ESSAYS ON APPLIED FINANCIAL MODELING

by

Hao Zhang

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Business Administration

Charlotte

2023

Approved by:

____

Dr. Steven P. Clark

____

Dr. I-Hsuan Ethan Chiang

____

Dr. Kiplan Womack

____

Dr. Jaya Bishwal

ABSTRACT

HAO ZHANG. Essays on applied financial modeling. (Under the direction of DR. STEVEN P. CLARK)

This dissertation explores the application of quantitative modeling to analyze the features of different financial assets. This study investigates the dynamics of financial assets using advanced mathematical, statistical approaches, and Machine Learning method. We aim at contributing to the field of finance in empirical asset pricing, risk management, and investment.

In the first chapter, we provide empirical analysis on Fixed Index Annuity (FIA) and Fixed Index Linked Annuity (FILA) with insights on utility gains of different types of investors. As one of the most recent financial product in the market, we find that this financial asset provides higher and secured returns for investors and could be an alternative investment especially in the era of low yield market. We also construct a multi-period utility framework to model the utility preferences which provide many intuitive findings in the insurance industry.

We investigate the conditional betas in the U.S. stock market based on individual stocks. In this study, we not only use econometric modeling but also Machine Learning approach to capture the conditional betas in stock market. Different from previous literature, we include a comprehensive list of variables to model time-varying betas and examine the asset pricing models such as Capital Asset Pricing Model (CAPM), Fama French models, and Q5 models from the asset pricing model tests perspective.

As one of the most important investment in the market, Real Estate Investment Trusts (REITs) has been constantly regarded as a diversification investment for many fund managers. Especially in the 20008 financial crisis and 2020 Covid pandemic, REITs has played an important role and achieved defensive role in the portfolio optimization. The findings of this dissertation contribute to the understanding of quantitative modeling in finance and

offer practical implications for asset pricing, risk management and investment. In summary, these essays provide intuitive, economically insightful and interesting findings on financial modeling for wide applications.

DEDICATION

I would like to dedicate this dissertation to my dear parents for their support and encouragement throughout my life. I am deeply grateful for their patience, understanding, and sacrifices that motivated my pursuing academic aspirations. I dedicate this work to them with heartfelt appreciation and love. I also dedicate this dissertation to all of my family, friends and every faculty member in my program. Your support helped me achieve such an amazing academic achievement.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and appreciation to all those who have supported me throughout the completion of this dissertation.

First and foremost, I am deeply grateful to my advisor, Dr.Clark, for his invaluable guidance, expertise, and continuous support. His insightful feedback have been instrumental in shaping the direction and quality of this research. He always directed me towards an efficient way to solve problems. I would also like to extend my heartfelt thanks to the members of my dissertation committee, Dr. Chiang, Dr.Womack, and Dr.Bishwal. Their time, expertise, and valuable input have significantly enriched this work. I am also grateful to Dr.Mauer, whose kind advice and understanding have supported my academic growth and provided a conductive research environment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Narrow Framing application in Fixed Index Annuity and Fixed Index Linked Annuity

ABSTRACT

HAO ZHANG. Essays on applied financial modeling. (Under the direction of DR. STEVEN P. CLARK)

Recently, there has been increasing interest in Fixed Indexed Annuities (FIA). These insurance products, specifically designed for retirement saving, are well-known for their ability to offer higher potential returns compared to portfolios of traditional fixed income securities while maintaining a lower level of risk relative to simply holding equities. This unique combination positions FIAs as attractive alternative investments, especially during periods of low interest rates. However, the industry still lacks comprehensive research on the return features of a particular variant of FIA known as the Fixed Indexed Linked Annuity (FILA). Our findings suggest that FILAs can deliver superior return profiles relative to FIAs. Moreover, FILAs offer increased flexibility to policyholders, a characteristic essential for managing long-term investments like retirement funds. This study applies a numerical method to examine the performance of FIAs and FILAs under various stock market conditions. In addition, this study also introduces the examination of FIA performance within a recursive utility framework. This innovative approach allows us to assess the performance of FIAs in a dynamic, recursive context, providing a more comprehensive and nuanced understanding of the benefits FIAs offer as retirement savings products.

1.1 Introduction

In recent years, there has been a growing interest in Fixed Index Annuities (FIAs) among both academics and practitioners. Increasingly, investors are realizing that FIAs could be an ideal investment, allowing them to take advantage of high equity market returns while securing their principal without downside risk. In a low-interest market, FIAs can be particularly advantageous, offering better performance than bonds. As FIA products do not carry the risk of losing principal and can potentially generate higher returns through option structures, more and more insurance companies are introducing them. However, the feature of generating stable cash flows year by year requires complicated option structures, resulting in higher premiums. Despite the benefits of FIA products, there are other costs associated with the contract, such as rider fees and surrender charges. Rider fees are annual fees charged based on the contract's value, while surrender charges refer to fees charged by insurance companies for withdrawing large sums of money from the original account.

Although the high return potential of Fixed Indexed Annuities (FIA) is widely acknowledged in the market, several pertinent questions remain unanswered. For example, what structuring choices should FIA holders select in order to optimize expected results? What are the determining factors that motivate investors to allocate a portion of their retirement savings to FIAs? At present, empirical evidence to guide these decisions is sparse.

Another important question concerns the utility gains of FIA products. While many studies have investigated the returns perspective of FIA, they often ignore the utility aspect. Specifically, the option allocation feature may be attractive to a group of investors, but not to those who prefer to keep their money secure. Furthermore, most investments cannot consistently perform well, meaning that performance tends to fluctuate over time. Would this also apply to FIA, and how would it impact different types of investors? With all of these questions in mind, we use a numerical method to gain a better understanding of this popular insurance product and answer these questions.

1.1.1 Structuring: From FIA to FILA

Every FIA contract is an FILA, but not every FILA contract is a FIA. The way that the FILA structure generalizes the FIA structure is by allowing investors to put previously earned interest credits at risk by selecting a floor rate below zero during the subsequent tracking period. This modification gives FILA investors the flexibility to choose their desired tradeoff between risk and return contingent on the realized crediting history of the contract.

1.1.1.1 Fixed Index Annuity(FIA)

In its most basic form, a FIA enables the holder to earn risk premia from an underlying index of publicly-traded securities, while eliminating downside risk to principal from declines in index value. Each year, a certain amount of the accumulation value is invested in the insurance company's general account, and the remainder (net of expenses) is used to purchase options with one year expirations on the underlying index. The portion in the general account guarantees the return of the accumulation value of the contract, and the options provide upside exposure to the index. The amount available to purchase options, called the options budget, depends on the expected return of the general account. Since general accounts primarily hold fixed-income securities, the options budget will depend, in large part, on the interest rate environment prevailing when the FIA is written. As interest rates decrease (increase), option budgets become lower (higher) resulting in decreased (increased) upside exposure to the index.

