applied in social science research. For example, they are used in psychology to identify factors that are important in evaluating life stress and well-being in later life (Scott et al., 2013; Wallace et al., 2002) as well as assessing caregiver stability (Proctor et al., 2011). In clinical research, decision trees have been utilized in suicide attempt predictions (Mann et al., 2008) and the analysis of how competence affects children depressive symptoms (AD et al., 1997).

Staying within tree structures, a machine learning tool that can be applied in linear relationship but goes further for broader usage as well as having a high prediction power is Gradient Boosting Machine (GBM hereafter). GBM is a popular and successful application of gradient boosting in additive tree models, meaning that learning is conducted to learn the residuals of previous trees following a gradient descent direction. It was introduced by Friedman (Friedman, 2001) and has been proven powerful in building predictive models. It is a gradient descent method that explores and finds an optimal model in the function space, which is further translated into finding optimal parameters in the parameter space. The model is built by optimizing arbitrary differentiable loss functions in iterative stages. The prediction model is an ensemble of weak prediction models, typically decision trees for their easy interpretation. These weak "learners" becomes a strong learner as a whole after the training process and makes the algorithm a powerful tool in fitting generic non-parametric predictive models empirically.

To more fully and easily explain GBM model, the least-squares regression setting is chosen to illustrate how this model works. The goal of this setting is to minimize the mean squared error $(\hat{y} - y)^2$ for the prediction function $\hat{y} = F(x)$. At each stage m with $1 \leq m \leq M$ of gradient boosting, an imperfect model F_m is assumed in each step and an estimator h is added to F_m to improve model performance (Friedman, 2001). The gradient boosting solution starts with the observation that a perfect model would imply

$$F_{m+1}(x) = F_m(x) + h(x) = y,$$

or, equivalently,

$$h(x) = F_{m+1}(x) - F_m(x),$$

where X is a set of predictors and h is fit to the residual $y - F_m(x)$. Thus, F_{m+1} is learned to correct the predecessor F_m in each step. The algorithm proposes that residuals of the loss function can be efficiently mitigated by following the direction to which the negative gradient points to. Furthermore, the idea can be generalized to other loss functions and classification problems.

Another way to capture non-linear fixed effects is recursive partitioning approaches. These approaches can estimate a large number of predictors quickly. Some studies fit linear mixed models for each region of a decision tree. The fixed effects are modeled using decision trees while random effects are trained separately with linear mixed models (Abdolell et al., 2002). The resulting trees are easy to interpret and this is helpful in exploratory regression analysis. However, this method requires the mixed effects model to be specified a-priori. To allow group-specific nonlinear effects, splines are utilized. However, the group-specific effects are only allowed for a single covariate only.

In general, current machine learning method in handling panel data with regard to fixed effects can be simply categorized into the following ways: 1) introduce dummy variables either by one-hot-encoding, a process where categorical variables are converted into multiple 0-1 dummy variables, each of which represent an individual category, as inputs to ML algorithms; or 2) loop between panel regression and machine learning tools until certain stopping criteria are met (Miller et al., 2014); 3) group observations using independent variables by clustering and run predictive machine learning tools separately for each group (Verma et al., 2018). There are also machine learning methods that handle random effects by bringing in the expectation-maximization method into the learning process (Hajjem et al., 2011, 2014). To illustrate these methods, three papers are used as key references with the models designed in the paper used as benchmark models.

The first method is to modify machine learning to adapt to panel data by integrating panel data structure into machine learning techniques in a loop structure. Metboost (multivariate tree boosting) (Miller et al., 2014) is an example. The metboost method is an extension of boosted decision trees that is designed specifically to deal with data with a hierarchical structure (Bühlmann and Hothorn, 2007). The basic logic behind the algorithm is to shape the tree structure to be the same across groups but allows the difference among groups to be reflected in the terminal node. Univariate trees are fitted to multivariate outcomes by capturing the most covariance coming from predictors. Furthermore, Metboost addresses model selection problems without any specification of nonlinear effects a-priori and approximates these effects by additive models of trees smoothly. The algorithm works well for a medium sample with hundreds of predictors or a big data with a limited number of predictors. As an extension of the boosted decision trees, Metboost is different from the boosted decision trees in that Metboost incorporate the effects of grouping variables in each tree. With this difference, the algorithm works better than boosted decision trees in both variable selection and prediction performance. One limitation of the algorithm is that it treats clustering variables as fixed, which may generate poor model performance when predicting observations at unknown levels of clustering variables or identifying important factors in the process of generalizing the model to other possible levels of the grouping variables.

Instead of retrieving differences after non-linearity exploration in the Metboost method, another concept in handling panel data is to group individuals for their difference before the training process begins. Based on this thought, KMeansXGBoost (Verma et al., 2018) is designed to adapt the machine learning technique to panel data. KMeansXGBoost is a hybrid machine learning model that combines K-means with XGBoost. This method clusters and groups different individuals first and then runs prediction models differently for each individual cluster. In detail, K-Means clustering groups all the observations into multiple groups based on the patterns. Then XGBoost model trains each group separately to extract the unique pattern for each group.

Besides the two methods mentioned above, another logic is to introduce hidden factors by adopting the expectation–maximization algorithm in handling random effects, such as “mixed-effects random forest” (MERF hereafter) (Hajjem et al., 2011, 2014). MERF is an integration of Linear mixed effects (LME) model and random forest model. The model replaces the component linear model with random forest model to relax the pre-assumptions of function forms. Each random forest is trained after learned random effect is subtracted from the original value of dependent variables. MERF has proven to improve model performance significantly when random effects are important.

Though these existing machine learning techniques help improve model performance by conquering some disadvantages in a traditional panel data analysis method, they all have shortcomings. KMeansXGBoost is constrained by the pre-assumption of the number of clusters and by the assumption that independent variables help decide individual differences. MERF relies on the assumption of hidden factors driving random effects. All of the three have the limitation of pre-deciding how many clusters are included in the model. For typical panel data in finance study, we often face a large number of cross sections with short time periods. In this case, the degree of freedom issue brought by numerous number of dummies would hurt the learning process. Furthermore, representing each section with one dummy variable is detrimental to some degree as different sections can be within the same group for slight differences that can be ignored. Finally, for clustering using the K-Means method based on independent variables, the grouping cannot catch the specific effects not inherited in independent variables but in some unknown factors. To relax all the assumptions mentioned above for these existing machine learning methods, we propose a new model machine learning method called “PanelGBM”, which will relaxes the assumptions these existing machine learning methods make on clusters.

The PanelGBM method is designed specifically for panel data without any assumptions in number of clusters (i.e., dummy variables for fixed effects) and what drives those fixed effects. The basic idea is to derive dummy variables that cluster each individual unit to different groups based on the effects which are unrelated with independent variables. To exploit the prediction contribution of independent variables, GBM is utilized as the underlying machine learning model. The unique characteristic of this method that differentiate itself from all the existing machine learning techniques mentioned above is that the number of clusters are not known beforehand but is learning from the training process. One possible concern of this proposed method is overfitting issue, which is controlled with cross-validation method and hyper-parameters re-tuning in the model training process. More details of the algorithm, training process and discussion will be described in the next chapter.

We further test the performance of the proposed algorithm via simulation data and empirical data. The basic idea of data generation process follows existing studies (Miller et al., 2014; Hajjem et al., 2011, 2014). Multiple rounds of simulation data are generated where at the same time data driven by different underlying pattern to ensure that the test results we get from the simulation study are stable and can be generalized. For each test, a panel data with N sections with each section possess M points are generated. Thus, there are MN data points involved for each test. Independent variables or predictors X are generated based on some distribution assumptions, such as uniform distribution or normal distribution. The response variable y is then the outcome of some designed patterns acting on these independent variables along with a random component, which is often assumed as a white noise. In other words, for each generated panel data, we have:

$$y_{i,j} = Pattern(X) + \mu_i (i = 1, 2, ..., N; j = 1, 2, ..., M),$$

where $Pattern$ is the predefined rules or rules we use to define the characteristics of data. Furthermore, what patterns or how many patterns are needed would depend on topics under study.

The model training and test process of this dissertation follows existing literature. For the data sets we generate, not all of them are used to train the model; instead a proportion of the data is randomly picked for model training and the rest are used for model evaluation. For each data point, there are multiple sampling methods to split the data for training/testing. In general, we have: 1) simple random sampling: selecting points for training randomly with each data point picked at the same probability; 2) systemic sampling: pre-defining a pattern to select data points (e.g. at intervals); and

3) stratefied sampling: having some pre-specified groups that need to be considered in order to ensure the sampled data follows the same pattern as the whole population. In our case, we will adopt simple random sampling since our proposed algorithm is expected to work in general with different characteristics of data, including balanced data and imbalanced data. Besides the design of different data patterns in the simulation, we also need to run the whole set multiple times for each pattern to ensure the model performance is stable. In our dissertation, we will follow what current studies do and run simulation 100 times (Hajjem et al., 2011, 2014; Wu et al., 2011).

To measure the performance of regression model, The following three performance metrics are often considered: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and $R-squared$ or $Adjusted-R-squared$. The logic of how each performance metric works are described in details below:

1) Mean Absolute Error(MAE)

MAE is the average of the absolute prediction errors where prediction errors are the differences between predicted values and observed value. Small MAE implies good model performance

2) Root Mean Square Error (RMSE)

RMSE is the standard deviation of prediction errors. Following the statistic meaning of standard deviation, RMSE states the dispersion of prediction errors. Models with lower RMSE are preferred. Since RMSE is the squared root of the average of squared prediction errors, RMSE is differentiable and easier to do mathematical analysis with, which makes it a better performance metric than MAE. However, due to the square operation in the calculation, RMSE weights more on big errors and thus makes it sensitive to outliers.

3) $R-squared$ or $Adjusted-R-squared$

R-squared ($R-squared$) tells how much the variation of an dependent variables can be explained by independent variables. The equation below shows how $R-squared$ is calculated:

$$R-squared = 1 - \frac{Unexplainedvariation}{TotalVariation}$$

where Unexplained variation is usually measured by Sum Square of Residuals (SSR for short) which is the squared sum of prediction errors, and Total variation is defined as Sum Square of Total (SST for short) which is the sample variance of the dependent variable. Thus, $R - squared$ usually ranges between 0 to 1. However, if regression models work worse than average estimate, then the value could be negative. The higher $R - squared$ is, the better regression models work. The logic behind this is that low $R - squared$ implies that regression models fit dataset pretty well or the data points that are not well predicted are distributed with low variance. One drawback of $R - squared$ is that adding new independent variables into models would help improve $R - squared$ regardless of whether new independent variables improve model prediction power. To solve this problem, $Adjusted - R - squared$ is proposed by penalizing the model with more variables.

Among these commonly used performance metric of regression models, which one is preferred is subjective to data under exploration and model adopted. In most cases, these three metric are aligned with each other. In our study, these metrics are consistent with each other with difference in the scale when it comes to rank model performance of different algorithms. However, we prefer to use $R - squared$ since it is the most common performance metric for linear regression and panel regressions and is also commonly used in machine learning (Tang et al., 2021; Banik et al., 2017; Wang et al., 2011).

Chapter 8

PROPOSED ALGORITHM PANELGBM

We propose a new method called PanelGBM to help improve model performance and grouping efficiencies for panel data studies. We call the procedure PanelGBM as it is specially designed for panel data and utilizes GBM to tackle non-linearity and interactive effects among independent variables. The algorithm is conducted by extending the gradient boosting tree to take hierarchical clustering into accounts. Gradient boosting tree develops trees sequentially with each tree trying to reduce prediction residuals that are not caught by preceding trees. For each round of the model training process, i.e., after extracting information from independent variables, dummy variables are introduced sequentially to catch possible exogenous factors that are inherent in individual sections but cannot be predicted from models with existing explanatory variables.

We narrow down multiple non-linearity conditions to focus on the non-linearity between independent and dependent variables only. The basic form is:

$$Y_{it} = GBM + \mu_{it}(i = 1, 2, ..., N; t = 1, 2, ..., T)$$

with $\mu_{it} = \mu_i + \lambda_t + \upsilon_{it}$ and υ_{it} is independent and identically distributed normal distribution, i.e., N(0,σ^2). μ_i and λ_t are section-level fixed effects and time-level fixed effects. These two types of fixed effects are handled in the same way in PanelGBM, thus we will illustrate how the algorithm extract μ_i as an example. The dummy variables, which are used to indicate fixed-effects are introduced sequentially for the following advantages: 1) assumption-free distribution of fixed effects; 2) far fewer dummy variables needed to be introduced, thus avoiding degree of freedom issue; 3) handling fixed effects caused by exogeneity other than the known independent variables. Since this method groups individuals based on the residuals of training models, two mechanisms are designed to avoid over-fitting and ensure the stability of model performance: cross validation and hyperparameter tuning after multiple rounds of training.

The training process of PanelGBM is described below:

---

1. GBM training with existing independent variables. The GBM model is trained from a grid search process of a set of combined hyperparameters P_1.

2. Calculate error $e_{1,i,t} = Y_{i,t} - GBM(X)$ for each section i at time t where GBM(X) is the prediction from the trained GBM model. .

3. Calculate the mean error for each section i in round 1, $\bar{e}_{1,i}$, and the sample mean in the first round, $\bar{e}_1$, across sections.

4. Introduce a dummy variable $d_{1,i}$ with

$$d_{1,i} = \begin{cases} 1, & \text{if } \bar{e}_{1,i} \geq \bar{e}_1, \\ 0, & \text{if } \bar{e}_{1,i} < \bar{e}_1. \end{cases}$$

   For sections whose average residual $\bar{e}_{1,i}$ is below the sample residual $\bar{e}_1$, 0 is assigned to the dummy variable $d_{1,i}$. Otherwise, $d_{1,i}$ is set as 1. Train a GBM model (using the P_1 set of hyperparameters) with the existing independent variables and the newly added dummy variable d_1; repeat the process 2 and 3 by m times.

5. Re-train the GBM model with independent variables and the introducerd dummy variable set $d_1, ...d_m$ to find newly adjusted hyperparameters P_2.

6. Repeat the process steps 2, 3, 4 and 5 until the model cannot be improved. A stopping rule can be adopted as an option here to stop the training process. The initial choice is that the model performance improvement reaches 0 at a certain

step. In our research, we do not use the stopping rule to illustrate how the model works in each learning loop for different simulation data.

Seen from the detailed training process described above, PanelGBM is more flexible than other machine learning methods we discussed earlier in the literature review since we do not need to make too many assumptions. For example, we do not need to make such assumptions that are pre-requisites for other existing machine learning methods: number of dummies needed to add to the model, assumptions that independent/hidden factors drive individual differences/random effects, and distribution assumptions about individual effects. Instead, the number of dummy variables is decided by the data itself in how many rounds of individual training process have been conducted in the process. Furthermore, both exogenous factors and known independent factors in the model could contribute to the individual dummies so no assumptions are made in which drives the individual effects in the PanelGBM model.

PanelGBM is expected to work well in prediction tasks in that it can exploit non-linearity among variables and utilize individual fixed effects from introduced dummy variables. GBM outperforms linear regressions when there exists non-linearity between independent variables and dependent variables. Furthermore, PanelGBM is expected to work better than plain GBM in cases where fixed effect matters. For panel data where fixed effects do not exist, PanelGBM works as plain GBM model since the very early stage of the PanleGBM model is plain GBM model. The more contribution the fixed effects make to dependent variables, the more important these dummy variables become and the better performance PanelGBM can exhibit. The number of dummy variables needed to be introduced is affected by how variant fixed effects are. More variant fixed effects are means more groups are needed, and leads to more dummy variables. To test the performance of the PanelGBM algorithm and its property on the different situations discussed above, a simulation study where variations of simulation data are generated.

One limitation of this method is that the performance of the algorithm heavily relies on the match between data and the base model adopted; in our case, the base model adopted is GBM. For data/topics that can be well explored by GBM, PanelGBM would work better than others. However, when GBM method cannot fully exploit data information, the introduction of dummy variables learned through our design would hurt model performance. However, the way dummy variables are introduced can be easily applied as a wrapper for other machine learning methods, which can conquer this limitation. For instance, it can be integrated with the Hidden Markov Model when dealing with topic inference or integrated with neural network in dealing with temporal relationships where these algorithms can catch data information better than GBM.

8.1 Simulation Process

In this section, we will compare model performance of PanelGBM with benchmark models on simulation data. The benchmark models include linear model, panel model, GBM, MERF model, and multivariate variable boosting ("MVBoost") model. Simulation data cover multiple data sets that follow similar pattern with slight variations in model linearity, fixed effects and variable correlations. The generating process of baseline data set is described below:

The simulated data has 100,000 observations that belong to 4,000 sections with each section possessing 25 observations.

1. For independent variables, we generate 9 random variables $(X_1, ..., X_9)$ from a multivariate normal distribution $\sim N(0, \Sigma)$ where Σ is the covariance matrix of these 9 variables. In our simulation, the variances of each variable is 1 and correlation coefficient of each pair of variables is 0.6 (Hajjem et al., 2014). The reason why we chose multivariate normal distribution is to simulate actual data.

2. Fixed effects Fix_i are uniformly selected from the range (0,10) with an interval of 0.001.

3. The dependent variable $Y0$ and Y are generated using only 3 variables generated in the first step and the fixed effects in the second step following the formula below ((Hajjem et al., 2014)):

$$Y0_{ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij}\varepsilon_{ij}.$$

$$Y_{ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + Fix_i + \varepsilon_{ij}.$$

Some function forms such as *log* is introduced to make the overall model complicated such that our methods can be tested to be flexible enough to deal with complicated data generation process.

To test the robustness of the algorithm, we make some variations to the baseline data set in data linearity, the relationship among independent variables and properties of fixed effects that include range, distribution, and variance. We get the following different sets of simulation data for model performance tests:

1.) Model variations in independent variable correlations and fixed effects

(A) Interval of the uniform distribution fixed effect. The tested data are different in the interval of the fixed effects. The interval of fixed effects decides the minimal difference among two different fixed effects.

(a) F_i is uniformly selected from the range (0,10) with interval be 0.0001:

$$Y_{M1,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(b) F_i is uniformly selected from the range (0,10) with interval be 0.01:

$$Y_{M2,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(B) Distribution of fixed effects. The fixed effects of the tested data are sampled from two distributions: normal distribution and Poisson distribution.

(a) F_i is selected from a normal distribution $N(10, 1)$:

$$Y_{D1,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(b) F_i is selected from a Poisson distribution with mean at 10:

$$Y_{D2,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(c) F_i is a combination value that gives same weights to both a Poisson distribution with mean at 10 and a normal distribution $N(10,1)$:

$$Y_{D3,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(C) Range of the distribution fixed effects. The fixed effects of the tested data are sampled from two similar types of uniform distributions. The only difference between the two is their range, meaning the max values to be retrieved from the distribution are different.

(a) F_i is uniformly selected from the range (0,1) with an interval of 0.001:

$$Y_{V1,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(b) F_i is uniformly selected from the range (0,100) with an interval of 0.001:

$$Y_{V2,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(D) Correlation between independent variables. The independent variables of tested data have different levels of correlation with each other.

(a) Correlations among variables are 0 :

$$Y_{C1,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

(b) Correlations among variables are 0.9 :

$$Y_{C2,ij} = 2 * X_{1,ij} + X_{2,ij}^2 + 4(X_{3,ij} \geq 0) + 2log|X_{1,ij}|X_{3,ij} + F_i + \varepsilon_{ij}.$$

2.) Model variations in model linearity

We also explore how well the PanelGBM model performs in different scenarios where then underlying model of the baseline data sets is more linear, where is described below:

$Y_{R1,ij} = 2 * X_{1,ij} + 8 * X_{2,ij} + 5 * X_{3,ij} + 4(X_{3,ij} \geq 0) + F_i + \varepsilon_{ij}.$

Standing on this relationship. we test performance of different algorithms mentioned above in different scenarios of fixed effect conditions. That is, we test the new model performance with different fixed effects interval ($YR1M1$ and $YR1M2$), different magnitudes of fixed effects (*YR1V1* and *YR1V2*), and with different distributions of fixed effects (*YR1D1*, *YR1D2* and *YR1D3*).

3.) Summary: simulation schema

To summarize all the scenarios we mentioned above, the schema of the simulation data is shown below:

Table 8.1: Data simulation schema

The table illustrates the schema of data simulation process. The first step of the simulation is to explore model performance for fixed effects under different conditions, including the fixed effects under different distributions (*YD1, YD3, YD13*), fixed effects generated from distributions with different range (*YV1, YV3*), and interval (*YM1, YM3*). Moreover, models with independent variables with different correlations (*YC1, YC3*) are tested. *Y0* is the dependent variable without any fixed effects, and Y is the base model. The second part is to repeat the simulation process with the underlying model of baseline data set possessing more linearity than the benchmark models in the first part.

Simulation Schema

Model Linearity	Base Model		Correlation		Range		Interval		Distribution		
Base	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
More Linearity	-	YR1	-	-	YR1V1	YR1V2	YR1M1	YR1M2	YR1D1	YR1D2	YR1D3

8.2 Model Training Process

The simulation data is split into training data and testing data with the proportion of 70% to 30%. For each section, data is split randomly such that each data is kept as unbalanced panel data, which matches better than balanced splitting with the real data research encountered in finance study. Following the training process introduced in the literature review, the detailed training process of the algorithm is described below:

---

1. GBM training with existing independent variables. In our case, the GBM training case is aimed to find the best set of hyperparameters in the number of trees, depth, and learning rate. The number of trees are chosen from 100 to 600 with steps at 100, the depth is chosen from 2 to 8 with steps at 1, while the learning rate is chosen from 0.05 to 0.1 with steps at 0.01. The final values of these key hyperparameters P_1 are chosen based on model performance employing a grid search process. We calculate the error of each observation as $e_{1,it} = Y_{it} - GBM(X)$, where GBM(X) is the prediction from the trained GBM model.