The potential share of the index return that the FIA will credit is specified either as a cap or as a participation rate. In the case of a cap, the structuring agent buys at-the-money (ATM) call options (European exercise style) on the underlying index in a notional amount equal to the contract's accumulation value, however the options budget will never be sufficient to cover the full cost of these options. So the agent simultaneously sells out-of-the-money call options on the underlying index so that the net option premium exactly equals the options budget (minus fees). Thus the cap is achieved using a long position in a call spread. In the case of a participation rate, only ATM call options are used, but in a notional amount that is less than the contract's accumulation value. The notional value of the call options is set so that the cost of the options exactly equals the options budget (minus fees). The participation rate is then the ratio of the notional value of the call options to the contract's accumulation value. At the end of each crediting period, the index value realized over the term is used to determine the option structure payoff and therefore the interest-credit realized by the contract for this period. We construct the FIA structures following Clark and Dickson(2021)[1]. We compare two FIA structures in this paper, CAP and PAR. PAR structure could achieve higher returns since it allows for investing the FIA premiums in at-the-money options as a whole. In contrast, CAP structure invests all of the premiums in at-the-money options while sells the out-of-money options to secure the minimum level of returns. Since both of these two FIA structures have the equity index as the underlying asset,i.e., for the same return of the equity return, PAR would be able to earn higher potential returns. Based on the structures, FIA has become an attractive investment especially when the interest rate is low. Except for the features of higher returns and secured minimum returns, the investors will also need to pay rider fees and surrender charge. The policyholder need to pay the rider fee for riders to guarantee a minimum level of returns within this contract. And if a policyholder would like to cancel the annuity contract or withdraw a large amount of money from the contract within the first couple of years, a surrender fee will be charged.

1.1.2 Fixed Index Linked Annuity (FILA)

Moenig and Samuelson(2023)[2] compares different Index Linked Annuities including FILA and considers FILA as a sub-type of FIA product. In the FILA contract, when the holder chooses an annual floor below zero, the floor is guaranteed by selling an at-the-money put option and buying an out-of-the-money put option with strike price corresponding to the chosen floor level. The combination of these two option legs is a short put spread. The net premium collected by the structuring agent from selling the put spread is added to the options budget, thereby allowing the strike price of the OTM call in the call spread to be increased in the case of a cap rate, or allowing an increase in the notional value of the ATM call options purchased in the case of a participation rate. The effect of these changes is to increase the cap rates and participation rates, respectively, and thus to increase the potential interest credit for that year.

Given historical time-series data on relevant interest rates and credit spreads, along with corresponding historical prices for options on a given index, it is possible to replicate historical FILA structures and, assuming competitive pricing and interest crediting terms, examine the realized historical performance of FILA contracts that could have been offered by insurance companies.[1] In this study, we will generate realistic hypothetical FILAs on the S&P 500 index, of various contract lengths, commencing on each day starting on January 4th, 1996. We will then be able to compare and contrast various FILA structures based on statistics calculated from the realized performance of these contracts. FILA is a type of FIA with unique features. FILA has the same features as FIA, investing in option positions and in general account, securing the minimum level of returns, keeping the principal. The unique characteristis include: first, different from traditional FIA strcture, FILA is a multi-year contract. FILA links the underlying asset, i.e., the option structure has the equity index and the allocation between option structures and equity index depends on the return achieved in the previous' year. We call the maximum level to invest in the option structures as 'floor rate'. Second, the FILA allows the investor to choose the minimum level of return. Specifically, for investors are willing to take higher risk, for which would like to choose lower minimum level of returns, they allocate more in the option positions and thus could achieve higher rate of returns.

[1] A similar exercise is carried out by Clark and Dickson(2021)[1] for FIA contracts

1.1.2.1 Options Structures for FILAs and FILA returns

The periodic interest credit for a FILA is a function $G_j(\cdot)$, $j = cap, par$ of the index return, $y = \frac{S_{t+1}}{S_t}$. At the end of the period, the accumulation value of the contract is reset to

$$A_{t+1} = A_t G_{cap}\left(\frac{S_{t+1}}{S_t}\right).$$

Let $C(S, K, \tau)$ and $P(S, K, \tau)$ be, respectively, the prices of European call and put index options, when the level of the underlying index is S, strike price is K, and time until expiration (in years) is τ. The risk-free rate of interest is r and the credit spread for the general account is ν.

The periodic payoff function for a FILA with cap rate, Γ, and floor rate, F, is

$$G_{cap}(y) = \begin{cases} 1+\Gamma, & y \geq 1+\Gamma \\ y, & 1-F \leq y < 1+\Gamma \\ 1-F, & y < 1-F \end{cases}$$

Given option prices at time t, and choice of floor rate, the cap rate Γ is such that

$$C(S_t, S_t, \tau) - C(S_t, S_t(1+\Gamma), \tau) + P(S_t, S_t(1-F), \tau) - P(S_t, S_t, \tau) = S_t(1 - e^{-(r+\nu)\tau}),$$

which can be solved numerically for the cap rate Γ.

The payoff function for a FILA with participation rate, δ, is

$$G_{par}(y) = \begin{cases} 1+\delta(y-1), & y \geq 1 \\ y, & 1-F \leq y < 1 \\ 1-F & y < 1-F \end{cases}$$

The participation rate, δ, satisfies

$$\delta C(S_t, S_t, \tau) + P(S_t, S_t(1-F), \tau) - P(S_t, S_t, \tau) = S_t(1 - e^{-(r+\nu)\tau}),$$

Using homogeneity of degree one of the call and put price functions, this simplifies to

$$\delta C(1, 1, \tau) + P(1, 1(1-F), \tau) - P(1, 1, \tau) = 1 - e^{-(r+\nu)\tau},$$

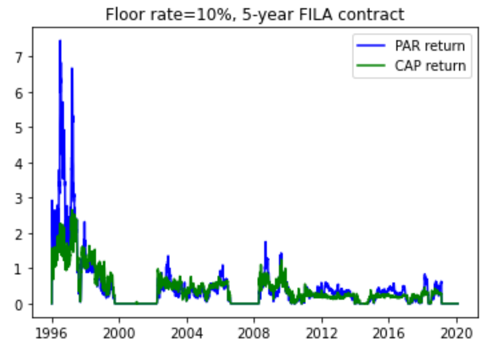
which can be solved explicitly for δ,

$$\delta = \frac{1 - e^{-(r+\nu)\tau} - P(1, 1(1-F), \tau) + P(1, 1, \tau)}{C(1, 1, \tau)}.$$

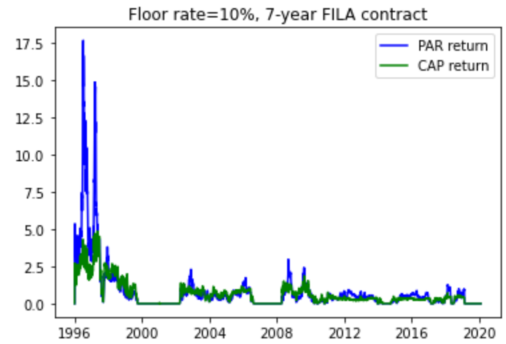
The Figure 1,2, and 3 represent the historical returns of FILA of 5,7, and 10 years, respectively. From these three figures, we can clearly see that, first, with the contract term increases, the returns of either Cap rate and Par rate increase. The average returns increase from 0.35 of a 5-year FILA contract to 0.95 of a 10-year contract. Second, when the absolute value of the floor rates increase, i.e., the potential to pursue higher returns increase, we can clearly find that the returns of FILA increase, whatever Cap or Par. The FILA exactly proves to be more valuable for investors since allowing more arbitrage opportunity to pursue higher returns in the following contract terms. Another important finding is that FILA is more valuable than the multi-year FIA contract. Compared with the same length of FIA, FILA can always achieve higher returns.

1.2 Methodology

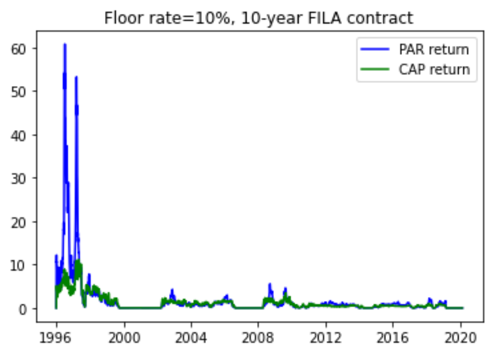
Fixed Indexed Linked Annuities (FILAs) have yet to become mainstream in the market. However, considering their ability to generalize the terms present in Fixed Indexed Annuities (FIAs), it is conceivable that most contract variations commonly found among FIA offerings could also be applied to FILAs. This study focuses on FILAs with a one-year point-to-point crediting method, based on either a participation rate (par rate) or a cap, with an annual



(a) 5-year FILA returns with floor rates= 10%



(b) 7-year FILA returns with floor rates= 10%



(c) 10-year FILA returns with floor rates= 10%

reset, serving as a standard for comparison. Modeling the historical performance of an FIA is a complex task as it requires more than merely observing the historical returns of an underlying index. As in the FIA case, the evolution of interest rates and index volatility are the primary determinants of FILA parameters, which can fluctuate considerably over time.

1.2.1 Modeling Historically Realistic Hypothetical FILAs

The participation rates and caps for newly written FILAs are influenced by the prevailing index volatility as well as the interest rate and credit environment. Historical option prices encapsulate relevant index volatility information (and vice versa). Capturing the effects of interest rates and credit spreads on historical FILA structuring accurately presents a considerable challenge. For FIAs, Ibbostson and Sinquefield(2018)[3] employed simulated participation rates supplied by AnnGen Development, LLC, a prominent authority in annuity product structuring. Regrettably, the specific calculation methodology is not publicly disclosed. Recently, Phau, in his paper, utilized the Moody's Seasoned BAA Corporate Bond Yield to approximate the performance of a representative insurer's general account.

He subtracted a constant spread of 2.28% to account for internal expenses. Clark and Dickson(2021)[1] approximated the general account return net of expenses using the yield to worst (YTW) of the Bloomberg Barclays US Investment Grade Corporate Bond Index. This index represents both a higher credit quality and shorter duration compared to the index referenced by Pfau(2017)[4]. In this study, a slightly more conservative approach is adopted, using the YTW of the Bloomberg Barclays Aggregate Bond Index (AGG). As a reference, the average fixed rate as of May 26, 2021, was 2.39%, while the YTW of our chosen proxy index stood at 2.11% on the same date.

To construct historical FILAs based on the S&P 500, we collected daily historical options volatility surface data from January 1996 through December 2021. Using the YTW of AGG plus the premium earned from selling the put spread as the options budget, we estimated both a participation rate and cap rate for each day on a rolling 1-year term linked to the S&P 500 price index. We created return series for such 1-year contracts using constant floor levels of 0%, −2.5%, −5.0%, −7.5%, and −10.0%. Subsequently, we developed a sequence of multi-year FILA contracts, starting each day in the sample, by determining annual interest credits referencing the 1-year contract with a chosen floor. At the beginning of each year of the multi-year contracts, we assumed that the holder opted for the lowest allowable floor to ensure that the accumulation value at year-end would not be less than the initial premium.

1.3 Empirical results

1.3.1 Multi-year contract comparison

Figures 1.2 and 1.3 graph the estimated cap and participation rates, respectively, for newly-written one-year FILAs for given floor values. Figures 1.4, 1.5, 1.6, and 1.7 graph the ending accumulation values for the given floor rates for 3-, 5-, 7-, and 10-year FILAs, respectively.

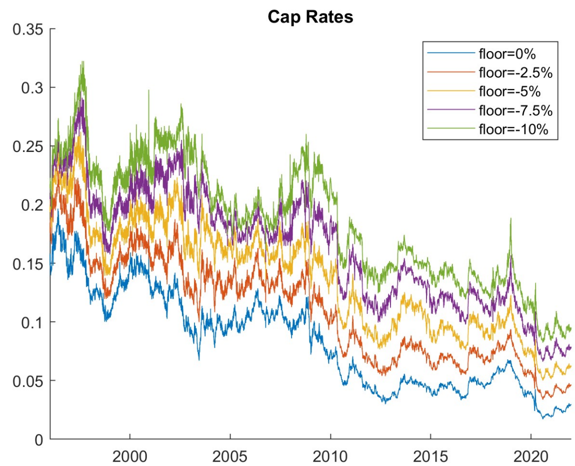


Figure 1.2: Cap rates for one-year point to point FILAs

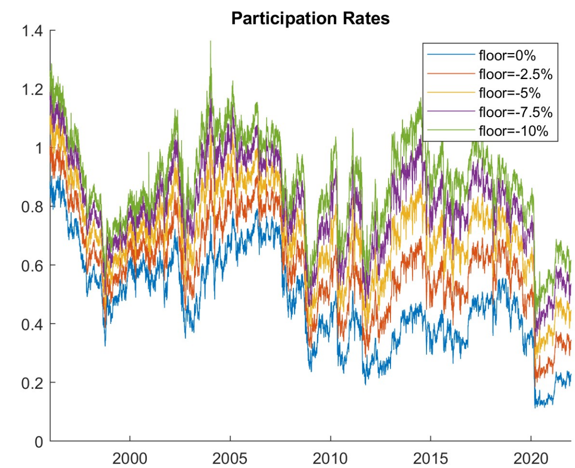


Figure 1.3: Participation rates for one-year point to point FILAs

1.3.2 Participation vs. Cap Rates

Table 1.2 reports the mean ending accumulation values per $1 of initial premium, the standard deviation of the accumulation value, and the mean compound annual growth rate, by number of contract years for both cap and par rate FILAs. Compared to cap rate FILAs, par rate FILAs produce greater mean ending accumulation values (and mean CAGRs) with lower standard deviations across all four contract maturities.