2. Calculate the mean error for each section i in round 1, $\bar{e}_{1,i}$, and the sample mean at the first round, $\bar{e}_1$ across sections. These two error measures are calculated as follows:

$$\bar{e}_{1,i} = \sum_{t=1}^{n} \frac{e_{1,it}}{n}$$
$$\bar{e}_1 = \sum_{i=1}^{m} \frac{e_{1,i}}{m}$$

where n is the number each section has in the training data, and m is the number of unique sections in the training data.

3. The dummy variable $d_{1,i}$ is introduced following the rule shown below:

$$d_{1,i} = \begin{cases} 1, & \text{if } \bar{e}_{1,i} \geq \bar{e}_1 \\ 0, & \text{if } \bar{e}_{1,i} < \bar{e}_1 \end{cases}$$

For the sections whose average residual $\bar{e}_{1,i}$ is below the sample residual $\bar{e}_1$, 0 is assigned to the dummy variable $d_{1,i}$. Otherwise, $d_{1,i}$ is set as 1. Train the GBM model (using the P_1 set of hyperparameters) with the existing independent variables and the newly added dummy variable d_1; repeat the process 2, 3 for m times. In our dissertation, we set m to be 3.

4. We re-train the GBM model with the original independent variables and the three newly generated dummy variable sets $d_1, ...d_m$ and retune the hyperparameters with the new combination of P_2.

5. We repeat the process steps 2, 3, and 4 until the model performance cannot be improved. In our dissertation, we get our final model after 3 repetitions of step 2 and step 3.

8.3 Model Performance Comparison and Grouping Performance

8.3.1 Model Performance in Simulated Data

The performance of PanelGBM models and benchmark models are reported in Table 8.2. Benchmark models include linear model, panel model, GBM, and MERF model. For each model and each condition covered, the performance of both the training and testing data are revealed. The model performance is evaluated using $R-squared$ instead of MSE (Mean squared error) for two reasons. Firstly, $R-squared$ is the most common used performance metric for linear models and panel regression models and is also a common metric in machine learning studies (Tang et al., 2021; Banik et al., 2017; Wang et al., 2011). Secondly, with same data, models show higher MSE would also imply higher $R-squared$, and results from $R-squared$ are consistent with the ones from MSE. The model performance in the training data and testing data is shown in Panel A and Panel B, respectively. Panel C shows the model deterioration from the training to testing data. Column $Y0$ shows model performance with simulated data without fixed effects and Column Y shows model performance of baseline line data. Model performance with different correlations is shown in $YC1$ and $YC2$; model performance with different ranges is shown in $YC1$ and $YC2$; model performance with different intervals is shown in $YM1$ and $YM2$; and model performance with different distributions is shown in $YD1$ $YD2$, $YD3$.[1]

[1] Graphical illustration of model performance comparison is Appendix Figure A.3.

Table 8.2: Model performance comparison from simulations

Model performance refers to R-squared. They are "Linear Regression", "Panel Regression", "GBM", "K-means GBM", "MERF", "MvtBoost", and "PanelGBM". From left to right, we illustrate the model performance in data without fixed effects, baseline data, and in simulation data with different independent variable correlation, and with fixed effects varying in range, interval and distribution. The model performance in the training data is shown in Panel A, and the model performance in the testing data is illustrated in Panel B. Panel C shows the model deterioration from the training to testing data.

Panel A: Training

	Base Model		Correlation		Range		Interval		Distribution		
Training	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
Linear Regression	0.844	0.613	0.619	0.612	0.841	0.021	0.614	0.615	0.809	0.586	0.752
Panel Regression	0.853	0.894	0.894	0.895	0.854	0.996	0.893	0.893	0.859	0.899	0.869
GBM	0.955	0.694	0.699	0.691	0.952	0.026	0.694	0.697	0.915	0.665	0.850
K-meansGBM	0.915	0.695	0.699	0.691	0.951	0.027	0.695	0.697	0.915	0.665	0.850
MERF	0.991	0.994	0.994	0.994	0.9913	0.9998	0.994	0.994	0.9919	0.9943	0.993
MvtBoost	0.955	0.694	0.700	0.691	0.951	0.026	0.694	0.697	0.915	0.665	0.850
PanelGBM	0.958	0.969	0.969	0.969	0.958	0.999	0.969	0.969	0.959	0.970	0.962

Panel B: Testing

Testing	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
Linear Regression	0.847	0.615	0.612	0.610	0.844	0.019	0.619	0.614	0.811	0.586	0.750
Panel Regression	0.838	0.884	0.880	0.881	0.839	0.996	0.884	0.883	0.845	0.887	0.857
GBM	0.955	0.692	0.696	0.689	0.951	0.021	0.694	0.694	0.915	0.658	0.847
K-meansGBM	0.914	0.692	0.694	0.688	0.951	0.021	0.694	0.693	0.914	0.658	0.847
MERF	0.9337	0.9497	0.949	0.949	0.9318	0.9982	0.9494	0.9489	0.9329	0.951	0.938
MvtBoost	0.954	0.692	0.696	0.689	0.951	0.021	0.694	0.694	0.915	0.659	0.848
PanelGBM	0.951	0.965	0.964	0.964	0.952	0.999	0.965	0.965	0.953	0.965	0.957

Panel C: Deterioration

	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
Linear Regression	0.003	0.003	-0.011	-0.004	0.003	-0.119	0.009	-0.001	0.003	-0.001	-0.002
Panel Regression	-0.018	-0.011	-0.016	-0.016	-0.018	0.000	-0.011	-0.012	-0.016	-0.013	-0.014
GBM	-0.001	-0.003	-0.005	-0.003	-0.001	-0.199	0.000	-0.006	-0.001	-0.009	-0.003
K-meansGBM	-0.001	-0.004	-0.007	-0.004	0.000	-0.226	-0.001	-0.006	-0.001	-0.011	-0.004
MERF	-0.058	-0.045	-0.045	-0.045	-0.060	-0.002	-0.045	-0.045	-0.059	-0.044	-0.055
MvtBoost	0.000	-0.002	-0.005	-0.003	0.000	-0.188	0.001	-0.005	0.000	-0.009	-0.003
PanelGBM	-0.007	-0.004	-0.005	-0.005	-0.006	0.000	-0.004	-0.004	-0.007	-0.006	-0.005

Seen from Panel A and Panel B in Table 8.2, models perform differently in different cases. "Linear Regression" model performs the worst in all models, which makes sense as it only captures the linear relationship between independent and dependent variables. The "Panel Regression" model performs similarly to the "Linear Regression" model when fix effects do not exist, as seen from Column $Y0$, but performs better in other cases when fixed effects do exist, as seen in other Columns. The performances of the three GBM-related benchmark models, i.e., "GBM", "K-means GBM", and "MvtBoost" are quite similar for almost all the simulated cases. We can conclude, therefore, that the relative performances between the three GBM-Benchmark models and the "Panel Regression" model differ in conditions where fixed effects vary. GBM-related benchmark models perform better in *YM1*, *YD1*, *YV1*, and *YR1*, the cases where fixed effects are smaller, less varied, and more normally distributed. One possible explanation is that the inherited assumptions behind the "Panel Regression" and "GBM" models. The linearity assumption of "Panel Regression" renders its inability to catch the non-linearity between the independent and dependent variables, but can capture the fixed effects by bringing in dummy variables representing individual sections. On the contrary, "GBM" can catch the non-linearity but ignore the individual fixed effects. Thus, for fixed effects that are not so important in prediction, "GBM" performs better than "Panel Regression", as we see in our case, but performs worse otherwise. Comparison of model performance in testing data among all the models shows that "PanelGBM" is the best model in all the cases as it can capture both the non-linearity in the data as well as fixed effects.

Furthermore, the model performance of PanelGBM is more stable than other models at similar model performance across different scenarios as shown in Panel C. The model deterioration in Panel C is defined by the percentage of model performance change from the training to the testing data. The top two models that work the best among all models are "MERF" and "PanelGBM". "MERF" performs better than "PanelGBM" in most cases in the training data but worse in the testing data, as observed from the top two Panels.

We further show model performance of PanelGBM model in repeated experiments for all cases to test the stability of model performance. For each case, we repeat the data generating and model training process for 100 times and report the results in Table 8.3. Seen from "STD" in both Panel A regarding model performance in training data and Panel B regarding model performance in testing data, we can tell that the variance of model performance is quite small, meaning that model performance is quite stable over different simulated data.

Table 8.3: Model performance from repeated experiments

The table illustrates statistics of PanelGBM model performance on repeated experiments. Model performance refers to R-squared. From left to right, we illustrate the model performance in data without fixed effects, baseline data, and in simulation data with different independent variable correlation, and with fixed effects varying in range, interval and distribution. Model performance in the training data is shown in Panel A, and model performance in the testing data is illustrated in Panel B.

Panel A: Training Data

Statistics	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
N	100	100	100	100	100	100	100	100	100	100	100
MIN	0.956	0.968	0.968	0.968	0.956	0.999	0.968	0.968	0.958	0.969	0.961
MEDIAN	0.957	0.969	0.969	0.969	0.957	0.999	0.969	0.969	0.959	0.97	0.962
MEAN	0.957	0.969	0.969	0.969	0.957	0.999	0.969	0.969	0.959	0.97	0.962
MAX	0.959	0.97	0.969	0.97	0.959	0.999	0.969	0.969	0.96	0.971	0.963
STD	0.0005	0.0003	0.0003	0.0003	0.0005	0.00004	0.0003	0.0003	0.0005	0.0005	0.0004

Panel B: Testing Data

Statistics	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	D3
N	100	100	100	100	100	100	100	100	100	100	100
MIN	0.95	0.963	0.963	0.964	0.95	0.998	0.963	0.964	0.952	0.964	0.955
MEDIAN	0.952	0.965	0.965	0.964	0.952	0.998	0.965	0.965	0.953	0.965	0.956
MEAN	0.951	0.965	0.964	0.964	0.951	0.998	0.964	0.965	0.953	0.965	0.956
MAX	0.953	0.966	0.966	0.966	0.953	0.999	0.965	0.966	0.955	0.967	0.958
STD	0.0006	0.0004	0.0005	0.0004	0.0006	0.00002	0.0004	0.0004	0.0006	0.0005	0.0005

8.3.2 Model Performance in Simulated Data with More Linearity

Since the degree of non-linearity among data variables is important in model performance, we repeat the data generating process for different cases but with a more linear underlying baseline pattern as described in the Section 3.1.

Consistent results are found in the new set of simulated data shown in Table 8.4.

We still observe PanelGBM as the best performer with least performance deterioration. Comparing model performance in new simulated data set with that of previous more non-linear data, we can see that the difference in model performance between linear model and PanelGBM is less in new simulated data when compared with that in previous data sets. For example, in Table 8.2, the performance of PanelGBM is 0.957 versus 0.857 in Panel regression in YD3 case, illustrates an improvement of 0.1. However, in Table 8.4, the performance of PanelGBM is 0.994 compared with Panel data at 0.986, indicates an improvement of 0.008, which is much smaller than the 0.1 improvement shown in Table 8.2.

Table 8.4: Model comparison for model with higher linearity

we repeat the data generating process for different cases but with a more linear underlying baseline pattern as described in the Section 3.1. Model performance still refers to R-squared. They are "Linear Regression", "Panel Regression", "GBM", "K-means GBM", "MERF", "MvtBoost", and "PanelGBM". From left to right, we illustrate the model comparison in simulation with fixed effects varying in range, magnitude, and distribution. The model performance in the training data is shown in Panel A, and the model performance in the testing data is illustrated in Panel B. Panel C shows the model deterioration from the training to testing data.

Panel A: Training							
Training	YR1V1	YR1V2	YR1M1	YR1M2	YR1D1	YR1D2	YR1D3
Linear Regression	0.986	0.177	0.945	0.945	0.982	0.938	0.973
Panel Regression	0.988	0.998	0.988	0.988	0.988	0.988	0.988
GBM	0.994	0.180	0.953	0.953	0.990	0.946	0.980
K-meansGBM	0.994	0.182	0.953	0.953	0.990	0.946	0.981
MERF	0.999	1.000	0.999	0.999	0.999	0.999	0.999
MvtBoost	0.994	0.180	0.953	0.953	0.990	0.946	0.980
PanelGBM	0.995	0.999	0.995	0.995	0.995	0.995	0.995
Panel B: Testing							
Testing	YR1V1	YR1V2	YR1M1	YR1M2	YR1D1	YR1D2	YR1D3
Linear Regression	0.987	0.182	0.945	0.945	0.982	0.938	0.972
Panel Regression	0.986	0.998	0.987	0.987	0.986	0.987	0.986
GBM	0.994	0.182	0.952	0.952	0.990	0.945	0.980
K-meansGBM	0.994	0.184	0.952	0.952	0.989	0.945	0.980
MERF	0.992	0.999	0.992	0.992	0.992	0.992	0.992
MvtBoost	0.994	0.182	0.952	0.952	0.990	0.945	0.980
PanelGBM	0.994	0.999	0.995	0.994	0.994	0.994	0.994
Panel C: Deterioration							
Deterioration	YR1V1	YR1V2	YR1M1	YR1M2	YR1D1	YR1D2	YR1D3
Linear Regression	0.000	0.030	0.000	-0.001	0.000	0.000	0.000
Panel Regression	-0.001	0.000	-0.001	-0.001	-0.001	-0.001	-0.001
GBM	0.000	0.012	-0.001	-0.001	0.000	-0.001	0.000
K-meansGBM	0.000	0.009	-0.001	-0.001	0.000	-0.001	-0.001
MERF	-0.007	-0.001	-0.007	-0.007	-0.007	-0.007	-0.007
MvtBoost	0.000	0.016	0.000	-0.001	0.000	-0.001	0.000
PanelGBM	-0.001	0.000	-0.001	-0.001	-0.001	-0.001	-0.001

End of Table

8.3.3 Performance of Grouping Dummies in Simulated Data

To illustrate the roles of dummy variables in model performance, we show the model performance of "PanelGBM" model in each round in Table 8.5. Panel A reports model performance and model performance improvement in the training data set, and Panel B reports model performance and model performance improvement in the testing data. The introduction of dummy variables helps improve model performance significantly, especially in the cases where fixed effects are big and important in prediction. For instance, in Panel B, the model performance in *YV2*, the case where fix effects contribute greatly to the dependent variable, increases from 0.021 to 0.756, a rising of 3480.95%. However, the improvement for data with fewer fixed effects is much smaller: only around 29.48% for *Y*.

Further comparison of model improvement in each round shows that dummy variables introduced at the early stages of model training process make more significant contributions than those introduced in later stages (Table 8.5). For example, for *YM2* in Panel B, the first dummy variable improves model performance by 28.96% while the second dummy variable improves model performance by 5.92%. This finding is also found in training data sets in Panel A.

Furthermore, the less complex the relationship between independent and dependent variables and the less varied and smaller the fixed effects are, the less rounds that are needed to improve the model as shown in *YM1* and *YR1* case. This makes sense as concentrated fixed effects are easy to cluster, and simple relationships are easy to learn. Furthermore, in the *YM1* and *YR1* cases, we see that the introduction of the third dummy and dummies introduced afterward begin to reduce model performance in the testing data, meaning that the model begins to introduce dummies based on true errors from the training data. However, the deterioration of model performance from the third round on have been at 0.1%.

Table 8.5: Model performance for each round of simulation

This table demonstrates the model performance of “PanelGBM” in each round for different simulations. Panel A reports model performance and model performance improvement in the training data; Panel B reports model performance and model performance improvement in the testing dataset. From left to right, we illustrate the model performance in data without fixed effects, baseline data, and in simulation data with different independent variable correlation, and with fixed effects varying in range, interval and distribution.

Panel A: Training Data											
Performance (Training)											
Round	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
0	0.955	0.694	0.7	0.691	0.952	0.026	0.694	0.697	0.915	0.665	0.85
1	0.957	0.899	0.901	0.902	0.956	0.757	0.899	0.898	0.943	0.858	0.922
2	0.957	0.948	0.951	0.952	0.957	0.935	0.947	0.951	0.953	0.93	0.948
3	0.957	0.958	0.959	0.956	0.957	0.978	0.957	0.964	0.957	0.952	0.956
3 Adj	0.957	0.964	0.964	0.965	0.957	0.984	0.964	0.964	0.957	0.956	0.957
4	0.957	0.968	0.968	0.968	0.958	0.995	0.968	0.968	0.958	0.966	0.96
5	0.957	0.969	0.968	0.969	0.958	0.998	0.968	0.968	0.959	0.968	0.961
6	0.958	0.969	0.969	0.969	0.958	0.999	0.969	0.969	0.959	0.969	0.961
Final	0.958	0.969	0.969	0.969	0.958	0.999	0.969	0.969	0.959	0.97	0.962
Performance Improvement (Training)											
Round	Y	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
1	0.21%	29.54%	28.71%	30.54%	0.42%	2811.54%	29.54%	28.84%	3.06%	29.02%	8.47%
2	0	5.45%	5.55%	5.54%	0.1%	23.51%	5.34%	5.9%	1.06%	8.39%	2.82%
3	0	1.05%	0.84%	0.42%	0%	4.6%	1.06%	1.37%	0.42%	2.37%	0.84%
3 Adj	0	0.63%	0.52%	0.94%	0%	0.61%	0.73%	0%	0%	0.42%	0.1%
4	0	0.41%	0.41%	0.31%	0.1%	1.12%	0.41%	0.41%	0.1%	1.05%	0.31%
5	0	0.1%	0%	0.1%	0%	0.3%	0%	0%	0.1%	0.21%	0.1%
6	0.1%	0%	0.1%	0%	0%	0.1%	0.1%	0.1%	0%	0.1%	0%
Final	0	0%	0%	0%	0%	0%	0%	0%	0%	0.1%	0.1%
Panel B: Testing Dataset											
Performance (Testing)											
Round	Y0	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
0	0.955	0.692	0.6963	0.6888	0.951	0.021	0.694	0.694	0.915	0.658	0.847
1	0.953	0.896	0.8978	0.8982	0.952	0.752	0.897	0.895	0.939	0.851	0.918
2	0.953	0.945	0.9473	0.9475	0.952	0.934	0.945	0.948	0.948	0.925	0.944
3	0.952	0.955	0.959	0.956	0.952	0.978	0.955	0.96	0.951	0.947	0.953
3 Adj	0.952	0.961	0.96	0.96	0.952	0.983	0.961	0.96	0.951	0.95	0.953
4	0.952	0.964	0.9635	0.9635	0.952	0.995	0.964	0.964	0.953	0.961	0.955
5	0.951	0.9651	0.9642	0.9641	0.952	0.998	0.965	0.965	0.953	0.963	0.956
6	0.951	0.965	0.9643	0.9644	0.952	0.998	0.965	0.965	0.953	0.964	0.957
Final	0.951	0.965	0.9643	0.9644	0.952	0.999	0.965	0.965	0.953	0.965	0.957
Model Performance Improvement											

Continuation of Table 8.5:											
Round	Y	Y	YC1	YC2	YV1	YV2	YM1	YM2	YD1	YD2	YD3
1	-0.21%	29.48%	28.94%	30.4%	0.11%	3480.95%	29.25%	28.96%	2.62%	29.33%	8.38%
2	0%	5.47%	5.51%	5.49%	0%	24.2%	5.35%	5.92%	0.96%	8.7%	2.83%
3	-0.1%	1.06%	1.24%	0.9%	0%	4.71%	1.06%	1.27%	0.32%	2.38%	0.95%
3 Adj	0%	0.63%	0.1%	0.42%	0%	0.51%	0.63%	0%	0%	0.32%	0%
4	0%	0.31%	0.36%	0.36%	0%	1.22%	0.31%	0.42%	0.21%	1.16%	0.21%
5	-0.11%	0.11%	0.07%	0.06%	0%	0.3%	0.1%	0.1%	0%	0.21%	0.1%
6	0%	0%	0.01%	0.03%	0%	0%	0%	0%	0%	0.1%	0.1%
Final	0%	0%	0%	0%	0%	0.1%	0%	0%	0%	0.1%	0%

End of Table

Besides the contribution of dummy variables to model performance, the grouping performance of dummy variables in baseline data are also analyzed in Table 8.6. Panel A, Panel B, and Panel C report the grouping results of the first dummy variable, the first two dummy variables, and the first three dummy variables, respectively. In each panel, the "Frequency" Column reports number of sections that fall under a related category defined by dummy variables. For each category, we calculated the average ranking of sections from the model and from the true data in Column "Ranking_Model" and "Ranking_Expected", respectively. The difference between these two Columns are reported in the Column "Ranking_Diff". The bigger the absolute value of $Ranking\_Diff$ is, the more sections are wrongly classified or the more severely the sections are mis-classified; for instance, sections with small fixed effects are mistakenly classified into the groups where sections with high fixed effects belong to. The last Column shows how many observations are mis-classified under each category.

From the table, we can see that the dummy variables perform well in classifying different sections based on fixed effects generated. In Panel A, in 4,000 different sections, only 39 of them are mis-classified. 2,023 companies have been categorized in the group 0 and 1,977 in group 1. We sort the companies in terms of fixed effects from lowest to highest, and we find that the average ranking of firms in group 0 is 1,014.3 while it is 3,009.7 for firms in group 1. However, for the top 2,023 firms with low fixed effects, their expected ranking should be 1,012, which is lower than 1,014.3, meaning that some firms belonging to higher groups are mis-classified as group 0. The difference between the model grouping and the expectation is only around 2.3, meaning the mis-classification does not go too far and that there are not too many mis-classifications. Furthermore, the result that $Ranking\_Diff$ equals 2.3 indicates that the mis-classification occurs mostly for the sections with fixed effects ranked at the bottom of $Dummy1 = 0$ group or for the sections with fixed effects ranked at the top of $Dummy1 = 1$ group. Moreover, the earlier the dummy variables are introduced, the better the performance it has in grouping as illustrated in Panel B. There are only 37 sections out of 1024 that are mis-classified in the category where Dummy1 and Dummy2 are both 0. However, there are 76 sections out of 999 sections are mis-classiified in the next category, doubling the misclassification rate in the first category. This finding can be further confirmed for the three dummy variables case as shown in Panel C.