Over a slightly shorter time period (Jan 1996 - March 2021), Clark and Dickson(2021)[1] document average annual cap rate FIA returns of 5.5% (std 4.1%) and average annual par

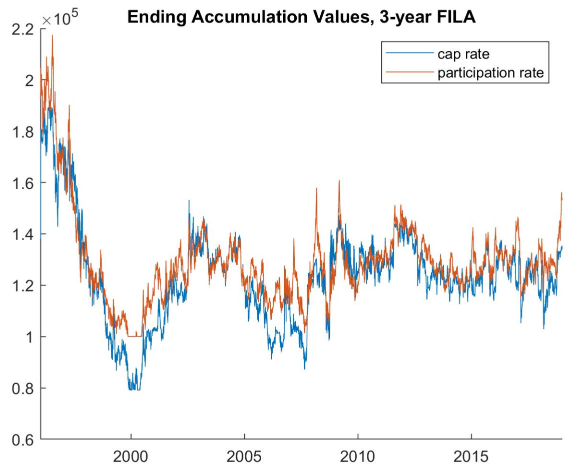


Figure 1.4: Ending Accumulation Values, 3-year FILA

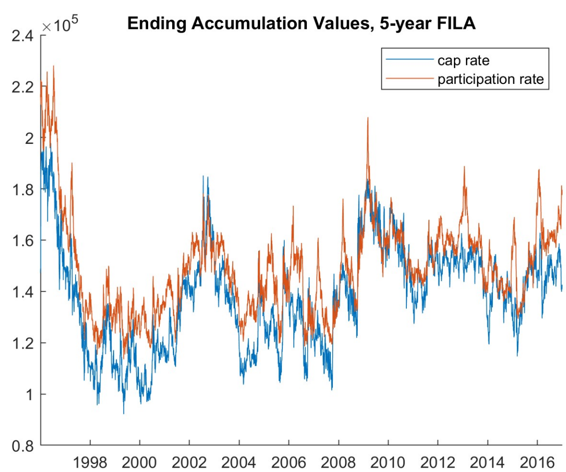


Figure 1.5: Ending Accumulation Values, 5-year FILA

rate FIA returns of 6.9% (std 6.9%). The FILA results in Table 1.2 illustrate the greater return potential and also the higher risk inherent in the FILA compared to the FIA. (To compare standard deviations reported in Table 1.2 to the annual return standard deviations reported in Clark and Dickson(2021)[1], we need to divide the standard deviations of ending accumulation values by $\sqrt{N}$ where N = contract years.) Table 1.2 documents that par rate FILAs across all contract lengths offer greater average returns while incurring lower risk when compared to the cap rate FILA structure. In terms of the relative performance of cap and par rate FILAs, these results are qualitatively similar to the cap and par rate comparison

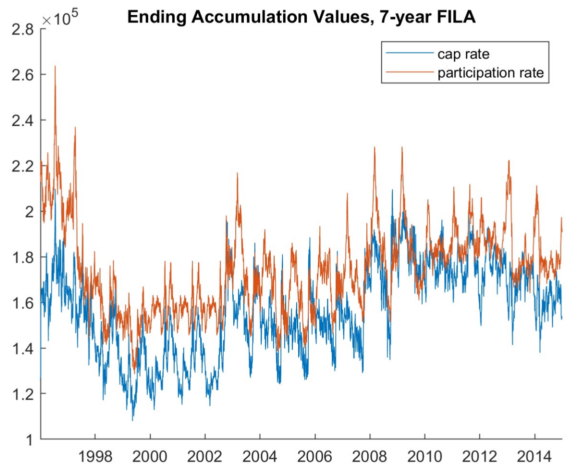


Figure 1.6: Ending Accumulation Values, 7-year FILA

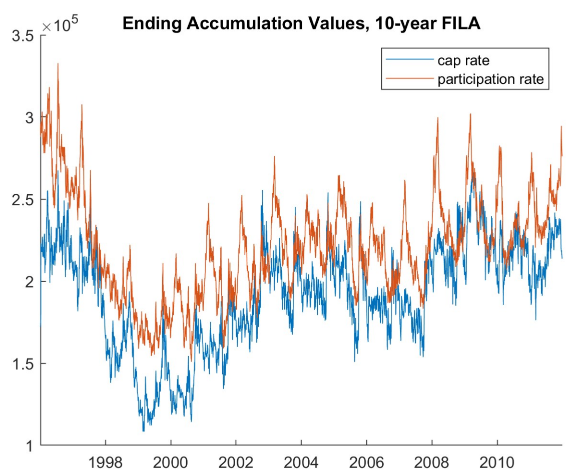


Figure 1.7: Ending Accumulation Values, 10-year FILA

for FIAs documented by Clark and Dickson(2021)[1].

Over a slightly shorter time period (Jan 1996 - March 2021), Clark and Dickson(2021)[1] document average annual cap rate FIA returns of 5.5% (std 4.1%) and average annual par rate FIA returns of 6.9% (std 6.9%). The FILA results in Table 1.2 illustrate the greater return potential and also the higher risk inherent in the FILA compared to the FIA. (To compare standard deviations reported in Table 1.2 to the annual return standard deviations reported in Clark and Dickson(2021)[1], we need to divide the standard deviations of ending accumulation values by $\sqrt{N}$ where N = contract years.) Table 1.2 documents that par rate

Table 1.1: Par vs. Cap Rates

		Cap Rate FILA			Par Rate FILA		
Contract Years	Number of Contracts	Mean Ending Value	Std Final Value	Mean CAGR	Mean Ending Value	Std Final Value	Mean CAGR
3	5789	1.2378	0.193	0.0738	1.2915	0.1736	0.089
5	5285	1.3835	0.202	0.0671	1.4948	0.1866	0.0837
7	4781	1.567	0.1975	0.0663	1.7425	0.1884	0.0826
10	4025	1.9159	0.3269	0.0672	2.2098	0.3122	0.0825

Table 1.2: Par vs. Cap Rates

		Cap Rate FILA			Par Rate FILA		
Contract Years	Number of Contracts	Mean Ending Value	Std Final Value	Mean CAGR	Mean Ending Value	Std Final Value	Mean CAGR
3	5789	1.2378	0.193	0.0738	1.2915	0.1736	0.089
5	5285	1.3835	0.202	0.0671	1.4948	0.1866	0.0837
7	4781	1.567	0.1975	0.0663	1.7425	0.1884	0.0826
10	4025	1.9159	0.3269	0.0672	2.2098	0.3122	0.0825

FILAs across all contract lengths offer greater average returns while incurring lower risk when compared to the cap rate FILA structure. In terms of the relative performance of cap and par rate FILAs, these results are qualitatively similar to the cap and par rate comparison for FIAs documented by Clark and Dickson(2021)[1].