Table 8.6: Grouping performance of "PanelGBM"

The table below demonstrates the grouping performance of "PanelGBM" model in terms of mis-classification. Panel A, Panel B, and Panel C report the grouping results of the first dummy variable introduced, the first two dummy variables, and the first three dummy variables respectively. In each panel, the "Frequency" Column reports the number of sections that fall under the relevant category. The average ranking based on model and real data are illustrated in the "Ranking_Model" and "Ranking_Expected" Columns, while the difference between the two are reported in the Column "Ranking_Diff". The last Column shows how many observations are mis-classified in each category.

Panel A							
Dummy1			Frequency	Ranking_Model	Ranking_Expected	Ranking_Diff	Misclassification
0			2023	1014.3	1012.0	2.3	39
1			1977	3009.7	3012.0	-2.3	39
Panel B							
Dummy1	Dummy2		Frequency	Ranking_Model	Ranking_Expected	Ranking_Diff	Misclassification
0	0		1024	517.3	512.5	4.8	37
0	1		999	1523.7	1524.0	-0.3	76
1	0		995	2521.6	2521.0	0.6	83
1	1		982	3504.2	3509.5	-5.3	44
Panel C							
Dummy1	Dummy2	Dummy3	Frequency	Ranking_Model	Ranking_Expected	Ranking_Diff	
0	0	0	528	274.5	264.5	10.0	
0	0	1	496	775.7	776.5	-0.8	
0	1	0	494	1271.1	1271.5	-0.4	
0	1	1	505	1770.8	1771.0	-0.2	
1	0	0	489	2265.8	2268.0	-2.2	
1	0	1	506	2768.8	2765.5	3.3	
1	1	0	455	3242.6	3246.0	-3.4	
1	1	1	527	3730.1	3737.0	-6.9	

End of Table

8.3.4 Conclusion

We test model performance of PanelGBM and grouping performance of related dummy variables in simulation data. The simulation data covers variant cases to test the robustness of model performance. Multiple models are utilized as benchmark models such as linear model, panel model, GBM, MERF model, and MVBoost model. We find that PanleGBM performs the best in model performance as well as in performance stability among all the models in variant situations, especially in cases where the underlying pattern is complex and fixed effects are varied across different sections and contributes much to predictions. The good results of PanelGBM in simulation data are inherited by the way how PanelGBM is trained. The underlying machine learning model of PanelGBM handles non-linearity among variables while the dummy variables introduced in the training process take care of individual fixed effects rooted in the exogeneity among different sections that cannot be explained by input variables.

What's more, PanelGBM model groups sections well according to the individual fixed effects of each section. We find that dummy variables that are introduced in early stages can separate sections better while the ones introduced later illustrate higher mis-classification rate. This makes sense since dummy variables that are introduced in later group sections are done so at a more granular level. Because of that, dummy variables that are introduced at the early stage of the training process help improve model performance more than others.

Further exploration of this algorithm would be to explore variants of PanelGBM. For example, for time-series data that illustrates temporal connections, we may explore this method by replacing the underlying GBM model with RNN model. The other direction is to explore the application of the method in data where fixed effects exist in temporal space instead of cross section or both two types of fixed effects exist.

Chapter 9

DATA ANALYSIS IN REAL DATA

In this chapter, the proposed algorithm PanelGBM as well as other benchmark models is tested in our empirical data that has been explored in finance section before. Both model performance and grouping efficiency of dummy variables are tested in this chapter.

9.1 Model Performance of Different Models in Real Data

We separate data into training data and testing data for model training and performance evaluation. The data is split based on year, i.e., data before 2011 are used as training data for model training, while data after 2011 are used as testing data for model performance evaluation. We split the data in this way out of such considerations that the splitting ratio of training data to testing data is around 0.8 to 0.2 and we include at least 1 decade of data after 2000 in the training data since we found in our empirical study that it is after 2000 that the role of production flexibility turns significantly and we want to be conservative in evaluating the performance of our algorithm by giving more roles to these independent variables. 5-fold cross-validation is adopted for hyper-parameter tuning for all machine learning models. For PanelGBM method, since there are some firms that do not exist in the training data but exist in the testing data, some adjustments need to be made to the training process. Two methods are adopted to handle such a situation. The first method is to randomly assign 1/0 to the dummy variables for the firms not in the training data but in the testing data, and the second method is to remove those extra firms from the testing data. We cannot resort to independent variables for grouping reference as the differences among different groups could are brought by unknown factors that are unrelated to independent variables.

Model results, including model performance in training data, model performance in testing data and model deterioration are reported in Table 9.1. We find that "linear regression" is the model that performs the worst among all the models, followed by panel regression methods. The model performance for both models are negative, meaning that the model is worse than using the mean value for prediction.[2] These results are consistent with our simulation study. MvtBoost is not applicable in our case due to too many dummy variables. The performances of the GBM and K-means GBM are quite similar to what is shown in the simulation study. PanleGBM performs better than MERF model in model prediction power and model stability as seen in the last two Columns. The PanelGBM (dummy dropped) performs better than PanelGBM, which makes sense since groups learned from trained data contribute more to prediction than random assignment, even though the performance lift is not too much.

Table 9.1: Model comparison from real data

Model performance refers to R-squared. Models include "Linear Regression", "Panel Regression", "GBM", "K-means GBM", "MERF", "MvtBoost", and "PanelGBM ". Lower R-squared indicates lower performance. We illustrate the model comparison in real data for both the training data and testing data.

Model	Training Data	Testing Data	Model deterioration
Linear Regression	0.143	-74.486	-52188.1%
Panel Regression	0.681	-4.752	-797.8%
MvtBoost	NA	NA	NA
GBM	0.343	0.199	-42.0%
K-meansGBM	0.316	0.199	-37.0%
MERF	0.9571	0.3887	-59.4%
PanelGBM	0.728	0.395	-45.7%
PanelGBM (dummy dropped)	0.728	0.4377	-39.9%

[2] For practical purposes, the lowest $R - squared$ you can get is zero, but if the regression model is not better than using the mean, then $R - squared$ will be negative

9.2 Grouping Efficiency on Real Data

We ran PanelGBM algorithm on real data and rank each variable in terms of their importance in Panel A of Table 9.2. The importance is measured by the gain each variable contributes to model performance. Seen from Column 1 to Column 6, the total gain of dummy variables as a whole is higher than each financial variable no matter how many dummy variables have been introduced into the model. For example, when there is only one dummy variable introduced as shown in Column 1, the dummy variable has the highest variable importance than other financial variables. The variable importance of dummy variable $Dummy1$ is decreased when the second variable is introduced into the model, but the whole contribution of all dummy variables is still the highest.

With more and more dummy variables being introduced into the model, the gain each dummy variable contributes decreases and becomes distributed among each dummy variable. Further comparison of variable importance from dummy variables indicates that the dummy variables which have been introduced earlier contribute more than the ones introduced later. For example, in Column 2, the importance of the first dummy variable is 0.2306 and that of the second dummy variable is 0.098. This finding can be applied to other dummy variables. Furthermore, comparing variable importance across different rounds, we see that the total importance of the dummy variables have been decreasing in the fourth rounds. The reason for this is the increasing contribution of the independent variables and decreasing contributions of the dummy variables. One possible reason for this is that after four rounds of the dummy variables being introduced into the model, these dummy variables help the independent variables explain the target better by splitting better in observations that share similar individual effects.

Table 9.2: Variable importance and the role of dummy variables in linear regression

Panel A shows the variable importance of financial variables and dummy variables trained from PanelGBM model for leverage data in each round. We further examine the prediction power of the learned dummy variables by introducing them into the linear regressions with other initial financial variables in Panel B.

Panel A: Variable Importance in PanelGBM for Real Data						
Variable	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
Tangibility	0.188	0.1686	0.1553	0.1697	0.1689	0.1699
FirmSize	0.148	0.13	0.1311	0.1646	0.1598	0.1743
Profitability	0.146	0.1313	0.1289	0.1562	0.1602	0.1621
Mktbk	0.1451	0.1395	0.1325	0.1522	0.1478	0.1478
ProductionFlexibility	0.1188	0.1018	0.0967	0.135	0.1293	0.136
Dummy1	0.2541	0.2306	0.223	0.1252	0.1224	0.0964
Dummy2		0.098	0.0945	0.0541	0.0542	0.0426
Dummy3			0.038	0.0257	0.0251	0.0224
Dummy4				0.0173	0.0172	0.0182
Dummy5					0.0151	0.0167
Dummy6						0.0137
Sum of Dummy Variables	0.2541	0.3287	0.3556	0.2223	0.234	0.21
Panel B: Linear Regression with six dummy variables						
FirmSize	0.0096***					
Mktbt	-0.0056***					
Tangibility	0.2034***					
Profitability	-0.0553***					
Productionflexibility	-0.002***					
dummy1	0.2327***					
dummy2	0.1359***					
dummy3	0.0797***					
dummy4	0.0457***					
dummy5	0.0268***					
dummy6	0.0179***					
Intercept	-0.0919**					
R_Squared: Training	0.6370					
R_Squared: Testing	0.3949					

We further run a simple linear regression by introducing the six learned dummy variables from PanelGBM into the previous five financial variables with results reported in Panel B. All the dummy variables are statistically significant in the linear regression. Compared to the results of the linear regression and panel regression in

Table 9.1, we find that the linear regression with six dummy variables can achieve similar performance with panel data in training but much better performance in testing. Thus, the six dummy variables do the same job as the 12,394 dummy variables in panel regression where we have 12,395 companies in the whole data. Furthermore, these six dummy variables have captured the pattern much better than the thousands of company dummies as seen from the better model performance in testing data. The same conclusion can be applied to the comparison with regressions in our empirical section in the finance chapter. For example, the $R-squared$ we got in Table 5.4 is much smaller than it is in our Table 9.1.

9.3 Conclusion

We test model performance and variable importance of dummy variables of proposed algorithm PanelGBM in real leverage data in this chapter. We find that PanelGBM works best in validation data with least deterioration rate among the all tested models. Furthermore, dummy variables introduced in the model training process show higher variable importance than other finance variables. Even though dummy variables introduced in early stages show higher importance than other dummy variables introduced in later stages, the whole contribution from all the dummy variables is always higher than individual contribution from any other financial variables.

We further test the grouping performance of the dummy variables introduced by PanelGBM by integrating the dummy variables into linear models. We find that these dummy variables help improve model performance a lot while at the same decreasing model over-fitting issue largely when compared with linear model and panel model.

Chapter 10

CONCLUSION OF MACHINE LEARNING

Motivated by the limitation that the traditional assumption of linearity between independent variables and dependent variables largely limits the prediction power of traditional panel data analysis models, and by the fact that introducing a dummy variable for each individual section when handling fixed effects is not efficient nor necessary when some sections have very tiny differences in their fixed effect, a modified algorithm called PanelGBM is designed to improve model prediction and to group individual sections efficiently in panel data. PanelGBM has been proven to outperform other benchmark models in model prediction and stability in both simulated data and real leverage data. The simulated data covers various cases where data differs in linearity, variable correlation, as well as properties of fixed effects such as range, interval and distribution. PanelGBM is found to be the best model in all the cases, especially in the situations where a model shows more non-linearity, fixed effects are more varied or make more contributions to prediction. For real data, PanelGBM also shows its superiority in performance in testing data after data is split based on year and model is trained from training data.

The grouping performance of dummy variables learned in the training process is also very promising. In the simulation data, we find that only around 1% of sections are mis-classified after first dummy variable is introduced, and the sections that are mis-classified are the ones that ranks in the middle wherein which group this section belongs to is ambiguous. In real data, variable importance of dummy variables as a whole is higher than other financial variables, meaning that dummy variables have learned to play important roles in prediction. This finding is further confirmed by a big improvement in model performance when we integrate these learned dummy variables into linear model.

The algorithm can be used widely in many aspects. One user case is in finance research when dealing with panel data with short length but broad sections. Moreover, it can be used in many business practice. For instance, when analyzing or predicting a bunch of customers' behavior when they are not on record for a long time but with a large sample size, this algorithm can be applied. One detailed example is the analysis of customers' credit/fraud risk when they are still early month on book, meaning that they have not shown a long traceable history for us to use. In this case, we need to explore the patterns shared among different users to extract common patterns.

One limitation of this method is that the performance of the algorithm heavily relies on the match between the data and base model adopted; in our case, it is GBM. For data/topics that can be well explored by GBM, PanelGBM would work better than others. However, when the GBM method does not fit in the data, the introduction of the dummy variables learned according to our design would further hurt the model performance. However, the way dummy variables are introduced can be easily applied as a wrapper for other machine learning method, which would easily conquer the limitations we mentioned earlier. For instance, it can be integrated with the Hidden Markov Model when dealing with topic inference or be integrated with neural network in dealing with temporal relationships.

One contribution of PanelGBM is that the algorithm gives a new thought in how to cluster different firms or units in the model training process. Instead of assigning grouping dummies based on some facts such as country, industry etc. that researchers think could affect dependent variables. We admit the limitation in fully understanding the mechanism that drives the heterogeneous effects by not assigning grouping dummies based on business understanding, but directly resort to data to derive clustering dummies. The concern of deriving clustering dummies in this way is over-fitting, meaning that clustering dummies may not represent the true grouping of these sectors.

One mechanism to mitigate that concern is to combine clustering method with some models which exploit the relationships among independent factors such that clustering dummies only catch the effects that cannot be explained by independent variables. In our essay, we adopt GBM as part of the training process since GBM has proven powerful and successful in exploring linear and non-linear relationships across multiple domains. In addition to that, we also bring is cross-validation and retrain GBM models in each round after new dummies are introduced. Furthermore, a stopping criteria can be adopted if needed to mitigate the over-fitting concern.

Future research can explore different machine learning methods with this approach to extend model scope. As mentioned earlier, the proposing clustering method works better with a model that can exploit data better. However, there is no singular machine learning model that can work well for different topics and data structures. Thus, we need to explore different machine learning models to exploit data information as much as possible. For example, when dealing with panel data with assumptions that information from prior time windows would affect the input and output in the later time points, Recurrent neural network ("RNN" for short), instead of GBM, would be thought as a better choice to be the embedded machine model in this case.

Another direction to go is to expand the clustering method in cross-section dimension to time series dimension. In our research, we only explore clustering individual firm into multiple groups. However, from temporal perspectives, there are also some fixed effects at temporal dimension that cannot be explained by independent variables. Current practice of this research topic is to add dummy variables for each year. However, similar to our arguments before in grouping cross sectional units, some years may have similar fixed effects, meaning that far fewer dummy variables are needed. Thus, we can apply the proposed clustering method into temporal space. The only modification we need to make is that the residual calculation needs to be conducted in temporal space instead of in cross-section space. Furthermore, we can conduct cross-sectional fixed effects and temporal fixed effects at the same time if needed. One challenge in such a case is that which dummy variables are conducted first may matter, which might call for some further studies and explorations.

We could also explore the economic meaning of these dummy variables to seek the possible business mechanism in driving these individual fixed effects. One recommendation would be to compare these dummy variables with the existing grouping methods based on business judgments. The other possible path is to explore whether these dummy variables coming from the proposed algorithm can be mimicked by certain combination of known grouping criteria for a good interpretation. There might be some other ways to explore this that require some deep exploration. Finally, whether introduction of these clustering dummy variables helps models fully utilizing the contribution of existing independent variables to dependent variables is also an open topic.

Part III

CONCLUSION

Motivated by the phenomenon of the increasing proportion of zero/low-leverage firms and the concerning growing debts, we explore the evolution of the tails of leverage distribution in a 60-year time window. We find that the left tails have been growing significantly over decades while right tails have been fluctuating and also recently increasing for listed firms in real industry. This finding is valid in both NYSE market and Nasdaq market, even though increasing zero/low-leverage phenomenon is more evident in Nasdaq market while increasing high-leverage phenomenon is more evident in NYSE market.

Production flexibility is proposed to explain this phenomenon based on literature research. Since there has never been an agreement among researchers regarding the relationship between production flexibility and leverage, a theoretical model is set up to illustrate whether and how the relationship between production flexibility and leverage could change under different conditions. Three factors, - financial constraint, agency problem concern, goals of management activities -, are explored in the theoretical discussion. We find that when firms aim to maximize the value of existing shareholders, production flexibility is positively related with leverage in most cases except when there is no agency problem, though there exists financial constraint. The reason behind this is that when there is no financial constraint, high production flexibility indicates low operation risk and enables firms to take more financial risk, henceforth resulting in a positive relationship between production flexibility and leverage. However, when there is financial constraint but no agency problem concern, the money needs to be raised from debtholders is less when production flexibility is high, exhibiting a negative relationship between production flexibility and leverage. However, this relationship is changed when there exists agency problem since risk-taking behavior reduces equity value and flips the sign of the relationship between production flexibility and leverage. The story further changes when firms aim to maximize firm value. When firms aim to maximize firm value but there exists no financial constraint, firms intend to prefer external equity instead of debt, thus resulting in a negative relationship between production flexibility and leverage. However, when there exists financial constraint, the relationship turns positive as debt is the only source of funds and high production flexibility enables firms to take higher financial risk.

The logic under equity value maximization goal in the theoretical model is further confirmed by empirical results. Empirical data comes from the merged annual Compustat and Center for Research in Security Prices (CRSP) data over the period 1962–2017 with financial companies being excluded, where production flexibility is shown to aid in the decrease of leverage in general. Further segmentation based on whether firms face financial constraints proves that production flexibility plays an important role with respect to leverage when the access to financial market is constrained. Building on this, the concern of agency problem, i.e., the conflicts of interest between shareholders and debtholders, helps in explaining the different roles of production flexibility in relation to leverage. When that concern is high, production flexibility is positively related with leverage, i.e., higher production flexibility leads to higher leverage. On the other hand, when the concern is low, production flexibility is negatively related with leverage, i.e., higher agency problem concern mitigates the negative relationship between production flexibility and leverage. The findings are more evident in Nasdaq firms in in the recent 20 years. The same results are found for low-leverage decision and low-leverage transition decision. Reasonably, the roles of production flexibility with respect to both high-leverage decisions and high-leverage transition decisions are reversed. The results remain robust even after controlling more firm characteristics and macro conditions.

After the finance study on the role of production flexibility in explaining leverage phenomenon, we further turn our eyes to model prediction performance, as predicting leverage appropriately is also important for regulators and investors. Motivated by the fact that the traditional assumption of linearity between independent variables and dependent variables largely limits the prediction power of traditional panel data analysis models, and that introducing a dummy variable for each individual section while handling fixed effects is neither efficient nor necessary when some sections have very tiny differences in fixed effects, PanelGBM is designed and tested in order to group individual sections efficiently and improve model prediction performance. It is proven that the newly designed algorithm outperforms other benchmark models in terms of prediction accuracy and stability in both simulation data and real empirical data. The algorithm is also effective in grouping sections well based on fixed effects with low mis-grouping rate. Furthermore, dummy variables learned in the training process illustrate a high variable importance when compared with other finance factors. When integrating these learned dummy variables with existing financial factors to run a simple linear regression, we find that model performance improves and outperforms linear regression model with no dummies and traditional panel regression model.

The algorithm exhibits high flexibility in data pattern and broad applications to industry practice. Since the way dummy variables are introduced can be easily applied as a wrapper for other machine learning method, the algorithm is flexible to adopt in various cases with varied data formats. For example, when tackling topic inference problems, GBM can be easily replaced by Hidden Markov Model and a revised algorithm with dummy variables introduced in the same way as described in PanelGBM can be applied. Besides the flexibility in choosing underling machine learning methods, the algorithm has broad applications to practice. One user case is in finance research when dealing with panel data with short length but broad sections. Moreover, it can be used in many business practices. For instance, when analyzing or predicting a bunch of customers' behavior when they are not on record for a long time but with a large sample size, this algorithm can be applied. One detailed example is the analysis of customers' credit/fraud risk when they are still early months on book, meaning that they have not shown a long traceable history for us to use. In this case, we need to explore the patterns shared among different users to extract common patterns.

For finance study, our contribution is that we explore the role of production flexibility on leverage, a firm characteristics that has not received enough attention. We also propose a new mechanism to help understanding the co-existence of growing low-leverage and high-leverage phenomena, the literature of which was developed separately. We further develop a theory to understand the mechanism and test the theory with empirical data to gain some evidence. Finally, the study helps understanding the difference of the patterns in NYSE and Nasdaq exchange markets. One future study could be to further explore the role of market timing on the relationship between production flexibility and leverage since market timing is a very important practical factor that managers consider in debt/stock issuance decisions, and it may help explain why in our empirical model, we do not find a statistically significant role of production flexibility on leverage when firms do not face financial constraint. The other study direction could be to delve into the relationship between the tails in real industry firms and those in the financial industry firms to understand how these two big sections interact with each other with regard to the relationship between leverage and production flexibility.

The contribution of PanelGBM proposed in our study is to provide a new thought to generate clustering dummies based on individual effects and to combine machine learning methods with these dummy variables to exploit non-linearity and fixed effects at the same time for better model performance and more efficient grouping results. Future research can explore different machine learning methods with this approach to extend model scope and its application on data with different characteristics. Furthermore, we can also expand the clustering method in cross-section dimension to time series dimension to generate dummy variables for temporal fixed effects. Another potential direction of future research is to explore the economic meaning of these dummy variables. One recommendation would be to compare these dummy variables with existing grouping methods. Furthermore, accessing the contribution of these clustering dummies in they help models exploiting the contribution of existing other independent variables maybe helpful to further understand the role of these clustering dummies.

Bibliography

Abdolell, M., LeBlanc, M., Stephens, D., and Harrison, R. V. (2002). Binary partitioning for continuous longitudinal data: categorizing a prognostic variable. *Statistics in Medicine*, 21(22):3395–3409.

AD, S., DA, C., and SE, M. (1997). Cumulative and compensatory effects of competence and incompetence on depressive symptoms in children. *J Abnorm Psychol.*, 106(4):586–597.