1.3.3 Consumption-Portfolio Problems

This section explores the optimal consumption strategies for investors who hold FIA contracts. The investment strategy is demonstrated by solving a dynamic portfolio maximization problem, which in essence, is a multi-period preference framework. At present, investors have the choice to either consume or save their money in the account and seek higher returns through the option structures. Investors can withdraw money from the insurance product within a certain range (which is lower than the filter generating surrender fee) and avoid the risk from the underlying asset, the equity market index. Alternatively, some investors might decide to retain all the capital gains in the insurance account to chase potentially higher returns in the future. Our objective is to determine how to achieve utility maximization for these investors.

Future returns depend on the underlying asset and the option structures in the FIA. The cap rate measures the ratio of the options value to the total insurance premiums value.

1.3.4 FIA Payoff Function

For a cap rate FIA, the payoff function can be expressed as:

$$G_{cap}(y) = \begin{cases} 1+\Gamma, & y \geq 1+\Gamma \\ y, & 1 \leq y < 1+\Gamma \\ 1, & y < 1 \end{cases}$$

Here, Γ denotes the cap rate and y refers to the underlying asset, the equity index.

Let $C(S, K, \tau)$ be the Black-Scholes call price. We can solve for Γ using the following equation. The cap rate FIA is constructed by selling the out-of-the-money options and buying the at-the-money options with the full FIA premiums to secure the principal. The call spreads are used to purchase the underlying asset:

$$C(K, K, \tau) - C(K, K(1+\Gamma), \tau) = K(1 - e^{-(r+\nu)\tau}),$$

which simplifies to

$$C(1, 1, \tau) - C(1, 1(1+\Gamma), \tau) = 1 - e^{-(r+\nu)\tau}.$$

We can solve for Γ numerically.

For a Par rate FIA, the payoff function can be expressed as

$$G_{par}(y) = \begin{cases} 1+\delta(y-1), & y \geq 1 \\ 1, & y < 1 \end{cases}$$

where δ refers to the participation rate. We can solve for δ explicitly. Since the Par rate

FIA allocates all of the money in the at-the-money options, i.e., the amount of money less than the FIA premiums, the rest of the money is invested in the underlying asset to secure the principal.

$$\delta C(S_t, S_t, \tau) = S_t(1 - e^{-(r+\nu)\tau}),$$

so that

$$\delta = \frac{1 - e^{-(r+\nu)\tau}}{C(1, 1, \tau)}.$$

1.3.4.1 The Distribution of Underlying Index

The performance of the underlying asset determines the payoff of the FIA product. In our numerical method, we model the equity index using two different distributions.

Normal Distribution We first assume that the equity index follows a normal distribution. This assumption provides several benefits. Firstly, it ensures computational efficiency as we can easily obtain the option prices using the Black Scholes model. Secondly, the normal distribution serves as our benchmark, which roughly depicts the stock market index.

Jump-Diffusion Process We also use the Jump-Diffusion process to model the equity index. This model offers a couple of advantages as it allows for multi-dimensional effects on the stock prices. The jump event is a Poisson process with intensity λ, a parameter of jump size, and a parameter that measures the jump volatility. Through this distribution, we can have a more comprehensive way to measure the market condition changes and thus observe the effect on the FIA product.

Unlike the normal distribution that allows the application of the Black-Scholes model, we need to use the Fourier Transform to estimate the expected value of option prices. Following Lewis (2001)[5], we use the characteristic function below $\phi_T = \mathrm{E}[exp(izX_T)]$

$$\exp\{iz\omega T - \frac{1}{2}z^2\sigma^2 T + \lambda T(e^{iz\alpha - z^2\delta^2/2} - 1)\},$$

With the characteristic function, we use the integration formula to estimate the call option price

$$C(S,K,T) = Se^{-qT} - \frac{1}{\pi}\sqrt{SK}e^{-(r+q)T/2}\int_0^\infty Re[e^{iuk}\phi_T(u-\frac{i}{2})]\frac{du}{u^2+\frac{1}{4}}$$

1.3.5 Portfolio Utility Maximization

In this section, we derive the optimal strategies for consumption and savings in the FIA account. In our discrete-time model, in addition to the utility gains from FIA investment, we also consider the associated costs, such as the rider fee and the surrender fee. For multi-year contract utility maximization, it becomes crucial to determine how much money policyholders should consume and save within the FIA contract. Different from Moenig and Samuelson(2023)[2], which applies Constant Relative Risk Aversion (CRRA) framework to proxy investors' utility gains, we apply a multi- period utility framework which incorporates more features to model investors' utility gains.

The Bellman equation for this scenario can be expressed as

$$J(W_t, I_t) = A_t W_t = A(I_t) W_t$$

The following maximization problem applies:

$$A_t W_t = \max_{C_t, j\in\{cap,par\}} \left[(1-\beta)C_t^\rho + \beta(W_t - C_t - m)^\rho \mu(A_{t+1}G_j(\widetilde{R}_{t+1})|I_t)\right].$$

Here, m represents the sum of the income rider premium and the surrender fee.[2] We can also define some other important variables:

$$\alpha_t = \frac{C_t}{W_t}, \quad d_t = \frac{m}{W_t}, \quad 1-\alpha_t-d_t = \frac{W_t - C_t - m}{W_t}$$

[2]In the context of a FIA, an *income rider* is an optional feature that policyholders can add to their annuity contract to ensure a guaranteed lifetime income, regardless of the performance of the FIA's underlying investments. This feature usually comes with an additional annual fee, which is a percentage of the account value. On the other hand, a *surrender fee* is a charge that policyholders must pay if they choose to withdraw a portion or all of their money from the annuity before a certain period, typically within the first 5 to 10 years of the contract. This fee is in place to discourage early withdrawals.

In this context, C_t is the consumption amount at time t, W_t is the total wealth at time t, α represents the fraction of consumption to wealth, and d_t is the proportion of the FIA cost (rider premium and surrender fee) to the total wealth.