Admati, A. R., M., P., DeMarzo, M. F. H., and Pfleiderer, P. (2018). The leverage ratchet effect. *The Journal of Finance*, 73(1):145–198.

Afshartous, D. and de Leeuw, J. (2005). Prediction in multilevel models. *Journal of Educational and Behavioral Statistic*, 3(2):109–139.

Agrawal, A. and Nagarajan, N. J. (1990). Corporate capital structure, agency costs, and ownership control: The case of all-equity firms. *The Journal of Finance*, 45(4):1325–1331.

Aivazian, V. and Berkowitz, M. (1998). Ex post production flexibility, asset specificity, and financial structure. *Journal of Accounting, Auditing and Finance*, 13(1):1–20.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4):589–609.

Anderson, R. and Carverhill, A. (2012). Corporate liquidity and capital structure. *The Review of Financial Studies*, 25(3):797–837.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: an empiricist's companion*. Princeton, NJ: Princeton University Press.

Antweiler, W. and Frank, M. (2005). Do us stock markets typically overreact to corporate news stories? Working Paper.

Baker, M. and Wurgler, J. (2002). Market timing and capital structure. *Journal of Finance*, 8(6):81–94.

Banik, P. P., Saha, R., and Kim, K.-D. (2017). Improvement of color detection by regression analysis of visual-mimo system. *2017 IEEE Globecom Workshops (GC Wkshps)*, pages 1–4.

Barr, D., Levy, Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. journal of memory and language. *HHS Public Access*, 68(3):1–43.

Bartram, S. M., Brown, G. W., and Waller, W. (2013). How important is financial risk? *Journal of Financial and Quantitative Analysis*, 50(4):801–824.

Bauer, D. and Cai, L. (2009). Consequences of unmodeled nonlinear effects in multilevel models. *Journal of Educational and Behavioral Statistics*, 34(1):97–114.

Bauer, D. and J. Cal, L. (2009). Consequences of unmodeled nonlinear effects in multilevel models. *Journal of Educational and Behavioral Statistics*, 34(1):97–114.

Beaver, W. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4:71–111.

Beaver, W. (1968a). Alternative accounting measures as predictors of failure. *Accounting Review*, 43(1):113–222.

Beaver, W. (1968b). Market prices, financial ratios, and the prediction of failure. *Journal of Accounting Research*, 6(2):179–192.

Becht, M., Bolton, P., and Roell, A. (2011). Why bank governance is different? *Oxford Review of Economic Policy*, 27(3):437–463.

Becker, G. (1981). Altruism in the family and selfishness in the market place. *Economica*, 48(189):1–15.

Berger, P. G., Ofek, E., and Yermack, D. L. (2012). Managerial entrenchment and capital structure decisions. *The Journal of Finance*, 52(4):1411–1438.

Bertrand, M. and Schoar, A. (2006). The role of family in family firms. *Journal of Economic Perspective*, 20(2):73–96.

Bertus, M., Jahera, J., and Yost, K. (2008). A note on foreign bank ownership and monitoring: An international comparison. *Journal of Banking & Finance*, 32(2):338–345.

Bharath, S., Pasquariello, P., and Wu, G. (2008). Does asymmetric information drive capital structure decisions? *The Review of Financial Studies*, 22(8):3211–3243.

Bizer, D. and DeMarzo, P. (1992). Sequential banking. *Journal of Political Economy*, 100(1):41–61.

Boot, A. and Thakor, A. (2011). Managerial autonomy, allocation of control rights, and optimal capital structure. *The Review of Financial Studies*, 24(10):3434–3485.

Bradley, M., Jarrell, G. A., and Kim, E. H. (1984). On the existence of an optimal capital structure: Theory and evidence. *The Journal of Finance*, 39(3):857–878.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24:123–140.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–35.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC.

Brennan, M. and Kraus, A. (1987). Efficient financing under asymmetric information. *Journal of Finance*, 42(5):1225–1243.

Brennan, M. and Schwartz, E. (1984). Optimal financial policy and firm valuation. *Journal of Financial Economics*, 39(3):593–607.

Burnham, K. P. and Anderson, D. R. (2003). *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach.* Springer.

Bühlmann, P. and Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statist. Sci.*, 22(4):477 – 505.

Choi, J. (2013). What drives the value premium?: the role of asset risk and leverage: The role of asset risk and leverage. 26(11):2845–2875. The Review of Financial Studies.

Cole, R. A. (2013). What do we know about the capital structure of privately held us firms? evidence from the surveys of small business finance. *Journal of Financial Management*, 4:31–42.

Dang, V. A. (2011). An empirical analysis of zero-leverage firms: Evidence from the uk. Working Paper.

Dangl, T. and Zechner, J. (2016). Debt maturity and the dynamics of leverage. working paper.

Dudley, W. C. (2010). The u.s. financial system: Where we have been, where we are and where we need to go. Accessed: 2010-02-08.

Dybvig, P. and Zender, J. (1991a). Capital structure and dividend irrelevance with asymmetric information. *Review of Financial Studies*, 4(1):201–219.

Dybvig, P. H. and Zender, J. F. (1991b). Capital structure and dividend irrelevance with asymmetric information. 4(1):201–219. The Review of Financial Studies.

D'Acunto, F., Liu, R., Pflueger, C., and Weber, M. (2017). Flexible prices and leverage. *Journal of Financial Economics*, 129(1):46–68.

Eells, E. (1991). *Probabilistic causality.* Cambridge University Press, Cambridge.

Ellul, A. and Pagano, M. (2019). Corporate leverage and employees' rights in bankruptcy. *Journal of Financial Economics*, 133(3):685–707.

Fama, E. and French, K. (2002). Testing trade-off and pecking order predictions about dividends and debt. *The Review of Financial Studies*, 15(1):1–33.

Fama, E. and French, K. (2005). Financing decisions: who issues stock? *The Journal of Financial Economics*, 76(3):549–582.

Faulkender, M. and Pertersen, M. A. (2006). Does the source of capital affect capital structure? *The Review of Financial Studies*, 19(1):45–79.

Fazzari, S., Hubbard, R., and Peterson, B. (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1(1):141–195.

Fisher, E., Heinkel, R., and Zechner, J. (1989). Dynamic capital structure choice: theory and tests. *The Journal of Finance*, 44(1):19–40.

Flannery, M. and Rangan, K. (2006). Partial adjustment toward target capital structure. *Journal of Financial Economics*, 79:469–506.

Flannery, M. J., Nikolova, S. S., and Oztekin, O. (2012). Leverage expectations and bond credit spreads. *Journal of Financial and Quantitative Analysis*, 47(4):689–714.

Fragapane, G., Ivanov, D., Peron, M., Sgarbossa, F., and Strandhagen, J. O. (2020). Increasing flexibility and productivity in industry 4.0 production networks with autonomous mobile robots and smart intralogistics. *Annals of Operations Research*.

Frank, M. and Goyal, V. (2003). Testing the pecking order theory of capital structure. *The Journal of Finance*, 67(2):217–248.

Frank, M. and Goyal, V. (2009). Capital structure decisions: Which factors are reliably important? *The Journal of Financial Management*, 38:1–37.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232.

Gelma, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, A. and Loken, E. (2014). The statistical crisis in science. *American Scientist*, 102:460–465.

Gornall, W. and Strebulaev, I. A. (2018). Financing as a supply chain: The capital structure of banks and borrowers. *Journal of Financial Economics*, 129(3):510–530.

Gorton, G. B. and Winton, A. (2002). Financial intermediation. Working Paper.

Graham, J. (2000). How big are the tax benefits of debt? *The Journal of Finance*, 55(4):1901–1942.

Graham, J. and Harvey, C. (2001). The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics*, 45(4):1019–1043.

Graham, J. R., Leary, M. T., and Roberts, M. R. (2015). A century of capital structure: The leveraging of corporate america. *Journal of Financial Economics*, 118(3):658–683.

Groll, A. and Tutz, G. (2014). Variable selection for generalized linear mixed models by l 1penalized estimation. *Statistics and Computing*, 24(2):137–154.

Hackbarth, D., Hennessy, C., and Leland, H. (2017). Can the trade-off theory explain debt structure. *The Review of Financial Studies*, 20(5):1389–1428.

Hackbarth, D. and Mauer, D. (2012). Optimal priority structure, capital structure and investment. *The Review of Financial Studies*, 25(3):747–796.

Hadlock, C. J. and Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the kz index. *The Review of Financial Studies*, 23(5):1909–1940.

Hajjem, A., Bellavance, F., and Larocque, D. (2011). Mixed effects regression trees for clustered data. *Statistics and Probability Letters*, 81(4):451–459.

Hajjem, A., Bellavance, F., and Larocque, D. (2014). Mixed effects regression trees for clustered data. *Journal of Statistical Computation and Simulation*, 81(4):451–459.

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., and Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *J Am Med Inform Assoc*, 13(1):16–23.

Harris, M. and Raviv, A. (1991). The theory of capital structure. *The Journal of Finance*, 46(1):297–355.

He, H. and Pindyck, R. S. (1992). Investments in flexible production capacity. *Journal of Economic Dynamics and Control*, 16:575–599.

Hocking, R. R. and Leslie, R. N. (1967). Selection of the best subset in regression analysis. *Technometrics*, 9(4):531–540.

Hovakimian, A., Kayhan, A., and Titman, S. (2012). Are corporate default probabilities consistent with the static trade-off theory? *Review of Financial Studies*, 25(2):315–340.

Jack, E. and Raturi, A. (2009). Measuring and comparing volume flexibility in the capital goods industry. *Production and Operations Management*, 12(4):480–501.

Jensen, M. (1986a). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, 76(2):323–329.

Jensen, M. (1986b). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review*, 76(2):323–329.

Jensen, M. and Meckling, W. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

José, P. and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.

Kahl, M., Lunn, J., and Nillson, M. (2014). Operating leverage and corporate financial policies. Working Paper.

Kahn, R. and Whited, T. M. (2018). Identification is not causality, and vice versa. *The Review of Corporate Finance Studies*, 7(1):1–21.

Kane, A., Marcus, A. J., and McDonald, R. L. (1984). How big is the tax advantage to debt? *The Journal of Finance*, 39(3):841–853.

Korajczyk, R. and Levy, A. (2003). Capital structure choice: macroeconomic conditions and financial constraints. *Journal of Financial Economics*, 68(1):75–109.

Kovenock, D. and Phillips, G. M. (1997). Capital structure and product market behavior: an examination of plant exit and investment decisions. *Review of Financial Studies*, 10(3):767–803.

Kraus, A. and Litzenberger, R. (1973). A state-preference model of optimal financial leverage. *The Journal of Finance*, 28(4):911–922.

Kulatilaka, N. (1988). Valuing the flexibility of flexible manufacturing systems. *IEEE Transactions in Engineering Management*, 35:250–257.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4):963–974.

Lamont, O., Polk, C., and Saaá-Requejo, J. (2015). Financial constraints and stock returns. *The Review of Financial Studies*, 14(2):529–554.

Leary, M. and Roberts, M. (2005). Do firms rebalance their capital structure. *The Journal of Finance*, 60(6):2575–2619.

Leary, M. and Roberts, M. (2014). Do peer firms affect corporate financial policy? *The Journal of Finance*, 69(1):139–178.

Lemmon, M., Roberts, M., and Zender, J. (2008). Back to the beginning: persistence and cross-section of corporate capital structure. *The Journal of Finance*, 63(4):1575–1608.

Lewis, D. (1973). Causation. *Journal of Philosophy*, 70:556–567.

Liu, T., Ungar, L., and Kording, K. (2021). Quantifying causality in data science with quasi-experiments. *Nature Computational Science*, 1:24–32.

Lucas, D. and McDonald, R. (1990). Equity issues and stock price dynamics. *Journal of Finance*, 45(4):1019–1043.

MacKay, P. (2003). Real flexibility and financial structure: An empirical analysis. *The Review of Financial Studies*, 16(4):1131–1165.

Maksimovic, V. and Titman, S. (1991). Financial policy and reputation for product quality. *Review of Financial Studies*, 4(1):175–200.

Malmendier, U., Tate, G., and Yan, J. (2011). Overconfidence and early- life experiences: The effect of managerial traits on corporate financial policies. *The Journal of Finance*, 66(5):1687–1733.

Mandelker, G. and Rhee, G. (1984). The impact of the degrees of operating and financial leverage on systematic risk of common stock. *Journal of Quantitative and Financial Analysis*, 19(1):45–57.

Mann, M., Zhang, Z., and Hughes, M. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proc. Natl. Acad. Sci.*, 105(36):13252–13257.

Mauer, D. C. and Triantis, A. J. (1994). Interactions of corporate financing and investment decisions: A dynamic framework. *Journal of Finance*, 49(4):1253–1277.

Miller, P., McArtor, D. B., and Lubke, G. H. (2014). Metboost: exploratory regression analysis with hierarchically clustered data. "Working Paper".

Minton, B. and Wruck, K. H. (2001). Financial conservatism: Evidence on capital structure from low leverage firms. Working Paper.

Modigliani, F. and Miller, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, 48(3):261–297.

Myers, S. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, 5(2):147–175.

Myers, S. and Majluf, N. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2):187–221.

Ni, X., Zhang, D., and Zhang, H. (2009). Variable selection for semiparametric mixed models in longitudinal studies. *Biometrics*, 66(1):79–88.

Parrino, R. and Weisbach, M. (1999). Measuring investment distortions arising from stockholder-bondholder conflicts. *Journal of Financial Economics*, 53(1):3–42.

Pearl, J. (2000). *Causality.* Cambridge University Press, Cambridge.

Pearl, J. (2019). Causal and counterfactual inference. *The handbook of rationality.*

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., and Hensley1, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*, 38(2):65–76.

Rajan, R. and Zingales, L. (1995). What do we know about capital structure? some evidence from international data. *Journal of Financial Economics*, 50(5):1421–1460.

Ramalho, J. J., Rita, R. M., and da Silva, J. V. (2018). The impact of family ownership on capital structure of firms: exploring the role of zero-leverage, size, location and the global financial crisis. *International Small Business Journal: Researching Entrepreneurship*, 36(5):574–604.

Reinartz, S. and Schmid, T. (2016). Production flexibility, product markets, and capital structure decisions. *The Review of Financial Studies*, 29(6):1501–1548.

Robertson, H. (2019). Soaring corporate debt levels threaten financial stability, imf warns. Accessed: 2019-10-16.

Rohart, F., Cristobal, M. S., and Laurent, B. (2014). Selection of fixed effects in high dimensional linear mixed models using a multicycle ecm algorithm. *Computational Statistics & Data Analysis*, 80:209–222.

Scheipl, F., Staicu, A.-M., and Greven, S. (2014). Functional additive mixed models. *J Comput Graph Stat*, 24(2):477–501.

Schelldorfer, J., Bühlmann, P., and van de Geer, S. (2010). Estimation for high-dimensional linear mixed-effects models using l_1-penalization. *Scandinavian Journal of Statistics*, 38:197–214.

Scott, S., Whitehead, B. R., Bergeman, C. S., and Pitzer, L. (2013). Combinations of stressors in midlife: Examining role and domain stressors using regression trees and random forests. *The Journals of Gerontology Series B Psychological Sciences and Social Sciences*, 68(3):464–475.

Sekhon, J. S. (2007). The neyman-rubin model of causal inference and estimation via matching methods. *The Oxford Handbook of Political Methodology.*

Sethi, A. and Sethi, S. (1990). Flexibility in manufacturing: A survey. *International Journal of Flexible Manufacturing Systems*, 2:289–328.

Shyam-Sunder, L. and Myers, S. C. (1999). Testing static tradeoff against pecking order models of capital structure. *The Journal of Financial Economics*, 51(2):219–244.

Skrondal, A. and Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, 68(2):267–287.

Strebulaev, I. A. and Yang, B. (2013). The mystery of zero-leverage firms. *Journal of Financial Economics*, 109(1):1–23.

Stulz, R. (1990). Managerial discretion and optimal financing policies. *Journal of Financial Economics*, 26:3–27.

Takayama, A. (1985). *Mathematical Economics (Second ed.)*. New York: Cambridge University Press.

Tang, Y., Gao, R., Han, S., Chen, Y., and Gao, D. (2021). Body part regression with self-supervision. *IEEE Transactions on Medical Imaging*, 40(5):1499–1507.

Titman, S. (1984). The effect of capital structure on a firm's liquidation decision. *Journal of Financial Economics*, 13(1):137–151.

Titman, S. and Roberto, W. (1988). The determinants of capital structure choice. *The Journal of Finance*, 43(1):1–19.

Torry, H. and Goldfarb, S. (2017). Imf warns high corporate leverage could threaten financial stability. Accessed: 2017-04-19.

Triantis, A. J. and Hodder, J. E. (1990). Valuing flexibility as a complex option. *Journal of Finance*, 45:549–565.

Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist*, 24(2):83–91.

Verma, P., Anwar, S., Khan, S., and Mane, D. S. B. (2018). Network intrusion detection using clustering and gradient boosting. Working Paper.

Wallace, K. A., Bergeman, C. S., and Maxwell, S. E. (2002). Predicting well-being outcomes in later life: An application of classification and regression tree (cart) analysis. *Advances in psychology research*, 17:71–92.

Wang, J., Qin, B., Zhang, W., and Shi, W. (2011). Regression rules extraction from artificial neural network based on least squares. *2011 Seventh International Conference on Natural Computation*, pages 203–207.

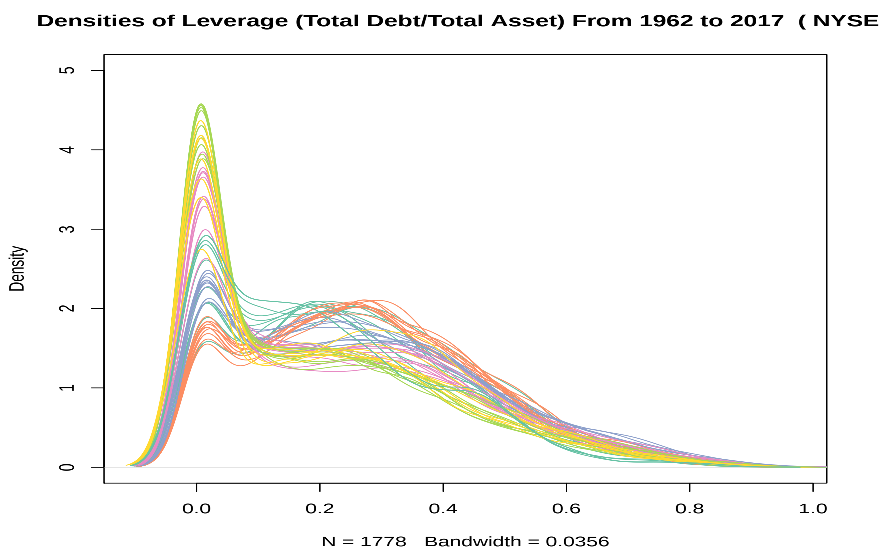
Wood, S. N. (2006). *Generalized Additive Models: An Introduction With R.* Boca Raton, Fl. : Chapman and Hall/CRC Press.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Wu, J., Gonzalez, R. E., Lan, Z., Gnedin, N. Y., Kravtsov, A. V., Rudd, D. H., and Yu, Y. (2011). Performance emulation of cell-based amr cosmology simulations. *2011 IEEE International Conference on Cluster Computing*.