With these variables, the consumption and FIA structure choices are separable. Given an optimal B_t^*, we can express the following maximization problem:

$$A_t^* = \max_{\alpha_t} \left[(1-\beta)\alpha_t^\rho + \beta(1-\alpha_t-d_t)^\rho (B_t^*)^\rho\right]^{\frac{1}{\rho}}$$

By solving the first order conditions for optimal consumption α_t^*, we can write A_t as:

$$A_t^* = (1-\beta)^{\frac{1}{\rho}}(\alpha_t^*)^{1-\frac{1}{\rho}}(1-d_t)^{\frac{1}{\rho}}.$$

Then, we have:

$$\begin{aligned}
B_t^* &= \max_{j\in\{cap,par\}} \mu(A_{t+1}^* G_j(\widetilde{R}_{t+1})|I_t) \\
&= \max_{j\in\{cap,par\}} \mu((1-\beta)^{\frac{1}{\rho}}(\alpha_{t+1}^*)^{1-\frac{1}{\rho}}(1-d_{t+1})^{\frac{1}{\rho}} G_j(\widetilde{R}_{t+1})|I_t) \\
&= \max_{j\in\{cap,par\}} E\left[((1-\beta)^{\frac{1}{\rho}}(\alpha_{t+1}^*)^{1-\frac{1}{\rho}}(1-d_{t+1})^{\frac{1}{\rho}} G_j(\widetilde{R}_{t+1}))^\xi|I_t\right]^{\frac{1}{\xi}} \\
&= \max_{j\in\{cap,par\}} \left[\int_{-\infty}^{\infty}((1-\beta)^{\frac{1}{\rho}}(\alpha_{t+1}^*)^{1-\frac{1}{\rho}}(1-d_{t+1})^{\frac{1}{\rho}} G_j(1+x))^\xi g(x)\,dx\right]^{\frac{1}{\xi}}
\end{aligned}$$

At this stage, given the parameters for the index return distribution and the parameters of the recursive utility specification for a particular agent, we can solve for the agent's optimal consumption strategy and the optimal FIA structure (either a cap or participation rate).

1.3.5.1 FILA optimal consumption-portfolio

The consumption-portfolio problem under consideration now includes a floor-indexed life annuity (FILA) contract, which introduces a safety net into the investment, guaranteeing a minimum amount (the floor, denoted by F) that investors receive regardless of how poorly the underlying asset performs.

The payoff function for the FILA contract is defined by the variable $G_F(y)$. If the underlying asset return y is greater than or equal to 1, the payoff includes a participation rate δ which is the proportion of the gain in the underlying asset that the policyholder receives. For returns that fall between $1-F$ and 1, the payoff is directly equivalent to the asset return y. If the return drops below $1-F$, the payoff will be at the floor value of $1-F$. Thus in each crediting period, the payoff function for a FILA contract is given by

$$G_F(y) = \begin{cases} 1+\delta(y-1), & y \geq 1 \\ y, & 1-F \leq y < 1 \\ 1-F, & y < 1-F \end{cases}$$

The FILA purchaser seeks to solve the following problem:

$$\begin{aligned} B_t^* &= \max_F \left\{ E\left[((1-\beta)^{\frac{1}{\rho}} \alpha_{t+1}^{1-\frac{1}{\rho}} G_F(\widetilde{R}_{t+1}))^{\xi} | I_t \right]^{\frac{1}{\xi}} + b_0 E\left[\bar{\nu}(G_F(\widetilde{R}_{t+1}) - 1) | I_t \right] \right\} \\ &= \max_F \left\{ \left[\int_{-\infty}^{\infty} ((1-\beta)^{\frac{1}{\rho}} \alpha_{t+1}^{1-\frac{1}{\rho}} G_F(1+x))^{\xi} g(x)\, dx \right]^{\frac{1}{\xi}} \right. \\ &+ b_0 \int_{-\infty}^{\infty} \bar{\nu}(G_F(1+x) - 1) g(x)\, dx \end{aligned}$$

The variable B_t^* represents the optimal utility, which is maximized across different floor values F. The first term in the maximization problem, which includes the expectation of the future utility, considers how the FILA contract payoff $G_F(\widetilde{R}_{t+1})$ affects the future consumption level α_{t+1}. The variable B_t^* represents the optimal utility, which is maximized across different floor values F. The first term in the maximization problem, which includes the expectation of the future utility, considers how the FILA contract payoff $G_F(\widetilde{R}_{t+1})$ affects the future consumption level α_{t+1}.

To determine the optimal floor value F, we consider both the normally distributed returns and the jump-diffusion process. With the parameters of the utility function and the distribution of the returns, we can integrate numerically to solve for the optimal floor value F. This yields the optimal consumption strategy and FILA contract that maximize the investor's utility.

1.3.6 Utility gains for investors in a Narrow Framing framework

Barberis and Huang(2009) [6] introduced the concept of Narrow Framing as a means to measure utility, and Barberis , Huang and Thaler (2006)[7] further applies this concept to model investors' preferences under various scenarios. In this section, we compare the optimal strategy for investors. Previous studies of FIA mention that FIA could bring higher returns but ignore the utility perspective. Clark and Dickson(2021)[1], discuss that FIA provide higher returns, higher Sharpe Ratio, and keep similar level of risk to US Corporate Bond Index. One study finds that optimal portfolio allocation on FIA depends on investors' risk aversion in Constant Relative Risk Aversion (CRRA). However, there are some limits on the application. First, FIA structures could provide investors different choices to choose their investment behaviors. Different levels of consumption levels will lead to different preference choice. Obviously, CRRA does not satisfy the FIA features. Second, as a multi-year contract, not only the investors' risk aversion should be considered, the patience, and time preference also matter in the utility function. CRRA can only reflect the risk aversion, which is not enough. Thus, we apply a recursive preference utility model to analyze FIA. Another contribution is that we investigate the optimal strategy in the FILA consumption. In the multi-year contract, to maximize the utility gains is important to both the investors and to the insurance company as well. To more accurately reflect the market conditions, we model the underlying asset (the equity index) following two different distributions: normal distribution as a benchmark, while the jump diffusion process as a more close-to-reality distribution to model the equity index.

1.3.6.1 Normal distribution model

In Table 1.3 and 1.4, we assume the equity index follows normal distribution. Table 3 and 4 reflect the optimal choice in PAR/CAP and the optimal consumption level, which is shown in the alpha column.