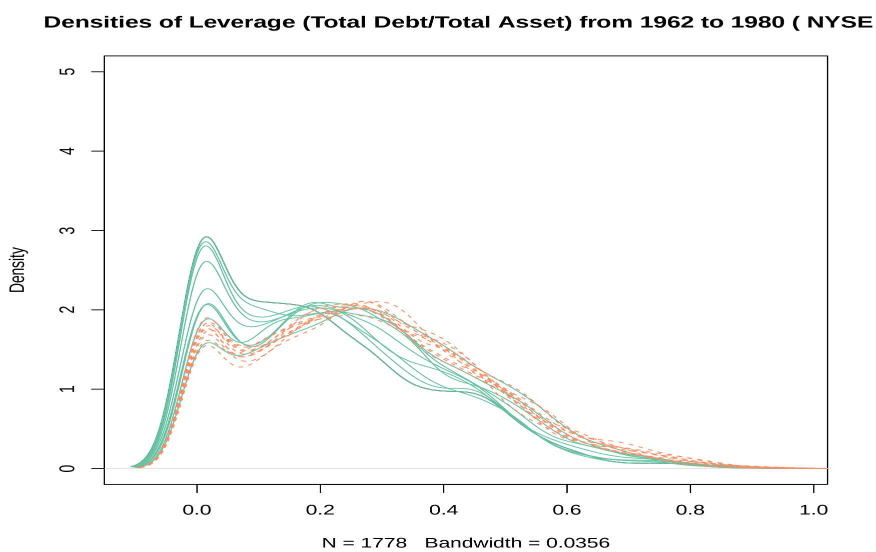
Appendix A

FIGURES



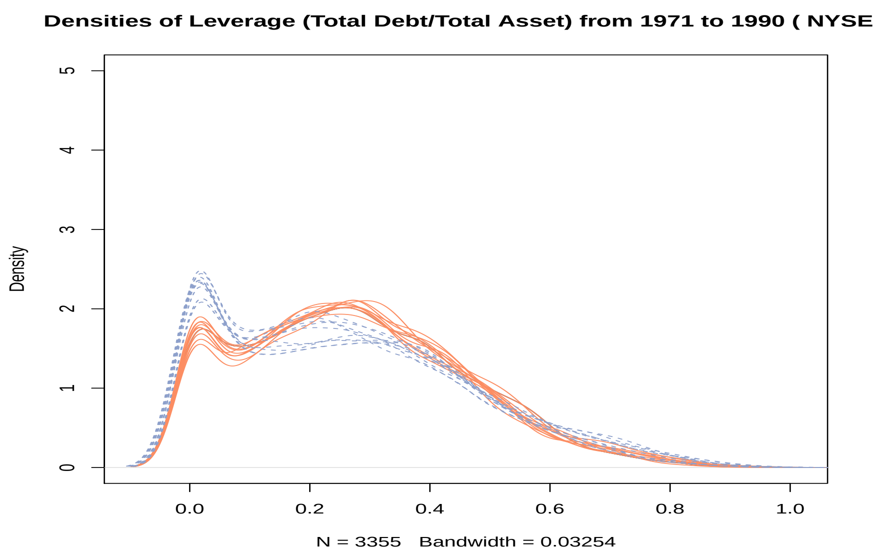
Densities of Leverage (Total Debt/Total Asset) From 1962 to 2017 (NYSE
Density
N = 1778 Bandwidth = 0.0356

(a)

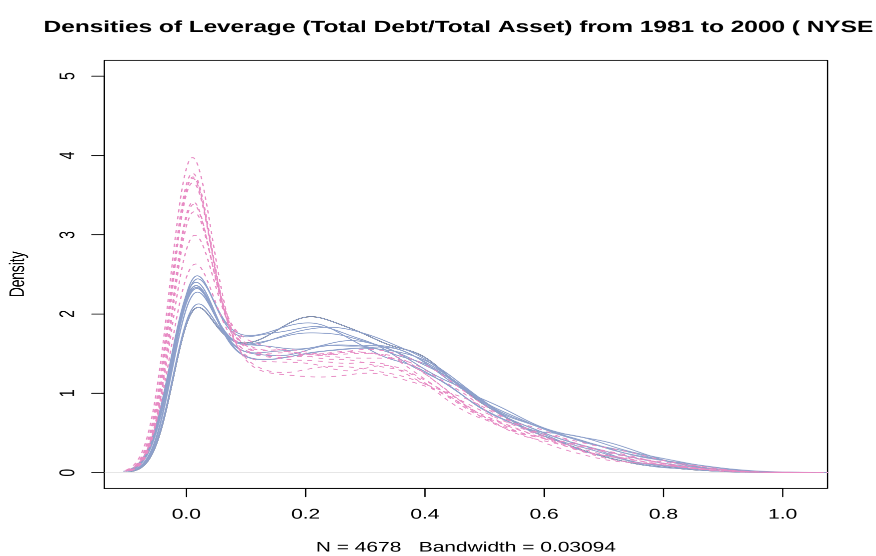


Densities of Leverage (Total Debt/Total Asset) from 1962 to 1980 (NYSE
Density
N = 1778 Bandwidth = 0.0356

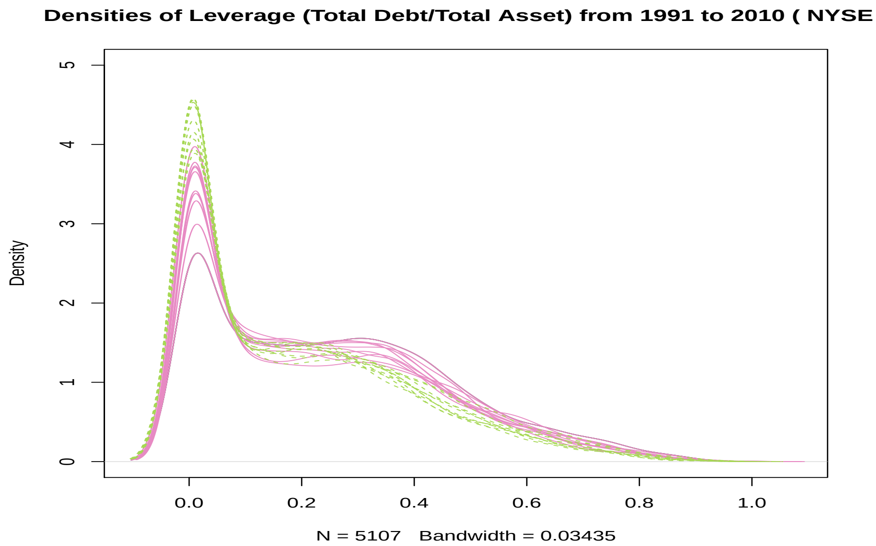
(b)


Densities of Leverage (Total Debt/Total Asset) from 1971 to 1990 (NYSE
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 3355 Bandwidth = 0.03254

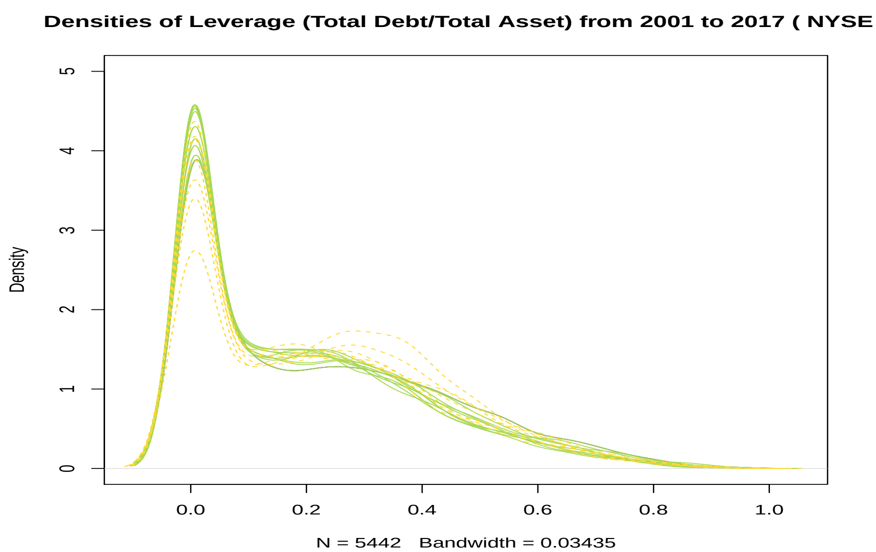
(c)


Densities of Leverage (Total Debt/Total Asset) from 1981 to 2000 (NYSE
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 4678 Bandwidth = 0.03094

(d)



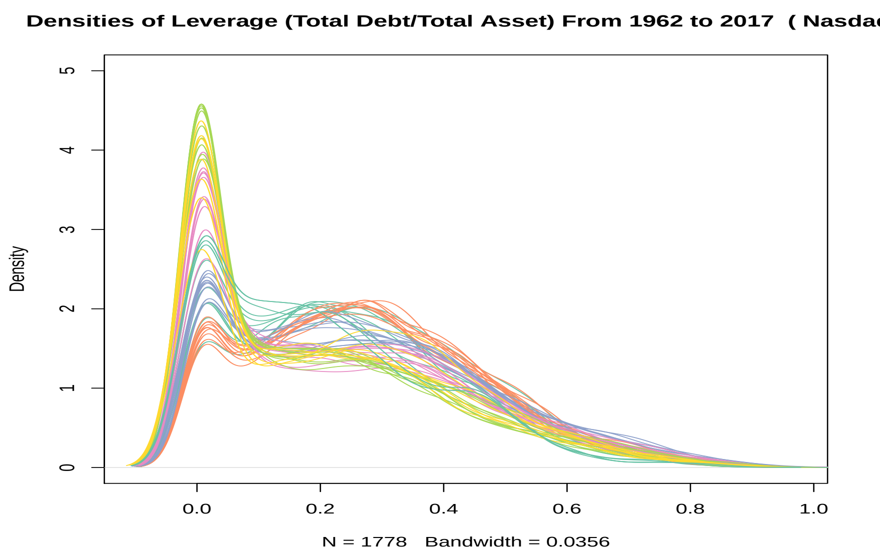
(e)



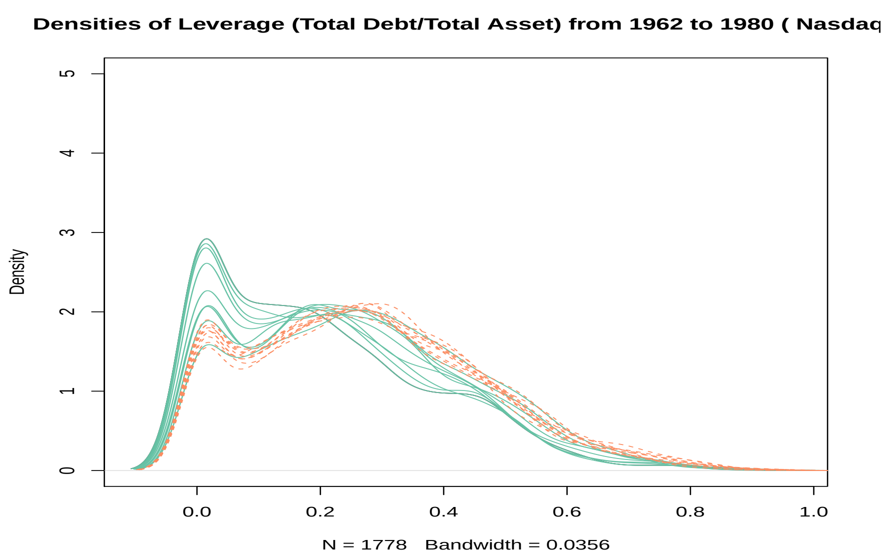
(f)

Figure A.1: Rolling comparison of leverage (total debt/total asset) from 1962 to 2017 for NYSE firms

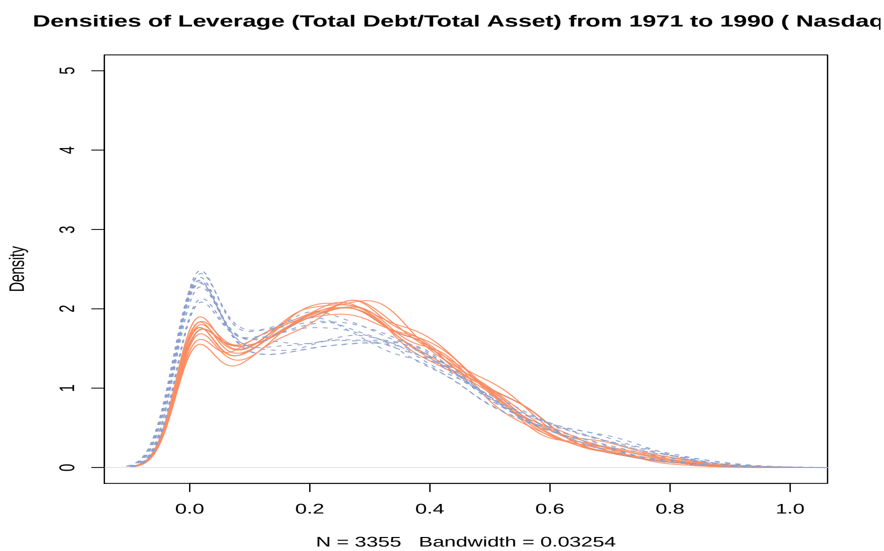
This figure exhibits a rolling comparison of leverage distributions every ten years. The leverage here is measured as the ratio of total debt to total asset. One line in each figure represents a leverage distribution of a certain year. The 57 years starting from 1962 to 2017 are grouped into six groups: "1962 to 1970", "1971 to 1980", "1981 to 1990", "1991 to 2000", "2001 to 2010", and "2011 to 2017". Years classified as one group are represented using a same color. The general comparison is shown in (a), and the comparison between every two ten years are shown in (b) to (f). Details about density color: densities of leverage from 1962 to 1970 are in lime green, densities of leverage from 1971 to 1980 are in orange, densities of leverage from 1981 to 1990 are in blue, densities of leverage from 1991 to 2000 are in pink. densities of leverage from 2001 to 2010 are in moderate green, densities of leverage from 2011 to 2017 are in yellow.


Densities of Leverage (Total Debt/Total Asset) From 1962 to 2017 (Nasda
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1778 Bandwidth = 0.0356

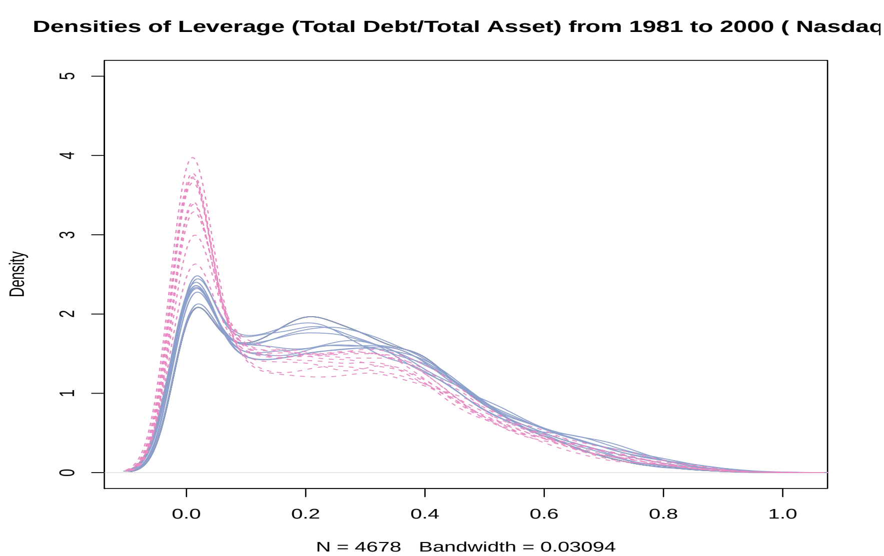
(a)


Densities of Leverage (Total Debt/Total Asset) from 1962 to 1980 (Nasdaq
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1778 Bandwidth = 0.0356

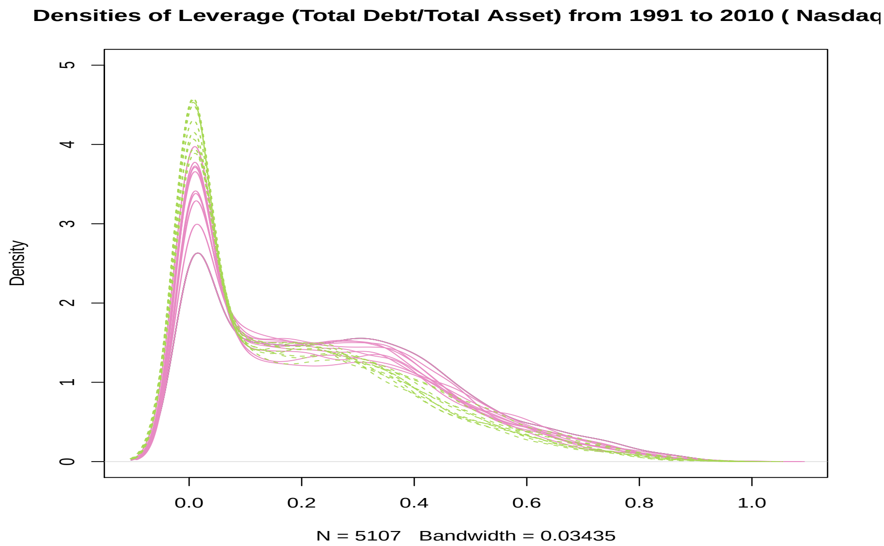
(b)


Densities of Leverage (Total Debt/Total Asset) from 1971 to 1990 (Nasdaq
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 3355 Bandwidth = 0.03254

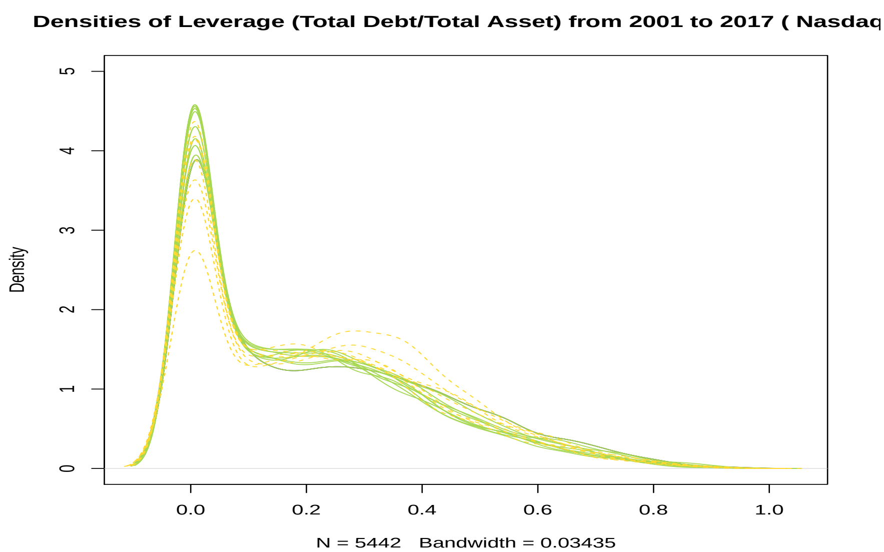
(c)


Densities of Leverage (Total Debt/Total Asset) from 1981 to 2000 (Nasdaq
Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
N = 4678 Bandwidth = 0.03094

(d)



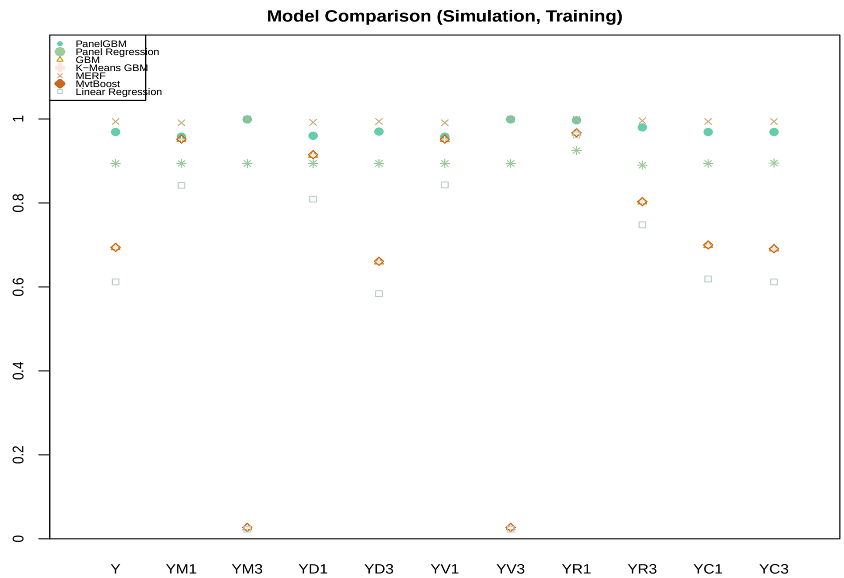
(e)



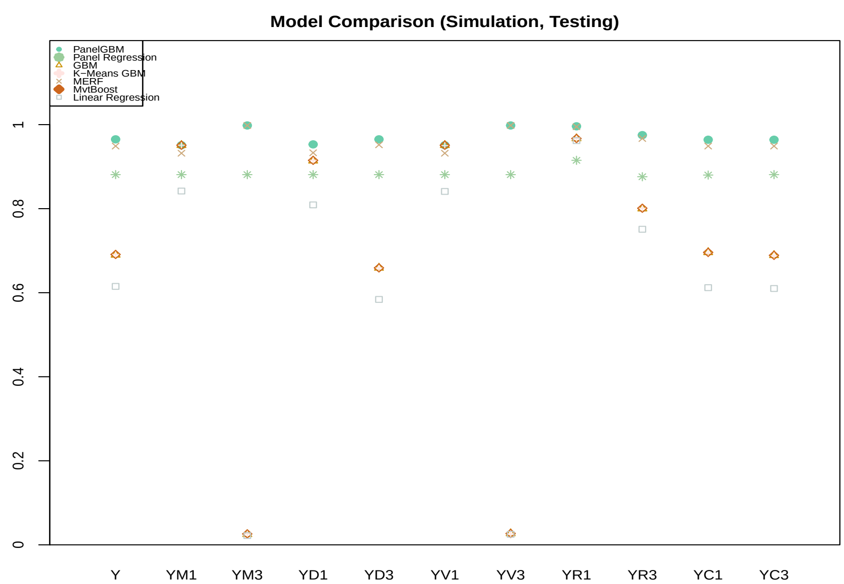
(f)

Figure A.2: Rolling comparison of leverage (total debt/total asset) from 1962 to 2017 for Nasdaq firms

This figure exhibits a rolling comparison of leverage distributions every ten years. The leverage here is measured as the ratio of total debt to total asset. One line in each figure represents a leverage distribution of a certain year. The 57 years starting from 1962 to 2017 are grouped into six groups: "1962 to 1970", "1971 to 1980", "1981 to 1990", "1991 to 2000", "2001 to 2010", and "2011 to 2017". Years classified as one group are represented using the same color. The general comparison is shown in (a), and the comparison between every two ten years are shown in (b) to (f). Details about density color: densities of leverage from 1962 to 1970 are in lime green, densities of leverage from 1971 to 1980 are in orange, densities of leverage from 1981 to 1990 are in blue, densities of leverage from 1991 to 2000 are in pink. densities of leverage from 2001 to 2010 are in moderate green, densities of leverage from 2011 to 2017 are in yellow.


Model Comparison (Simulation, Training)
PanelGBM
Panel Regression
GBM
K−Means GBM
MERF
MvtBoost
Linear Regression
0
0.2
0.4
0.6
0.8
1
Y
YM1
YM3
YD1
YD3
YV1
YV3
YR1
YR3
YC1
YC3

(a)


Model Comparison (Simulation, Testing)
PanelGBM
Panel Regression
GBM
K−Means GBM
MERF
MvtBoost
Linear Regression
0
0.2
0.4
0.6
0.8
1
Y
YM1
YM3
YD1
YD3
YV1
YV3
YR1
YR3
YC1
YC3

(b)

(c)

Figure A.3: Model comparison (simulation)

This figure demonstrates the comparison of model performance. Model performance refers to R-squared: “Linear Regression,” “Panel Regression,” “GBM”, “K-means GBM,” “MERF”, “MvtBoost”, and “PanelGBM.” From the top to the bottom, we illustrate the model comparison in training and testing the dataset; the last one comprises the model deterioration among different models, as mentioned above.