When the market volatility is high, i.e., σ=0.15, the results show that for more risk-averse investors($\xi = 0.1$)[3], they prefer to choose CAP over PAR. (Note:Banal and Yaron(2004)[**?**] that 10 is supposed to be the upper bound for a risk aversion parameter.) The preference is consistent with the FIA structures. Since CAP, which allocates all of the annuity premium into the ATM options and sells the OTM options, is exposed less to the market risk (the volatility of equity index), the CAP could provide a more secure investment for risk-averse investors. While, the more risk-tolerant investors ($\xi = 0.9$), they prefer to choose PAR which only buys ATM call options and could provide higher potential returns. It is also interesting to see the relation between the time preference, β, and the optimal consumption ratio to wealth, α. When the investors are more willing to consume more today, i.e., lower β, the optimal consumption ratio, i.e. α, is higher. When β is 0.1, the optimal consumption ratio exceeds 90%. On the other hand, more patient investors who would like to retain more today and invest in the future, consume only 10%. The results from Table 1.3 show that the preference among CAP and PAR FIA depends on the investors risk aversion.

Table 1.4 reflects the preference in a low-volatility market, PAR is always preferred to CAP, no matter how risk-averse the investors are. It is consistent with the FIA structure, commonly considered as an alternative to bonds, when the market risk is low, investors are always willing to find a way to get more return. In this case, PAR, which only buys ATM call options, can provide the chance. In addition, the investors prefer to consume more when they have more certainty today.

1.3.6.2 Jump diffusion model

We further investigate the optimal choice for investors under more complex and realistic market conditions. Empirically, the equity market index tends to exhibit fat tails and inconsistencies with the assumptions of the Black-Scholes model. Thus, we test investor preference between CAP and PAR when the underlying equity index follows the Jump-Diffusion model

[3]Our choice of risk aversion parameter is reasonable for a long-run investors.

Table 1.3: Preference between CAP and PAR FIA in a more volatile market condition.

Panel A: Preferences of risk-averse investors ($\xi = 0.1$)						
Investor Type	β=0.1	β=0.3	β=0.5	β=0.7	β=0.9	β=1.0
Alpha	0.90	0.60	0.40	0.25	0.15	0.10
Preference	CAP	CAP	CAP	CAP	CAP	CAP
Panel B: Preferences of risk-tolerant investors ($\xi = 0.9$)						
Alpha	0.90	0.60	0.40	0.25	0.15	0.10
Preference	PAR	PAR	PAR	PAR	PAR	PAR

Note. The upper panel presents the preference of risk-averse investors ($\xi = 0.1$). The lower panel presents the preference of risk-tolerant investors ($\xi = 0.9$). The solid dot represents the preferred FIA structure. See Table 3 for the details.

Table 1.4: Preference between CAP and PAR FIA in a less volatile market condition.

Panel A: Preferences of risk-averse investors ($\xi = 0.1$)						
Investor Type	β=0.1	β=0.3	β=0.5	β=0.7	β=0.9	β=1.0
Alpha	0.90	0.60	0.40	0.25	0.15	0.10
Preference	PAR	PAR	PAR	PAR	PAR	PAR
Panel B: Preferences of risk-tolerant investors ($\xi = 0.9$)						
Alpha	0.90	0.60	0.40	0.25	0.15	0.10
Preference	PAR	PAR	PAR	PAR	PAR	PAR

Note. The upper panel presents the preference of risk-averse investors ($\xi = 0.1$). The lower panel presents the preference of risk-tolerant investors ($\xi = 0.9$). The solid dot represents the preferred FIA structure. See Table 1.4 for the details.

as described in Merton(1976)[8]. Specifically, we examine how utility changes with varying equity market conditions, such as the distribution of the underlying index under different market volatility levels, jump intensities, and expected jump sizes. Another motivation for using the Merton Jump-Diffusion process is that the jump component reflects the change in the stock price when new information arrives. This distribution can also help us understand the optimal choice of investors and their responses to stock market changes. Table 3 demonstrates that preference conversion occurs when market volatility is at a mid-level, with a 7% annual risk, a jump intensity of 1, and a jump size ranging from -0.1 to -0.5.

This is assuming that investors have mid-level risk aversion, prefer future consumption, and exhibit mid-level elasticity of intertemporal substitution (which measures responsiveness). We set up this model to understand how such representative investors react to various FIA structures.

When the market volatility is low, i.e., at a 1% level, PAR is consistently preferred to CAP. Meanwhile, the optimal consumption to wealth ratio remains above 90%, indicating that consumers prefer to utilize most of their wealth. (Refer to Appendix table for the results).

During periods of medium and high market volatility, the results become intuitively interesting. Customers tend to choose PAR when the jump intensity is low. However, when the jump intensity is high with a low jump size, PAR is preferred; but, if the jump size is larger, CAP is favored. In essence, the preference switch depends on the equity market index distribution. The results suggest that market conditions play a critical role in shaping investors' preferences.

As the jump size reflects the importance of new information arrival, the changes in preference also indicate that investors are extremely sensitive to market condition changes. Hence, FIA can indeed be a compelling investment opportunity for investors.

The first panel of Table 1.5 presents different investors' FIA preferences. Under typical market conditions based on historical data (i.e., when the annual market volatility is 12%, the jump intensity is 1.0, and the jump size is -0.3), investors consistently choose PAR-style FIA, regardless of whether they are patient or impatient investors. This result aligns with the FIA structure wherein PAR generally provides higher returns compared to CAP, as shown in the empirical results in Clark and Dickson(2021)[1].

The mid-panel of Table 1.6 indicates that different investors, with varying levels of risk aversion, show divergent preferences among FIA structures. The less risk-averse investors, signified by lower risk aversion parameters, prefer PAR. Conversely, the most risk-tolerant investors, with a risk aversion parameter of 0.9, choose CAP. The optimal consumption-wealth ratio is above 70% for both CAP and PAR, suggesting that investors tend to spend their wealth quickly. If there is a 10% return on the underlying asset, then CAP yields a 1.3% return while the PAR will earn 6.8%, which is less likely to happen since the historical jump size is usually a negative value. While, if there is a decrease in the underlying index, i.e., a negative return, then, the CAP holder is more beneficial since they spend more of their money (very fast) and with a lower cap rate which means suffers less than PAR. From the investors' point of view, since the CAP holder could always spend most of their wealth and thus earn utilities, thus keep at the similar utility gains level with the PAR. While, CAP can effectively avoid unnecessary loss from the market index. Thus, we can understand why the risk-tolerant investors are more willing to take the CAP.

The third panel compares the preference of different time-elasticity investors. The time elasticity parameter actually measures the reaction of the comsumption growth to the real interest rate. Particularly, this parameter reflect the net effect of real interest on the concurrent consumption. Investors will decrease today's consumption since the increased interest rate motivates the saving, while the increased consumption today will also bring more utility gains to the investors. Thus, if this elasticity of intertemporal substitution is high, the consumption is very sensitive to the change of interest rate, while a low value means the consumption is insensitive to the real interest rate change.