Appendix B

TABLES

Table B.1: Debt with different measures across years

Column N shows the number of firms in the sample. Column Total Liability (Mean), Total Debt (Mean), and Long-term Debt (Mean) reports the average debt measured using total liability, total debt, and total long-term debts, respectively, while Column Total Liability (Sum), Total Debt (Sum), and Long-term Debt (Sum) report the summation property of debts in different measures. All these dollar values are measured in million dollars

Year	N	Total Liability (Mean)	Total Debt (Mean)	Long-term Debt (Mean)	Total Liability (Sum)	Total Debt (Sum)	Long-term Debt (Sum)
1962	1778	71.45	38.34	34.44	127,038.79	68,164.96	61,236.67
1963	1948	68.76	37.29	33.48	133,951.99	72,631.85	65,216.72
1964	2102	72.55	39.15	34.96	152,491.83	82,302.27	73,492.87
1965	2249	77.20	41.52	36.41	173,632.68	93,386.50	81,895.15
1966	2404	85.50	47.44	40.51	205,542.89	114,050.08	97,397.88
1967	2532	93.35	53.94	45.66	236,349.61	136,587.65	115,602.85
1968	3025	93.32	54.53	46.01	282,284.99	164,949.41	139,188.04
1969	3286	99.71	59.71	48.30	327,646.92	196,215.26	158,719.71
1970	3300	109.08	68.48	54.91	359,975.32	225,997.27	181,205.64
1971	3355	115.57	71.71	59.48	387,753.34	240,598.36	199,543.10
1972	3445	122.50	74.67	62.68	422,029.34	257,227.20	215,922.16
1973	3477	139.56	82.04	67.40	485,254.48	285,248.81	234,352.28
1974	4872	120.98	69.89	54.92	589,420.75	340,494.43	267,564.07
1975	4871	126.78	72.21	60.36	617,526.53	351,722.96	294,028.53
1976	4851	138.31	74.72	63.89	670,932.42	362,490.17	309,914.21
1977	4808	153.27	80.35	68.82	736,929.67	386,316.76	330,895.69
1978	4669	178.31	90.71	77.07	832,517.76	423,517.69	359,838.32
1979	4569	209.89	103.72	86.11	959,004.40	473,874.73	393,445.88
1980	4621	232.53	113.34	94.34	1,074,524.46	523,747.02	435,923.59
1981	4678	253.67	123.50	102.58	1,186,661.82	577,746.04	479,887.59
1982	4800	266.65	131.18	111.89	1,279,918.00	629,676.17	537,058.56
1983	4994	283.11	132.58	114.85	1,413,851.99	662,106.08	573,558.74
1984	5063	309.02	144.68	122.66	1,564,573.86	732,521.21	621,025.76
1985	5192	312.19	146.72	121.10	1,620,890.54	761,771.03	628,776.39
1986	5362	330.52	160.03	134.51	1,772,223.01	858,067.89	721,231.69
1987	5412	358.46	170.32	142.20	1,940,002.92	921,745.92	769,565.72
1988	5234	467.12	242.43	178.54	2,444,910.31	1,268,877.74	934,481.68
1989	5061	530.28	281.38	205.19	2,683,758.39	1,424,053.47	1,038,473.24
1990	5008	570.67	302.38	216.14	2,857,931.12	1,514,331.36	1,082,416.86
1991	5107	575.94	300.42	217.01	2,941,325.74	1,534,237.15	1,108,253.77
1992	5393	579.96	285.75	205.58	3,127,721.89	1,541,054.12	1,108,690.36
1993	5773	588.37	274.71	194.82	3,396,685.79	1,585,875.04	1,124,683.88

Continuation of Table B.1:

Year	N	Total Liability (Mean)	Total Debt (Mean)	Long-term Debt (Mean)	Total Liability (Sum)	Total Debt (Sum)	Long-term Debt (Sum)
1994	6060	569.06	261.93	195.33	3,448,526.26	1,587,317.44	1,183,720.12
1995	6640	557.79	265.52	196.16	3,703,702.55	1,763,080.53	1,302,522.68
1996	6927	585.14	278.94	208.91	4,053,240.30	1,932,185.25	1,447,102.66
1997	6717	653.09	319.30	239.19	4,386,793.50	2,144,771.10	1,606,643.54
1998	6663	724.28	360.06	270.35	4,825,875.34	2,399,073.77	1,801,327.48
1999	6569	829.50	417.41	311.11	5,448,996.41	2,741,997.32	2,043,696.15
2000	6153	985.95	487.15	363.76	6,066,523.36	2,997,453.98	2,238,240.07
2001	5442	1119.24	567.87	435.44	6,090,897.78	3,090,355.75	2,369,676.73
2002	5085	1214.27	598.70	475.46	6,174,538.35	3,044,412.99	2,417,708.57
2003	4895	1318.72	630.36	512.85	6,455,142.13	3,085,595.55	2,510,395.01
2004	4803	1386.90	648.49	515.40	6,661,257.03	3,114,721.27	2,475,469.49
2005	4635	1461.36	663.62	520.96	6,773,391.10	3,075,890.64	2,414,650.52
2006	4488	1468.10	649.60	531.21	6,588,833.52	2,915,391.18	2,384,059.88
2007	4331	1700.22	787.85	642.25	7,363,632.50	3,412,173.07	2,781,603.30
2008	4018	1786.57	836.67	688.06	7,178,446.76	3,361,757.03	2,764,617.66
2009	3782	1928.56	886.67	767.84	7,293,811.46	3,353,395.14	2,903,980.71
2010	3716	2055.58	917.20	783.96	7,638,534.30	3,408,303.52	2,913,191.12
2011	3660	2303.95	1023.31	863.43	8,432,451.10	3,745,296.85	3,160,170.07
2012	3662	2454.79	1110.94	950.65	8,989,445.89	4,068,271.14	3,481,297.39
2013	3677	2580.94	1210.73	1046.01	9,490,111.20	4,451,836.67	3,846,182.48
2014	3619	2818.98	1346.47	1165.94	10,201,899.50	4,872,879.95	4,219,527.91
2015	3401	3121.20	1545.59	1364.54	10,615,208.92	5,256,543.66	4,640,796.53
2016	3181	3439.81	1731.47	1531.05	10,942,047.33	5,507,805.67	4,870,255.89
2017	2203	5102.65	2629.21	2308.53	11,241,134.45	5,792,144.83	5,085,690.60

End of Table

Table B.2: Leverage and frequency of zero-leverage and high-leverage firms (NYSE)

Column N gives the number of firms in the sample. Column Lever_TLA , Lever_TDA, and Lever_LDA report the average leverage measured using total liability, total debt, and total long-term debts, respectively. ZL firms are firms that have zero book debt(DLTT + DLC = 0, where DLTT and DLC are the Compustat long-term debt and the debt in current liabilities, respectively). ZLTD firms have zero long-term debt (DLTT = 0). Almost zero-leverage (AZTD) firms have a book leverage that does not exceed 5% in a given year. NPND firms have non-positive net debt in a given year, i.e., $DLTT + DLC - CHE \leq 0$, where CHE refers to the Compustat cash holdings. Columns ZL, ZLTD, AZTD, and NPND report the corresponding fractions of firms relative to the total sample in each year, while HL, HLTD, and HLL report the percentage of firms that have measured leverage using total debt, long-term debt, and total liability exceeding 0.9 in each year.

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1962	1,296	0.20	0.16	0.41	24.23%	14.12%	19.68%	45.76%	0.00%	0.00%	0.54%
1963	1,386	0.21	0.17	0.41	22.80%	14.14%	18.90%	44.30%	0.00%	0.00%	0.36%
1964	1,471	0.21	0.17	0.42	22.23%	13.26%	18.83%	41.20%	0.07%	0.00%	0.48%
1965	1,542	0.22	0.18	0.44	19.97%	11.54%	16.34%	37.42%	0.13%	0.00%	0.52%
1966	1,639	0.24	0.19	0.46	16.72%	9.76%	14.46%	32.21%	0.18%	0.00%	0.67%
1967	1,697	0.25	0.20	0.46	15.62%	8.31%	12.61%	28.93%	0.06%	0.00%	0.65%
1968	1,822	0.26	0.20	0.47	14.87%	7.41%	11.14%	28.27%	0.00%	0.00%	0.66%
1969	1,918	0.26	0.20	0.47	12.93%	5.79%	9.59%	26.69%	0.10%	0.00%	0.57%
1970	1,928	0.29	0.21	0.48	10.58%	4.62%	7.94%	21.11%	0.00%	0.00%	0.62%
1971	1,956	0.28	0.21	0.48	11.30%	4.75%	8.03%	23.21%	0.05%	0.00%	0.77%
1972	2,010	0.27	0.21	0.48	12.19%	5.22%	8.26%	24.73%	0.05%	0.00%	1.04%
1973	2,015	0.28	0.21	0.50	10.42%	4.37%	7.74%	16.72%	0.00%	0.00%	0.55%
1974	2,064	0.29	0.21	0.52	8.62%	3.29%	6.40%	13.71%	0.05%	0.05%	1.07%
1975	2,027	0.28	0.21	0.50	10.56%	3.80%	6.56%	16.67%	0.05%	0.00%	0.84%
1976	1,994	0.26	0.21	0.50	11.33%	4.36%	7.02%	20.31%	0.05%	0.00%	1.05%
1977	1,934	0.26	0.21	0.50	12.10%	4.71%	6.67%	20.01%	0.00%	0.00%	0.93%
1978	1,849	0.26	0.21	0.51	10.65%	4.00%	5.62%	18.33%	0.00%	0.00%	0.76%
1979	1,787	0.27	0.21	0.53	9.74%	3.64%	5.65%	16.45%	0.00%	0.00%	0.95%
1980	1,764	0.26	0.21	0.52	10.03%	3.74%	5.61%	16.84%	0.06%	0.00%	1.08%
1981	1,713	0.25	0.20	0.51	10.51%	3.74%	5.55%	19.56%	0.00%	0.00%	0.82%
1982	1,690	0.26	0.21	0.51	11.30%	4.08%	5.80%	20.65%	0.00%	0.00%	1.48%
1983	1,704	0.25	0.20	0.51	13.62%	4.58%	6.22%	26.12%	0.06%	0.00%	1.76%
1984	1,648	0.25	0.20	0.51	12.26%	4.19%	5.89%	23.85%	0.00%	0.00%	0.97%
1985	1,626	0.25	0.20	0.52	12.73%	4.31%	6.09%	23.00%	0.06%	0.06%	1.35%
1986	1,610	0.27	0.21	0.53	12.80%	4.91%	7.39%	23.23%	0.00%	0.00%	2.86%
1987	1,607	0.27	0.22	0.54	12.26%	4.67%	7.41%	21.78%	0.12%	0.12%	2.43%
1988	1,551	0.28	0.22	0.54	11.67%	4.45%	6.51%	21.21%	0.00%	0.00%	2.90%
1989	1,519	0.29	0.23	0.55	12.24%	5.13%	7.37%	20.93%	0.07%	0.00%	3.69%
1990	1,521	0.29	0.23	0.55	13.02%	5.46%	7.69%	21.17%	0.13%	0.00%	3.48%
1991	1,568	0.28	0.22	0.54	14.80%	6.51%	9.50%	22.96%	0.06%	0.00%	2.87%

Continuation of Table B.3:

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1992	1,616	0.27	0.22	0.54	15.72%	6.68%	10.15%	23.14%	0.06%	0.06%	2.85%
1993	1,669	0.26	0.21	0.54	15.64%	7.55%	10.67%	23.31%	0.12%	0.06%	2.64%
1994	1,697	0.26	0.22	0.55	15.20%	6.78%	8.90%	21.69%	0.00%	0.00%	2.53%
1995	1,782	0.27	0.22	0.55	15.82%	6.85%	9.71%	21.55%	0.00%	0.00%	2.81%
1996	1,810	0.26	0.22	0.54	16.57%	6.57%	9.61%	23.31%	0.06%	0.00%	2.27%
1997	1,731	0.27	0.22	0.54	15.31%	7.28%	9.88%	22.13%	0.00%	0.00%	2.37%
1998	1,663	0.28	0.24	0.55	15.27%	7.04%	9.80%	21.05%	0.06%	0.00%	2.77%
1999	1,562	0.29	0.24	0.56	14.40%	6.85%	9.99%	18.25%	0.06%	0.06%	2.75%
2000	1,440	0.28	0.23	0.56	14.58%	7.22%	10.63%	19.93%	0.07%	0.00%	3.13%
2001	1,352	0.28	0.23	0.55	15.38%	6.80%	9.76%	21.60%	0.07%	0.00%	2.22%
2002	1,350	0.28	0.23	0.56	15.85%	7.11%	10.30%	22.59%	0.15%	0.15%	4.00%
2003	1,380	0.26	0.23	0.55	17.46%	8.33%	10.72%	26.01%	0.14%	0.00%	2.61%
2004	1,388	0.24	0.21	0.53	18.52%	8.21%	11.10%	27.88%	0.07%	0.07%	2.59%
2005	1,376	0.24	0.20	0.53	18.97%	9.38%	11.63%	29.22%	0.00%	0.00%	3.27%
2006	1,366	0.23	0.20	0.53	19.33%	10.61%	13.25%	29.14%	0.15%	0.07%	2.78%
2007	1,338	0.23	0.20	0.51	21.38%	11.88%	15.10%	30.34%	0.07%	0.00%	1.72%
2008	1,296	0.25	0.22	0.54	19.75%	11.11%	14.89%	28.16%	0.00%	0.00%	3.01%
2009	1,279	0.24	0.21	0.53	20.80%	11.34%	14.15%	31.27%	0.00%	0.00%	2.81%
2010	1,283	0.23	0.20	0.52	21.75%	11.15%	15.04%	32.81%	0.00%	0.00%	2.42%
2011	1,314	0.24	0.21	0.53	21.84%	11.34%	14.38%	30.37%	0.15%	0.08%	3.58%
2012	1,334	0.24	0.21	0.53	21.29%	11.69%	14.99%	30.88%	0.00%	0.00%	3.52%
2013	1,361	0.24	0.22	0.52	21.82%	12.78%	15.28%	31.59%	0.07%	0.00%	2.50%
2014	1,360	0.26	0.23	0.54	19.19%	11.40%	13.82%	28.38%	0.00%	0.00%	3.60%
2015	1,316	0.28	0.25	0.55	17.63%	10.79%	13.53%	26.82%	0.08%	0.00%	3.57%
2016	1,251	0.28	0.26	0.56	16.39%	10.15%	12.71%	25.50%	0.24%	0.24%	4.24%
2017	1,034	0.30	0.27	0.57	13.06%	7.64%	9.38%	21.57%	0.19%	0.19%	4.74%
Mean	1,593	0.26	0.21	0.52	15.41%	7.45%	10.39%	25.10%	0.06%	0.02%	2.07%

End of Table

Table B.3: Leverage and frequency of zero-leverage and high-leverage firms (Nasdaq)

Column N gives the number of firms in the sample. Column Lever_TLA , Lever_TDA, and Lever_LDA report the average leverage measured using total liability, total debt, and total long-term debts, respectively. ZL firms have zero book debt (DLTT + DLC = 0, where DLTT and DLC are the Compustat long-term debt and debt in current liabilities, respectively). ZLTD firms have zero long-term debt (DLTT = 0). Almost all zero-leverage (AZTD) firms have a book leverage not exceeding 5% in a given year. NPND firms have non-positive net debt in a given year, i.e., $DLTT + DLC - CHE \leq 0$, where CHE is the Compustat cash holdings. Columns ZL, ZLTD, AZTD, and NPND report corresponding fractions of firms relative to the total sample in each year, while HL, HLTD, HLL and reports the percentage of firms that have their leverage measured using total debt, long-term debt, and total liability exceeding 0.9 in each year.

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1962	63	0.22	0.17	0.42	23.81%	17.46%	20.63%	55.56%	0.00%	0.00%	0.00%
1963	88	0.21	0.16	0.41	23.86%	14.77%	21.59%	47.73%	0.00%	0.00%	0.00%
1964	100	0.21	0.17	0.42	25.00%	18.00%	23.00%	41.00%	0.00%	0.00%	0.00%
1965	108	0.21	0.16	0.43	22.22%	14.81%	21.30%	45.37%	0.00%	0.00%	0.00%
1966	118	0.22	0.16	0.44	16.95%	10.17%	18.64%	35.59%	0.00%	0.00%	0.00%
1967	128	0.24	0.18	0.45	15.63%	10.16%	14.06%	31.25%	0.00%	0.00%	0.00%
1968	182	0.23	0.17	0.44	19.78%	12.09%	15.38%	35.16%	0.00%	0.00%	0.00%
1969	208	0.24	0.17	0.44	17.31%	10.10%	14.42%	32.21%	0.00%	0.00%	0.00%
1970	219	0.25	0.17	0.44	15.53%	9.13%	11.87%	29.68%	0.00%	0.00%	0.91%
1971	239	0.24	0.17	0.43	20.08%	10.88%	13.81%	30.54%	0.00%	0.00%	0.84%
1972	272	0.23	0.16	0.43	22.06%	12.13%	16.18%	32.35%	0.00%	0.00%	1.10%
1973	294	0.24	0.17	0.46	18.37%	11.22%	16.33%	26.87%	0.00%	0.00%	0.68%
1974	482	0.27	0.17	0.49	16.39%	6.64%	14.11%	22.20%	0.00%	0.00%	1.24%
1975	500	0.26	0.18	0.48	18.60%	7.80%	13.40%	25.40%	0.20%	0.00%	1.40%
1976	517	0.25	0.18	0.48	18.57%	8.70%	13.93%	29.01%	0.58%	0.19%	1.16%
1977	533	0.24	0.17	0.49	17.82%	8.44%	13.88%	27.77%	0.38%	0.19%	1.50%
1978	548	0.25	0.19	0.50	16.97%	8.21%	12.77%	25.18%	0.18%	0.00%	1.64%
1979	575	0.26	0.19	0.51	15.65%	6.26%	10.78%	24.35%	0.17%	0.00%	1.04%
1980	646	0.25	0.19	0.50	16.10%	7.28%	11.92%	24.92%	0.00%	0.00%	1.24%
1981	715	0.23	0.17	0.47	20.42%	8.81%	13.85%	33.57%	0.00%	0.00%	1.82%
1982	827	0.23	0.18	0.47	20.07%	8.46%	10.88%	33.98%	0.00%	0.00%	1.33%
1983	960	0.20	0.16	0.44	28.02%	10.63%	13.65%	44.06%	0.00%	0.00%	1.25%
1984	1,042	0.21	0.16	0.44	28.02%	10.27%	14.20%	42.13%	0.00%	0.00%	1.63%
1985	1,168	0.23	0.17	0.46	23.20%	8.90%	13.36%	37.59%	0.09%	0.09%	2.31%
1986	1,274	0.24	0.17	0.47	24.02%	9.97%	14.29%	40.74%	0.00%	0.00%	2.43%
1987	1,308	0.23	0.17	0.46	24.16%	10.02%	14.53%	40.90%	0.00%	0.00%	2.45%
1988	1,294	0.23	0.17	0.47	24.03%	9.66%	14.91%	40.73%	0.00%	0.00%	3.86%
1989	1,268	0.23	0.16	0.46	25.63%	10.80%	16.56%	41.17%	0.24%	0.08%	2.21%
1990	1,358	0.22	0.16	0.46	27.98%	12.74%	18.11%	43.15%	0.15%	0.07%	2.80%
1991	1,479	0.21	0.15	0.45	30.29%	13.46%	18.86%	46.79%	0.07%	0.00%	3.58%

Continuation of Table B.3:

Year	N	Lever_TDA	Lever_LDA	Lever_TLA	AZTD	ZL	ZLTD	NPND_D	HTD	HLD	HTL
1992	1,667	0.19	0.14	0.43	33.89%	14.94%	20.04%	48.41%	0.12%	0.12%	2.40%
1993	1,849	0.18	0.13	0.42	35.97%	16.71%	23.15%	50.89%	0.05%	0.05%	2.60%
1994	2,016	0.18	0.13	0.42	35.91%	16.91%	23.31%	50.79%	0.10%	0.10%	1.93%
1995	2,327	0.18	0.13	0.43	36.40%	18.48%	24.84%	51.96%	0.17%	0.04%	2.79%
1996	2,426	0.17	0.12	0.40	41.34%	19.70%	26.75%	55.69%	0.04%	0.04%	2.14%
1997	2,389	0.16	0.12	0.40	43.32%	21.10%	27.46%	57.47%	0.08%	0.08%	1.84%
1998	2,504	0.16	0.13	0.41	42.45%	21.41%	28.31%	59.03%	0.08%	0.04%	2.40%
1999	2,432	0.16	0.12	0.40	44.04%	20.76%	27.88%	59.99%	0.08%	0.04%	1.93%
2000	2,227	0.16	0.12	0.38	46.92%	21.24%	28.38%	60.48%	0.00%	0.00%	1.71%
2001	2,051	0.15	0.12	0.38	46.42%	23.45%	30.38%	61.53%	0.05%	0.00%	2.00%
2002	2,049	0.15	0.12	0.40	45.78%	24.26%	31.53%	62.37%	0.00%	0.00%	2.59%
2003	2,053	0.15	0.12	0.40	46.96%	27.72%	34.83%	63.71%	0.00%	0.00%	2.29%
2004	2,015	0.14	0.11	0.39	48.98%	28.54%	34.59%	64.86%	0.05%	0.00%	2.13%
2005	1,986	0.13	0.11	0.39	51.66%	29.76%	36.66%	66.21%	0.00%	0.00%	2.22%
2006	1,904	0.13	0.11	0.39	49.42%	28.94%	35.82%	65.49%	0.00%	0.00%	1.68%
2007	1,805	0.14	0.11	0.40	49.42%	29.03%	36.34%	64.82%	0.00%	0.00%	1.77%
2008	1,721	0.16	0.12	0.43	46.08%	28.47%	36.03%	61.01%	0.00%	0.00%	3.25%
2009	1,651	0.14	0.11	0.41	48.27%	29.56%	37.67%	65.60%	0.00%	0.00%	2.91%
2010	1,602	0.13	0.11	0.41	50.62%	30.65%	39.39%	68.23%	0.06%	0.06%	2.68%
2011	1,597	0.14	0.11	0.41	48.22%	31.62%	39.82%	66.69%	0.00%	0.00%	2.13%
2012	1,606	0.15	0.12	0.43	46.01%	31.07%	38.54%	64.88%	0.06%	0.06%	2.93%
2013	1,628	0.15	0.12	0.42	45.82%	31.08%	37.78%	66.03%	0.12%	0.06%	3.07%
2014	1,651	0.16	0.14	0.42	43.85%	30.59%	36.95%	63.96%	0.00%	0.00%	3.09%
2015	1,626	0.17	0.14	0.42	43.23%	29.77%	36.35%	63.16%	0.06%	0.00%	3.26%
2016	1,579	0.18	0.15	0.45	40.60%	28.56%	35.15%	59.78%	0.06%	0.00%	3.61%
2017	1,037	0.19	0.17	0.46	36.93%	25.75%	31.05%	55.54%	0.00%	0.00%	3.28%
Mean	1,213	0.20	0.15	0.44	31.88%	17.46%	23.22%	47.24%	0.06%	0.02%	1.84%

End of Table

Table B.4: Detailed definition of variables used for empirical studies

The table below shows the detailed definition of variables utilized in empirical studies.

Variable	Variable Description
	Panel A: Independent variables
$Lever\_TDA$	$Lever\_TDA = \frac{TotalDebt}{TotalAsset}$, where $TotalDebt$= long term debt and debt in current liabilities. Long term debt is Compustat item "DLTT" and debt in current liabilities is Comp $TotalAsset$ comes from "AT" from Compustat.
$AZTD$	AZTD $= \begin{cases} 1, & \text{if} \quad Lever\_TDA <= 0.05, \\ 0, & \text{if} \quad Lever\_TDA > 0.05. \end{cases}$
$AZTD\_C$	AZTD_C $= \begin{cases} 1, & \text{if} \quad Lever\_TDA\_Lag > 0.05 \quad \text{and} \quad Lever\_TDA <= 0.05 \\ 0, & \text{if} \quad Lever\_TDA\_Lag > 0.05 \quad \text{and} \quad Lever\_TDA > 0.05. \end{cases}$
HTD	HTD $= \begin{cases} 1, & \text{if} \quad Lever\_TDA >= 0.33595 \\ 0, & \text{if} \quad Lever\_TDA < 0.33595 \end{cases}$
HTD_C	HTD_C $= \begin{cases} 1, & \text{if} \quad Lever\_TDA\_Lag < 0.33595 \quad \text{and} \quad Lever\_TDA >= 0.33595 \\ 0, & \text{if} \quad Lever\_TDA\_Lag < 0.33595 \quad \text{and} \quad Lever\_TDA < 0.33595 \end{cases}$
	Panel B: Dependent variables
$ProductionFlexibility$	$ProductionFlexibility = ROS * Log(\frac{SD_{sales}}{\sqrt{(SD^2_{inventory} + SD^2_{CGS})}})$, where ROS is return on sales, the ratio of Compustat item "EBIT" to Compustat item "SALE", $SD^2_{inventory}$ is the 3-year variance of inventory , i.e., Compustat item"ACQINVT" and SD^2_{CGS}) is the 3-year variance of cost of goods, the Compustat item "COGS"
$FinancialConstraintDummy$	FinancialConstraintDummy$= \begin{cases} 0, & \text{if} \quad Debt_rating < 13 \\ 1, & \text{if} \quad Debt_rating >= 13. \end{cases}$
$Z-Dummy$	Z-Dummy$= \begin{cases} 0, & \text{if} \quad Z-score < 1.85, \\ 1, & \text{if} \quad Z-score > 2.99. \end{cases}$
$KZ-Dummy$	KZ-Dummy $= \begin{cases} 1, & \text{if } KZ-Index \geq Median\ of\ the\ sample \\ 0, & Otherwise. \end{cases}$
$SA-Dummy$	SA-Dummy$= \begin{cases} 1, & \text{if } SA \geq Median\ of\ the\ sample, \\ 0, & Otherwise. \end{cases}$
$Capex$	Is the ratio of Compustat item capital expenditure "CAPX" to total assets "TA".
$ChgAsset$	Is change in log of total assets "TA".
$CPIExpec$	Is the expected change in the consumer price index over the coming year.
$Debt\_Rating$	Is a dummy variable that takes a value of one if Compustat item 280, senior debt rating, or item 320, subordinated debt rating "SUBDBT", have a value of less than 13 (i.e., S&P rates the debt investment grade). Rating takes a value of zero if the debt is not rated or if it is rated less than investment grade. Compustat does not report data on bond ratings before 1985. Thus, the variable is set equal to zero for all firms prior to 1985.
$FirmSize$	Is the log of Compustat item total assets "TA".
$InduChgAsset$	Is the median of change in the log of Compustat item "TA" by SIC code and by year.
$InduLeverTDA$	Is the median of total debt to market value of assets "MKVALT" by SIC code and by year. Industry is defined at the four-digit SIC code level.

InvTaxCr	Is the ratio of Compustat item "ITCB" (investment tax credi (balance sheet)) to total assets "TA".
Leverage_Lag	Is leverage in previous year where leverage is the ratio of total debt (the total of debt in current liabilities "LCT" and long-term debt "DLTT") to total assets "TA".
MacroGr	Is differences in log of real Gross Domestic Product in 1996 dollars.(Source: U.S. Department of Commerce, Bureau of Economic Analysis.).
MacroProf	Is differences in log of aggregate annual corporate profits after tax for non-financial firms.(Source: U.S. Department of Commerce, Bureau of Economic Analysis.)
MktBk	Is the ratio of market value of assets (MVA) to total assets "TA". MVA is obtained as the sum of the market value of equity (price-close "PRCC_C" * shares outstanding "CSHO") + debt in current liabilities "LCT" + long-term debt "DLTT" + preferred stock liquidating value "PSTKL", - deferred taxes and investment tax credit "TXDITC".
Profitability	Is the ratio of operating income before depreciation "OIBDP", to total assets "TA".
ReturnVar	Is the annual variance of asset returns that is obtained by unleveraging the variance of equity returns. Return variance is coded as missing if CRSP has less than 100 valid daily return observations in a fiscal year.
RnD	Is the ratio of research & development expense "XGD" to sales ("SALE").
SGA	Is the ratio of selling, general and administration expenses "XSGA" to sales ("SALE").
StockMarketReturn	Is annual CRSP Value-Weighted Index return.
StockMeanRet	Is average raw stock return obtained from monthly returns from CRSP.
Tangibility	Is the ratio of Compustat item net property, plant and equipment "PPEGT" to total assets "TA".
TaxRate	Is the top statutory tax rate. It was 42 percent in 1950, 51 percent in 1951, 52 percent from 1952-1963, 50 percent in 1964, 48 percent from 1965 to 1967, 52.8 percent from 1968 to 1969, 49.2 percent in 1970, 48 percent from 1971 to 1978, 46 percent from 1979 to 1986, 40 percent in 1987, 34 percent from 1988 to 1992, and 35 percent from 1993 to 2003.
TermSprd	Is the difference between the ten-year interest series and the one-yearinterest series. (Source: Federal Reserve files at http:www.federalreserve.govreleases.).

Table B.5: Leverage on production flexibility for NYSE Firms

This table reports the result of panel data analysis on the sample listed in NYSE over 1962 to 2017, in terms of decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include $ProductionFlexibility$ and other control variables. Panel regressions with fixed firm effects are estimated each decade. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.0814***	0.0276***	0.0517***	0.0450***	0.00758***	0.0678***
	(26.07)	(15.19)	(22.41)	(18.92)	(3.49)	(17.55)
MktBk	-0.000349	0.00776***	-0.00138	-0.00301**	-0.00587***	0.00582**
	(-0.20)	(7.24)	(-0.86)	(-3.02)	(-6.09)	(2.91)
Tangibility	0.132***	0.204***	0.106***	0.165***	0.0924***	0.131***
	(6.29)	(16.94)	(7.11)	(9.38)	(6.46)	(5.65)
Profitability	-0.219***	-0.341***	-0.181***	-0.147***	-0.0747***	-0.0521**
	(-8.60)	(-30.01)	(-11.72)	(-7.51)	(-7.19)	(-2.86)
ProductionFlexibility	0.00212	0.00248	-0.00330	-0.00772**	-0.00525**	-0.00562*
	(0.72)	(1.38)	(-1.36)	(-3.01)	(-2.87)	(-2.20)
Intercept	-0.137***	0.101***	-0.0432**	-0.0495**	0.183***	-0.307***
	(-8.55)	(10.33)	(-2.86)	(-2.84)	(10.57)	(-9.42)
N	5105	15716	9704	7951	11130	5183
Adjusted R^2	0.188	0.096	0.090	0.076	0.018	0.080
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.6: Leverage on production flexibility for Nasdaq Firms

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017, in terms of decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.0929***	0.0380***	0.0443***	0.0268***	0.0258***	0.0667***
	(4.75)	(7.42)	(12.10)	(11.66)	(13.65)	(18.11)
MktBk	0.00514	0.00483	0.000115	-0.00299***	-0.00170**	-0.0000580
	(0.66)	(1.57)	(0.08)	(-4.80)	(-3.12)	(-0.05)
Tangibility	0.0771	0.239***	0.147***	0.252***	0.198***	0.143***
	(0.57)	(7.39)	(6.64)	(13.49)	(13.70)	(4.96)
Profitability	-0.502**	-0.236***	-0.130***	-0.0577***	-0.0444***	-0.0586***
	(-2.72)	(-10.68)	(-8.36)	(-5.70)	(-7.39)	(-5.45)
ProductionFlexibility	0.0144	-0.00313	-0.000234	-0.000625	-0.00127	-0.00287
	(0.69)	(-0.63)	(-0.06)	(-0.26)	(-0.98)	(-1.38)
Intercept	-0.121	0.0655**	0.0248	0.00669	-0.0309**	-0.250***
	(-1.16)	(3.13)	(1.53)	(0.58)	(-2.79)	(-11.00)
N	156	2293	4969	8160	14904	5858
Adjusted R^2	0.244	0.118	0.072	0.060	0.030	0.078
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.7: Leverage on production flexibility and financial constraints for NYSE firms

This table reports the result of panel data analysis on the sample listed in NYSE over 1962 to 2017, in terms of decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. Regressions with firms facing debt market constraints are reported in Panel A, and regressions with firms not facing debt market constraints are reported in Panel B. In Panel B, the regressions for the first two decades are missing due to insufficient observations. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Panel A: Market Access Constrained						
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.0814***	0.0276***	0.0468***	0.0513***	0.0233***	0.0838***
	(26.07)	(15.19)	(18.14)	(15.23)	(6.92)	(12.15)
MktBk	-0.000349	0.00776***	-0.00291	-0.00184	-0.00504***	0.000571
	(-0.20)	(7.24)	(-1.80)	(-1.64)	(-4.15)	(0.24)
Tangibility	0.132***	0.204***	0.129***	0.213***	0.0855***	0.164***
	(6.29)	(16.94)	(7.79)	(9.01)	(3.99)	(4.32)
Profitability	-0.219***	-0.341***	-0.161***	-0.136***	-0.0581***	-0.0336
	(-8.60)	(-30.01)	(-10.21)	(-5.96)	(-4.69)	(-1.40)
ProductionFlexibility	0.00212	0.00248	-0.00221	-0.0108**	-0.00435	-0.0128**
	(0.72)	(1.38)	(-0.85)	(-3.09)	(-1.62)	(-3.23)
Intercept	-0.137***	0.101***	-0.0159	-0.0808***	0.0334	-0.374***
	(-8.55)	(10.33)	(-1.00)	(-3.95)	(1.56)	(-7.61)
N	5105	15716	8525	5023	5060	2247
Adjusted R^2	0.188	0.096	0.081	0.089	0.027	0.095
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
Panel B: Market Access Not Constrained						
FirmSize			0.0438***	0.0232***	-0.0276***	0.0639***
			(4.03)	(5.21)	(-8.79)	(12.09)
MktBk			0.0372**	-0.00348	-0.00487**	0.0332***
			(2.75)	(-1.07)	(-2.86)	(7.62)
Tangibility			-0.0234	0.0172	0.0829***	0.122***
			(-0.47)	(0.61)	(4.29)	(3.98)

Continuation of Table B.7:

	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Profitability			-0.106	-0.163**	-0.175***	-0.0989**
			(-1.26)	(-3.18)	(-8.46)	(-3.19)
ProductionFlexibility			-0.00895	0.00468	-0.00417	-0.000710
			(-1.40)	(1.27)	(-1.73)	(-0.22)
Intercept			-0.0498	0.150***	0.539***	-0.301***
			(-0.51)	(3.58)	(18.37)	(-6.16)
N			1179	2928	6070	2936
Adjusted R^2			0.037	0.024	0.036	0.083
Firm Fixed Effect			yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.8: Leverage on production flexibility and financial constraints with for Nasdaq Firms

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017, in terms of decades. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility* and other control variables. Panel regressions with fixed firm effects are estimated for each decade. Regressions with firms facing debt market constraints are reported in Panel A, while regressions with firms not facing debt market constraints are reported in Panel B. In Panel B, the regressions for the first two decades are missing due to insufficient observations. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	Panel A: Market Access Constrained					
	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.0929***	0.0380***	0.0444***	0.0259***	0.0226***	0.0585***
	(4.75)	(7.42)	(11.80)	(10.55)	(11.33)	(14.18)
MktBk	0.00514	0.00483	0.0000531	-0.00263***	-0.00230***	-0.000939
	(0.66)	(1.57)	(0.04)	(-4.16)	(-4.19)	(-0.75)
Tangibility	0.0771	0.239***	0.150***	0.268***	0.206***	0.170***
	(0.57)	(7.39)	(6.68)	(13.76)	(13.62)	(5.32)
Profitability	-0.502**	-0.236***	-0.130***	-0.0526***	-0.0382***	-0.0502***
	(-2.72)	(-10.68)	(-8.28)	(-5.13)	(-6.46)	(-4.60)
ProductionFlexibility	0.0144	-0.00313	-0.0000305	-0.000436	-0.00134	-0.00246
	(0.69)	(-0.63)	(-0.01)	(-0.18)	(-1.01)	(-1.12)
Intercept	-0.121	0.0655**	0.0250	0.00232	-0.0317**	-0.213***
	(-1.16)	(3.13)	(1.54)	(0.20)	(-2.89)	(-9.12)
N	156	2293	4854	7728	13027	4928
Adjusted R^2	0.244	0.118	0.072	0.060	0.031	0.064
Firm Fixed Effect	yes	yes	yes	yes	yes	yes
	Panel B: Market Access Not Constrained					
FirmSize			0.115*	0.0360**	0.00656	0.0581***
			(2.23)	(2.85)	(0.87)	(5.38)
MktBk			0.0501	-0.00965*	-0.000282	0.0229***
			(0.88)	(-2.23)	(-0.12)	(3.62)
Tangibility			0.0160	0.106	0.156***	0.127
			(0.08)	(1.26)	(3.34)	(1.81)

Continuation of Table B.8:

	(1962-1970)	(1971-1980)	(1981-1990)	(1991-2000)	(2001-2010)	(2011-2017)
Profitability			0.0477	-0.194	-0.144**	-0.0701
			(0.17)	(-1.95)	(-3.12)	(-0.83)
ProductionFlexibility			0.00781	-0.00942	0.00312	0.00317
			(0.37)	(-0.91)	(0.60)	(0.52)
Intercept			-0.561	0.0784	0.245***	-0.191*
			(-1.32)	(0.80)	(3.93)	(-2.03)
N			115	432	1877	930
Adjusted R^2			0.067	0.061	0.011	0.056
Firm Fixed Effect			yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.9: Correlation statistics

The table below presents the correlation matrix of independent variables included in our research. it is used for variable reduction later on. Variables with correlations exceeding 0.4 are deemed as highly correlated, and at least one of them need to be removed from the sample. Which of them is removed depend on which one is less correlated with the target variable.

	Lever_TDA	FirmSize	MktBk	Tangibility	Profitability	ProductionFlexibility
Lever_TDA	1					
FirmSize	0.372	1				
MktBk	-0.2515	-0.2051	1			
Tangibility	0.3238	0.2116	-0.1805	1		
Profitability	-0.0554	0.1024	0.1604	0.1207	1	
ProductionFlexibility	-0.1529	-0.0789	0.2267	-0.1112	0.1801	1
AgencyDummy	0.071	0.0336	0.0428	0.157	0.0377	-0.0506
StockMarketRet	0.0222	-0.0315	0.064	0.041	0.028	-0.0228
Grow_ChgAsset	-0.018	-0.0249	0.236	-0.0617	0.1657	0.1076
Indu_ChgAsset	-0.0284	-0.0326	0.1631	0.0103	0.1313	0.0658
CPIExpec	0.0126	-0.0104	-0.0379	-0.0099	-0.026	0.0027
MacroGr	0.0445	-0.0155	0.1039	0.0529	0.0647	-0.0175
MacroProf	-0.0348	0.0188	-0.0159	-0.0039	0.0053	-0.0189
RnD	0.0167	0.0243	-0.0536	0.0336	0.1471	-0.009
SGA	-0.0586	-0.0711	0.0461	-0.0827	-0.0831	0.0569
ReturnVar	-0.1073	-0.3267	0.1235	-0.1441	-0.2317	0.0641
InvTaxCr	0.1364	0.1297	-0.0844	0.2689	-0.022	-0.0761
TermSprd	-0.0342	0.008	-0.0798	-0.0308	-0.0639	-0.0049
TaxRate	-0.0134	0.048	0.0052	-0.0743	-0.0105	0.0289
StockMeanRet	-0.0918	-0.1116	0.3036	-0.0492	0.1136	0.011
Lever_TDA_Lag	0.8863	0.379	-0.274	0.3266	-0.0701	-0.168
NatureAsset_Unique	-0.1711	-0.1276	0.0461	-0.2769	-0.0328	0.0304
IndusRegulated	0.0971	0.0877	-0.0469	0.2018	0.0343	-0.0389

	AgencyDummy	StockMarketRet	Grow_ChgAsset	Indu_ChgAsset	CPIExpec	MacroGr
AgencyDummy	1					
StockMarketRet	0.0379	1				
Grow_ChgAsset	0.0722	0.0559	1			
Indu_ChgAsset	0.1152	0.1625	0.3643	1		
CPIExpec	0.0257	-0.5258	-0.0442	-0.1424	1	
MacroGr	0.4551	0.1782	0.1619	0.3451	-0.1734	1
MacroProf	0.1333	0.2509	0.0158	0.0877	-0.4569	0.427
RnD	-0.0143	0.0113	-0.0288	0.0034	-0.0138	-0.0099
SGA	-0.0229	-0.0029	-0.0135	-0.0177	0.0018	-0.0264
ReturnVar	0.0365	-0.1253	0.0261	-0.0479	0.1151	-0.1362

Continuation of Table B.9:

InvTaxCr	0.1063	0.0602	-0.0318	-0.0365	-0.0298	0.0236
TermSprd	-0.1927	0.0121	-0.1399	-0.3216	0.2772	-0.6632
TaxRate	-0.1218	-0.2039	0.0061	0.045	0.1418	0.2529
StockMeanRet	0.0614	0.258	0.239	0.1518	-0.1735	0.064
Lever_TDA_Lag	0.0855	-0.004	-0.0341	-0.0443	0.0269	0.0161
NatureAsset_Unique	-0.0112	-0.0109	-0.0101	-0.0404	0.0033	-0.0057
IndusRegulated	0.0274	0.0017	0.0042	0.0205	0.0009	0.0068
	MacroProf	RnD	SGA	ReturnVar	InvTaxCr	TermSprd
MacroProf	1					
RnD	-0.006	1				
SGA	-0.01	-0.0038	1			
ReturnVar	-0.2036	-0.0308	0.0559	1		
InvTaxCr	0.0046	0.0069	-0.0316	-0.1351	1	
TermSprd	-0.061	0.0061	0.0148	0.1212	-0.0184	1
TaxRate	0.0685	-0.0081	0.007	0.0237	-0.1204	-0.1696
StockMeanRet	0.0843	0.0085	0.003	0.2581	-0.0274	0.0373
Lever_TDA_Lag	-0.0367	0.0149	-0.066	-0.08	0.1408	-0.0025
NatureAsset_Unique	0.0031	-0.0059	0.0173	0.0902	-0.0935	0.0047
IndusRegulated	-0.0043	0.0063	-0.0221	-0.0193	0.0172	-0.0078
	TaxRate	StockMeanRet	Lever_TDA_Lag	NatureAsset_Unique	IndusRegulated	
StockMeanRet	-0.0475	1				
Lever_TDA_Lag	-0.036	-0.086	1			
NatureAsset_Unique	0.0061	0.0213	-0.1663	1		
IndusRegulated	-0.0108	-0.0113	0.0947	-0.1115	1	

End of Table

Table B.10: Robustness tests: leverage on production flexibility, financial constraints and agency problem dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017, in terms of decades, with more control variables being introduced for the robustness test. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility*, Agency Problem Dummy *AgencyDummy*, the interaction between the two *ProFlex_Agen*, and other control variables. Panel regressions with fixed firm effects are estimated on both debt-access-constrained firms and firms with non-constrained debt access for all firms (Columns 1 and 2), NYSE firms (Columns 3 and 4), and Nasdaq firms (Columns 5 and 6). All the 12 regressions are estimated with firm effects remaining fixed. . Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA	Lever_TDA
FirmSize	0.0128***	0.000696	0.0121***	-0.00104	0.0122***	0.00964
	(9.17)	(0.40)	(4.83)	(-0.56)	(6.84)	(1.84)
MktBk	0.000634	0.00190	-0.000774	0.00166	0.000649	0.00391
	(1.44)	(1.81)	(-0.59)	(1.36)	(1.26)	(1.64)
Tangibility	0.0518***	0.00904	0.0156	0.00242	0.0687***	0.00875
	(4.26)	(0.79)	(0.86)	(0.20)	(3.92)	(0.25)
Profitability	-0.0184*	0.0354*	0.00320	0.0269	-0.00879	0.0883*
	(-2.30)	(2.23)	(0.17)	(1.56)	(-0.85)	(1.97)
ProductionFlexibility	0.00158	0.00233	-0.00392	0.000640	0.00193	0.00762
	(1.05)	(1.33)	(-1.39)	(0.33)	(1.04)	(1.66)
AgencyDummy	-0.00174	-0.00395**	-0.00209	-0.00489***	-0.0000692	-0.00159
	(-1.34)	(-3.21)	(-1.04)	(-3.75)	(-0.04)	(-0.43)
ProFlex_Agen	-0.00275*	-0.000649	-0.000280	0.000897	-0.00418**	-0.00340
	(-2.14)	(-0.42)	(-0.11)	(0.51)	(-2.67)	(-0.86)
StockMarketRet	0.0223***	0.0155***	0.0297***	0.0158***	0.0130*	0.0190
	(5.36)	(3.63)	(4.61)	(3.53)	(2.30)	(1.47)
Grow_Chgasset	0.00757*	0.0139***	0.0210***	0.00570	0.00189	0.0296***
	(2.39)	(4.10)	(3.45)	(1.47)	(0.48)	(3.76)
Indu_Chgasset	-0.00728	-0.0142	-0.00811	-0.00751	-0.00192	-0.0155

Continuation of Table B.10:

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(-0.86)	(-1.70)	(-0.62)	(-0.82)	(-0.16)	(-0.65)
MacroGr	3.335***	2.086***	3.115***	2.241***	3.633***	1.328
	(6.57)	(4.05)	(3.79)	(4.19)	(5.33)	(0.83)
RnD	0.0128	0.0162	-0.0265	0.0235	0.0125	0.0122
	(1.32)	(0.52)	(-0.14)	(0.43)	(1.23)	(0.26)
SGA	0.00118*	0.0397**	0.00139**	0.0313	0.000260	0.0595**
	(2.43)	(3.12)	(2.86)	(1.63)	(0.17)	(2.69)
RetrunVar	-0.00152	0.0217	-0.0211	0.0205	0.00665	0.0441
	(-0.13)	(1.46)	(-0.95)	(1.23)	(0.45)	(1.19)
InvTaxcr	-0.561	0.136	-0.777	0.157	0.180	-0.344
	(-0.47)	(0.40)	(-0.68)	(0.48)	(0.02)	(-0.16)
TaxRate	0.501	0.564	-0.140	0.589	0.497	-0.287
	(1.35)	(1.59)	(-0.23)	(1.71)	(0.93)	(-0.14)
StockMeanRet	-0.0659***	-0.193***	-0.103**	-0.197***	-0.0432	-0.246***
	(-3.55)	(-9.00)	(-3.12)	(-7.98)	(-1.81)	(-4.69)
Lever_Tda_Lag	0.600***	0.618***	0.626***	0.651***	0.575***	0.558***
	(75.25)	(72.36)	(50.09)	(69.05)	(51.84)	(26.14)
Natureasset_Unique	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)
IndusRegulated	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)
Intercept	-0.214	-0.112	0.0308	-0.109	-0.222	0.101
	(-1.67)	(-0.92)	(0.15)	(-0.92)	(-1.20)	(0.14)
N	13585	10414	5285	8182	7207	1800
Adjusted R^2	0.365	0.392	0.407	0.428	0.330	0.333
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.11: Robustness Tests: Low leverage on production flexibility, financial constraints and agency problem dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017 in terms of decades, with more control variables being introduced for robustness test. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include $ProductionFlexibility$, Agency Problem Dummy $AgencyDummy$, the interaction between the two $ProFlex\_Agen$, and other control variables. Panel regressions with fixed firm effects are estimated on both debt-access-constrained firms and firms with non-constrained debt access for all firms (Columns 1 and 2), NYSE firms (Columns 3 and 4), and Nasdaq firms (Columns 5 and 6). All the 12 regressions are estimated with firm effects remaining fixed. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
FirmSize	-0.369***	-0.289	-0.213	-0.560*	-0.450***	-0.494
	(-5.89)	(-1.49)	(-1.67)	(-2.04)	(-5.79)	(-1.54)
MktBk	0.0168	0.0668	0.124	0.0866	0.00384	-0.0663
	(0.73)	(0.97)	(1.92)	(0.88)	(0.16)	(-0.58)
Tangibility	-2.572***	-5.484**	-1.089	-2.939	-3.273***	-11.02***
	(-4.57)	(-3.26)	(-1.11)	(-1.51)	(-4.35)	(-3.32)
Profitability	0.390	-0.0560	-1.072	-0.319	0.0113	1.081
	(1.14)	(-0.04)	(-1.11)	(-0.15)	(0.03)	(0.46)
ProductionFlexibility	0.0358	-0.279	0.341*	-0.00655	0.00102	-0.392
	(0.55)	(-1.32)	(2.39)	(-0.02)	(0.01)	(-1.21)
AgencyDummy	0.130*	0.159	0.245*	-0.00861	0.0353	0.391
	(2.26)	(1.12)	(2.31)	(-0.05)	(0.48)	(1.54)
ProFlex_Agen	0.0833	0.0706	-0.0434	0.136	0.150*	-0.00188
	(1.49)	(0.49)	(-0.33)	(0.64)	(2.30)	(-0.01)
StockMarketRet	-0.601**	0.363	-1.305***	-0.153	-0.160	1.237
	(-3.25)	(0.75)	(-3.94)	(-0.24)	(-0.67)	(1.56)
Grow_Chgasset	-0.142	-0.579	-0.514	-0.101	-0.0471	-0.650
	(-1.02)	(-1.47)	(-1.50)	(-0.16)	(-0.29)	(-1.06)
Indu_Chgasset	0.131	-1.993*	0.136	-1.383	-0.0350	-3.936*
	(0.36)	(-2.12)	(0.21)	(-1.04)	(-0.07)	(-2.31)

Continuation of Table B.11:

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
MacroGr	-96.48***	-0.419	-56.42	-50.23	-113.3***	79.96
	(-4.09)	(-0.01)	(-1.35)	(-0.63)	(-3.71)	(0.74)
RnD	-1.169	26.25	10.13	-8.740	-1.347	34.46
	(-1.48)	(1.90)	(1.22)	(-0.44)	(-1.53)	(1.93)
SGA	-0.0269	2.951	-0.395	5.652	-0.00214	2.955
	(-1.06)	(1.61)	(-1.01)	(1.58)	(-0.03)	(1.26)
RetrunVar	-0.658	-0.0303	-0.523	0.396	-0.673	-0.958
	(-1.30)	(-0.02)	(-0.45)	(0.14)	(-1.14)	(-0.40)
InvTaxcr	72.66	0	-12.21	0	2996.3	0
	(0.26)	(.)	(-0.04)	(.)	(0.01)	(.)
TaxRate	-6.132	-46.07	-27.08	-45.66	7.391	218.4
	(-0.28)	(-0.86)	(-0.77)	(-0.76)	(0.23)	(1.36)
StockMeanRet	1.746*	2.734	1.464	5.816	1.970*	1.370
	(2.19)	(1.20)	(0.87)	(1.72)	(2.02)	(0.37)
lever_tda_lag	-14.77***	-14.68***	-18.48***	-15.60***	-13.25***	-14.69***
	(-30.34)	(-12.64)	(-19.11)	(-9.66)	(-21.98)	(-7.86)
N	7809	1538	2784	939	4455	558
ll	-2276.4	-364.7	-715.0	-214.9	-1363.9	-128.6
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.12: Robustness tests: high leverage on production flexibility, financial constraints and agency problem dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017 in terms of decades, with more control variables being introduced for robustness test. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility*, Agency Problem Dummy *AgencyDummy*, the interaction between the two *ProFlex_Agen*, and other control variables. Panel regressions with fixed firm effects are estimated on both debt-access-constrained firms and firms with non-constrained debt access for all firms (Columns 1 and 2), NYSE firms (Columns 3 and 4), and Nasdaq firms (Columns 5 and 6). All the 12 regressions are estimated with firm effects remaining fixed. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	HTD	HTD	HTD	HTD	HTD	HTD
FirmSize	0.151	-0.0600	0.0427	-0.156	0.196	0.199
	(1.76)	(-0.67)	(0.28)	(-1.56)	(1.60)	(0.78)
MktBk	0.0646*	0.134	0.00632	0.147	0.0793	0.231
	(1.97)	(1.93)	(0.06)	(1.82)	(1.65)	(1.89)
Tangibility	1.618*	-0.975	1.574	-1.012	2.439*	-2.182
	(2.47)	(-1.77)	(1.45)	(-1.61)	(2.46)	(-1.45)
Profitability	-0.366	1.340	2.133	0.554	-1.201	3.991*
	(-0.68)	(1.69)	(1.43)	(0.61)	(-1.55)	(2.12)
ProductionFlexibility	0.157	-0.0222	0.0781	-0.0329	0.203	-0.00421
	(1.63)	(-0.25)	(0.44)	(-0.31)	(1.65)	(-0.02)
AgencyDummy	-0.0705	-0.192**	-0.156	-0.230**	0.0548	-0.162
	(-0.89)	(-3.09)	(-1.27)	(-3.22)	(0.47)	(-1.03)
ProFlex_Agen	-0.163	0.173*	0.0519	0.221*	-0.319**	0.160
	(-1.85)	(2.12)	(0.34)	(2.29)	(-2.76)	(0.87)
StockMarketRet	0.949***	0.303	1.011**	0.398	0.646	-0.0484
	(3.78)	(1.45)	(2.63)	(1.69)	(1.67)	(-0.09)
Grow_Chgasset	0.499**	0.438*	0.339	0.271	0.426	0.435
	(2.97)	(2.55)	(1.06)	(1.32)	(1.89)	(1.50)
Indu_Chgasset	-0.124	-0.469	0.344	0.309	-0.224	-1.302

Continuation of Table B.12:

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(-0.25)	(-1.05)	(0.44)	(0.57)	(-0.29)	(-1.35)
MacroGr	135.8***	84.03**	116.9*	84.94**	155.5**	79.32
	(4.29)	(3.22)	(2.30)	(2.90)	(3.08)	(1.19)
RnD	2.910*	-0.881	-10.36	0.612	2.481	-4.130
	(2.24)	(-0.49)	(-0.77)	(0.18)	(1.91)	(-0.65)
SGA	-0.0565	1.762*	-0.287	0.884	-0.0207	2.531*
	(-0.35)	(2.37)	(-0.16)	(0.89)	(-0.12)	(2.48)
RetrunVar	0.496	1.249	-0.972	1.996*	0.651	-0.639
	(0.68)	(1.63)	(-0.76)	(2.15)	(0.60)	(-0.42)
InvTaxcr	41.54	0.486	15.22	2.649	0	-1858.6
	(0.97)	(0.04)	(0.35)	(0.19)	(.)	(-0.02)
TaxRate	12.33	-28.85	-5.197	-28.62	5.938	37.55
	(0.60)	(-1.59)	(-0.15)	(-1.48)	(0.12)	(0.47)
StockMeanRet	-3.444**	-5.798***	-4.865*	-6.052***	-1.342	-4.566*
	(-3.01)	(-5.35)	(-2.42)	(-4.56)	(-0.83)	(-2.10)
lever_tda_lag	8.098***	13.68***	9.974***	15.52***	6.769***	9.874***
	(20.45)	(26.71)	(13.83)	(24.94)	(13.17)	(9.86)
N	4398	6368	1955	5194	1958	964
ll	-1258.4	-1924.7	-567.1	-1533.8	-544.5	-282.6
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.13: Robustness tests: low leverage Change on production flexibility, financial constraints and agency problem dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017 in terms of decades, with more control variables being introduced for robustness test. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility*, Agency Problem Dummy *AgencyDummy*, the interaction between the two *ProFlex_Agen*, and other control variables. Panel regressions with fixed firm effects are estimated on both debt-access-constrained firms and firms with non-constrained debt access for all firms (Columns 1 and 2), NYSE firms (Columns 3 and 4), and Nasdaq firms (Columns 5 and 6). All the 12 regressions are estimated with firm effects remaining fixed. Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	AZTD_C	AZTD_C	AZTD_C	AZTD_C	AZTD_C	AZTD_C
FirmSize	-0.625***	-0.571*	-0.572*	-0.920*	-0.651***	-0.672
	(-5.27)	(-2.10)	(-2.20)	(-2.52)	(-4.37)	(-1.10)
MktBk	0.108	0.0682	0.139	0.127	0.0822	-0.547*
	(1.66)	(0.70)	(1.07)	(0.84)	(0.98)	(-1.98)
Tangibility	-1.623	-1.586	3.012	1.721	-4.970***	-10.14*
	(-1.66)	(-0.79)	(1.70)	(0.70)	(-3.54)	(-2.02)
Profitability	0.522	0.873	-0.964	0.495	-0.357	-1.747
	(0.66)	(0.50)	(-0.51)	(0.20)	(-0.35)	(-0.40)
ProductionFlexibility	0.206	-0.370	0.714*	-0.179	0.0777	-0.272
	(1.56)	(-1.32)	(2.42)	(-0.45)	(0.47)	(-0.66)
AgencyDummy	-0.128	0.229	0.0427	0.104	-0.267	0.310
	(-1.22)	(1.21)	(0.22)	(0.44)	(-1.92)	(0.84)
ProFlex_Agen	0.129	-0.264	-0.266	-0.0260	0.292*	-0.724
	(1.15)	(-1.17)	(-1.06)	(-0.09)	(2.13)	(-1.75)
StockMarketRet	-0.356	0.340	-1.379*	-0.291	0.489	1.825
	(-1.21)	(0.54)	(-2.55)	(-0.37)	(1.25)	(1.50)
Grow_Chgasset	-1.138***	-0.414	-0.995	-0.554	-1.133**	0.289
	(-3.79)	(-0.78)	(-1.40)	(-0.67)	(-3.13)	(0.32)
Indu_Chgasset	-0.476	-2.006	-0.684	-0.536	-0.744	-6.294*

Continuation of Table B.13:

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(-0.71)	(-1.72)	(-0.59)	(-0.33)	(-0.79)	(-2.13)
MacroGr	-117.0**	-6.515	-63.60	-95.44	-97.62	223.9
	(-2.98)	(-0.08)	(-0.82)	(-0.92)	(-1.95)	(1.27)
RnD	-2.677	33.65	-1.115	-26.95	-4.861	62.48
	(-1.27)	(1.44)	(-0.05)	(-1.04)	(-1.87)	(1.64)
SGA	-0.547	7.786**	-0.222	10.64*	-0.683	7.399
	(-0.94)	(3.20)	(-0.63)	(2.45)	(-1.03)	(1.72)
RetrunVar	-1.098	-1.069	-1.744	-5.180	-0.740	2.671
	(-1.37)	(-0.46)	(-0.90)	(-1.37)	(-0.86)	(0.69)
InvTaxcr	-448.6	0	-497.5	0	0	0
	(-0.60)	(.)	(-0.54)	(.)	(.)	(.)
TaxRate	47.13	-9.316	59.88	3.751	55.57	256.3
	(1.33)	(-0.15)	(0.95)	(0.05)	(1.05)	(1.28)
StockMeanRet	4.224**	2.191	4.530	8.162	3.803*	2.664
	(3.04)	(0.70)	(1.68)	(1.65)	(2.14)	(0.49)
lever_tda_lag	-14.93***	-20.49***	-23.09***	-24.66***	-13.60***	-22.95***
	(-14.22)	(-9.07)	(-9.89)	(-7.34)	(-10.18)	(-5.27)
N	2925	1076	1195	692	1551	360
ll	-664.1	-191.7	-210.7	-113.2	-374.3	-55.89
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Table B.14: Robustness tests: high leverage change on production flexibility, financial constraints and agency problem dummy

This table reports the result of panel data analysis on the sample listed in Nasdaq over 1962 to 2017 in terms of decades, with more control variables being introduced for robustness test. The sample includes all firms in the Compustat and CRSP merged database from 1962 to 2017, with financial firms being excluded. The dependent variable is leverage-measured using total debt divided by total asset. All dependent variables are lagged by one year. Independent variables include *ProductionFlexibility*, Agency Problem Dummy *AgencyDummy*, the interaction between the two *ProFlex_Agen*, and other control variables. Panel regressions with fixed firm effects are estimated on both debt-access-constrained firms and firms with non-constrained debt access for all firms (Columns 1 and 2), NYSE firms (Columns 3 and 4), and Nasdaq firms (Columns 5 and 6). All the 12 regressions are estimated with firm effects remaining fixed. . Coefficients marked with ***,**, and * are significant at 1%,5%, and 10%, respectively.

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	HTD_C	HTD_C	HTD_C	HTD_C	HTD_C	HTD_C
FirmSize	0.110	-0.154	0.00997	-0.158	0.0677	-0.537
	(1.01)	(-1.15)	(0.05)	(-1.05)	(0.44)	(-1.28)
MktBk	0.0885**	0.259***	-0.0660	0.202*	0.119*	0.413**
	(2.67)	(3.44)	(-0.52)	(2.15)	(2.46)	(2.69)
Tangibility	0.865	-1.166	-0.0190	-1.403	1.626	-0.383
	(0.99)	(-1.28)	(-0.01)	(-1.34)	(1.21)	(-0.15)
Profitability	0.418	1.905	3.961*	2.104	-0.130	1.534
	(0.67)	(1.52)	(1.98)	(1.45)	(-0.15)	(0.47)
ProductionFlexibility	0.196	0.0954	-0.0244	0.0178	0.392*	0.280
	(1.63)	(0.68)	(-0.11)	(0.11)	(2.40)	(0.73)
AgencyDummy	-0.110	-0.288**	-0.233	-0.331**	0.0563	-0.253
	(-1.07)	(-3.00)	(-1.56)	(-3.06)	(0.35)	(-0.98)
ProFlex_Agen	-0.290**	0.0844	-0.0712	0.125	-0.546***	0.297
	(-2.68)	(0.65)	(-0.40)	(0.86)	(-3.63)	(0.88)
StockMarketRet	1.156***	0.219	1.305*	0.204	0.820	-0.532
	(3.37)	(0.64)	(2.51)	(0.52)	(1.56)	(-0.62)
Grow_Chgasset	0.414*	-0.194	0.610	-0.0988	0.235	-0.747
	(2.00)	(-0.66)	(1.39)	(-0.29)	(0.84)	(-1.19)
Indu_Chgasset	-1.341*	-0.174	-0.768	-0.122	-1.779	0.907

Continuation of Table B.14:

	All Firms		NYSE Firms		Nasdaq Firms	
	Constrained	Not Constrained	Constrained	Not Constrained	Constrained	Not Constrained
	(-1.96)	(-0.26)	(-0.72)	(-0.16)	(-1.62)	(0.54)
MacroGr	166.0***	166.9***	138.1*	183.8***	176.8**	66.15
	(3.92)	(3.69)	(2.22)	(3.59)	(2.70)	(0.61)
RnD	3.726	-5.003	-6.129	-3.436	3.240	-24.68
	(1.35)	(-1.41)	(-0.39)	(-0.81)	(1.14)	(-1.28)
SGA	0.0105	1.108	-0.503	1.164	0.0729	1.410
	(0.08)	(0.89)	(-0.21)	(0.75)	(0.54)	(0.55)
RetrunVar	0.654	0.270	-1.548	-0.103	1.985	-0.0556
	(0.66)	(0.21)	(-0.87)	(-0.06)	(1.47)	(-0.02)
InvTaxcr	184.8	-12.84	161.8	-11.19	0	-2139.2
	(1.22)	(-0.45)	(1.04)	(-0.39)	(.)	(-0.01)
TaxRate	82.04*	-5.706	64.17	-24.33	1373.8	1500.3
	(2.48)	(-0.16)	(1.76)	(-0.63)	(0.03)	(0.01)
StockMeanRet	-2.365	-8.795***	-5.195	-9.323***	0.315	-6.529
	(-1.62)	(-4.79)	(-1.83)	(-4.18)	(0.16)	(-1.63)
lever_tda_lag	6.326***	12.53***	7.140***	12.55***	5.693***	14.10***
	(8.98)	(11.00)	(5.77)	(9.57)	(5.76)	(5.29)
N	2817	2936	1231	2378	1297	463
ll	-721.7	-787.2	-322.8	-637.7	-313.9	-113.8
Firm Fixed Effect	yes	yes	yes	yes	yes	yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

End of Table

Appendix C

MODEL NOTATIONS

Table C.1: Scenarios debtholders face and related price points

The following table summarizes scenarios debtholders face and related price point.

Panel A: Scenarios debtholders face	
Scenario a	Firms go bankrupt and nothing is left to debtholders after bankruptcy costs get paid.
Scenario b	Firms go bankrupt, and money left to debtholders after bankruptcy costs get paid is between 0 and $D(1+r)$.
Scenario c	Firms do not go bankrupt, and debtholders get $D(1+r)$.
Scenario d	After risky investments, firms go bankrupt and debtholders get 0 after bankruptcy costs get paid.
Scenario e	After risky investments, firms go bankrupt and debtholders get between 0 to $D(1+r)$ after bankruptcy costs get paid.
Scenario f	After risky investments, firms do not go bankrupt and debt holders get $D(1+r)$.
Panel B: Related price break point to the scenarios debtholders face	
θ_1	Firms go bankrupt, and the asset left is equal to the bankruptcy costs.
θ_2	Firms are on the verge of going bankrupt.
θ_3	The price at which firms reach gains enough profits to initiate the risky investment.
θ_4	Firms go bankrupt due to the risky investment, and the asset equals the bankruptcy costs.
θ_5	After the risky investment, firms are on the verge of bankruptcy.