Table 1.7 presents the utility comparison between CAP and PAR. From the first panel, PAR is always prefered to CAP when the jump intensity is low,i.e., at 0.5 level. Similarly, in a more volatile market (when the market volatility is relatively high), the jump intensity is deterministic to investors making decisions. When the jump is more intense, specifically, at 1.0 and at 2.0 levels, CAP will be prefered when jump size is above mid-level. While, PAR is chosen when the jump size is low.

Table 1.5: This table reports the preference between CAP and PAR FIA under different market conditions. The underlying equity index following Jump diffusion stochastic process. The columns alpha represent the optimal consumption-wealth ratio. The solid dot represents the preferred FIA strucuture.; see Table 1.3 for the details.

sigma=0.07

jump intensity =0.5	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.01486158	0.01050198
-0.3		•	0.05620521	0.04347361
-0.5		•	0.0859304	0.08029846

sigma=0.07

jump intensity =1.0	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.0309287	0.0231316
-0.3	•		0.0538723	0.0568619
-0.5	•		0.1470095	0.1517829

sigma=0.07

jump intensity =2.0	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.03885589	0.03130449
-0.3	•		0.08655362	0.09123504
-0.5	•		0.29342821	0.29876722

Table 1.6: Investors preference between PAR and CAP This table presents the preference between CAP and PAR FIA of different investors and the optimal consuption-wealth ratio. The underlying equity index following Jump diffusion stochastic process. The columns alpha represent the optimal consumption-wealth ratio. The solid dot represents the preferred FIA structure.; see Table 1.4 for the details.

volatility=0.12	FIA structures and difference		alpha	
time preference	cap	par	cap	par
0.1		•	0.98192003	0.98189649
0.5		•	0.72800084	0.72741232
0.9		•	0.13552268	0.1336159

volatility=0.12	FIA structures and difference		alpha	
risk aversion	cap	par	cap	par
0.1		•	0.72841561	0.72768453
0.5		•	0.72800084	0.72741232
0.9	•		0.72781898	0.72799122

volatility=0.12	FIA structures and difference		alpha	
time elasticity	cap	par	cap	par
0.1		•	0.52976058	0.5296463
0.5		•	0.72800084	0.72741232
0.9		•	0.97890315	0.97889885

Table 1.7: Investors preference between PAR and CAP.This table reports the preference between CAP and PAR FIA under more volatile and different market conditions. The underlying equity index following Jump diffusion stochastic process. The columns alpha represent the optimal consumption-wealth ratio. The solid dot represents the preferred FIA structure.; see Table 1.6 for the details.

volatility=0.12

jump intensity =0.5	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.033635	0.029962
-0.3		•	0.050594	0.045543
-0.5		•	0.082430	0.077919

volatility=0.12

jump intensity =1.0	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.037658	0.035533
-0.3	•		0.059127	0.060261
-0.5	•		0.146691	0.153156

volatility=0.12

jump intensity =2.0	FIA structures and difference		alpha	
jump size	cap	par	cap	par
-0.1		•	0.041828	0.040165
-0.3	•		0.093221	0.098795
-0.5	•		0.277351	0.285460

Table 1.8: FILA contract and 5-year FIA returns

floor rates(%)	2.5	5	7.5	10	FIA
cap	0.35356	0.38636	0.38652	0.38449	0.30539
par	0.42547	0.47504	0.47028	0.46864	0.36647
diff	-0.0719	-0.08869	-0.08376	-0.08415	-0.06107

Table 1.9: FILA contract and 7-year FIA returns

floor rates(%)	2.5	5	7.5	10	FIA
cap	0.56635	0.61034	0.61709	0.61774	0.469
par	0.73149	0.80779	0.81339	0.82099	0.60049
diff	-0.16514	-0.19745	-0.19629	-0.20325	-0.13149

Table 1.10: FILA contract and 10-year FIA returns

floor rates(%)	2.5	5	7.5	10	FIA
cap	0.959823	1.059808	1.093561	1.108854	0.77544
par	1.431704	1.639481	1.71323	1.775979	1.130808
diff	-0.47188	-0.57967	-0.61967	-0.66713	-0.35537

1.4 Conclusion

FILA has garnered substantial interest among investors, chiefly due to its potential for higher returns along with a secured floor rate of return. This paper offers both comprehensive guidance for investors with diverse profiles and nuanced investment suggestions. Our derived optimal investment strategy prescribes an consumption plan, the optimum choice between FIA structures, and an intuitive rationale for investment choices amidst shifting market conditions. The cumulative value of multi-year contracts may serve as a valuable guidepost for both the insurance industry and investors.

We pioneer the application of recursive utility preference setup to model the utilities of FILA structures. Through this framework, we delineate the superior choices for patient investors, risk-averse investors, and those sensitive to interest rates. Consideration of such investment choices is pivotal when providing investment advice. As such, we have applied various mathematical approaches to model the underlying assets, ensuring a holistic evaluation of market factors. Our empirical findings underscore the influence of dynamic economic indicators on investment decisions. Most crucially, our evidence indicates that FILA can be a profitable investment avenue, accommodating a diverse range of investors across various economic cycles.

REFERENCES

[1] S. P. Clark and M. Dickson, "Fixed indexed annuities: Dissecting performance expectations," *Retirement Management Journal*, vol. 10, no. 1, pp. 36–45, 2021.

[2] T. Moenig and B. Samuelson, "A comparison of index-linked annuities," *North American Actuarial Journal*, pp. 1–22, 2023.

[3] R. G. Ibbotson and R. A. Sinquefield, *2019 SBBI® Yearbook: Stocks, Bonds, Bills, and Inflation®: US Capital Markets Performance by Asset Class 1926-2018*. Duff & Phelps, 2018.

[4] W. D. Pfau, *Safety-first retirement planning: An integrated approach for a worry-free retirement.* JSTOR, 2017.

[5] A. L. Lewis, "A Simple Option Formula for General Jump-Diffusion and other Exponential Levy Processes," Related articles explevy, Finance Press, Aug. 2001.

[6] N. Barberis and M. Huang, "Preferences with frames: A new utility specification that allows for the framing of risks," *Journal of Economic Dynamics and Control*, vol. 33, no. 8, pp. 1555–1576, 2009.

[7] N. Barberis, M. Huang, and R. H. Thaler, "Individual preferences, monetary gambles, and stock market participation: A case for narrow framing," *American economic review*, vol. 96, no. 4, pp. 1069–1090, 2006.

[8] R. C. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of financial economics*, vol. 3, no. 1-2, pp. 125–144, 1976.

CHAPTER 2: Conditional betas

Conventional practice that assumes constant betas in asset pricing models is likely to be misleading because it ignores changing firm attributes and/or economic conditions that lead to time-variation in betas. Existing attempts to model time-varying betas are confined by limited scope, curse of dimensionality, estimation errors, or lack of economic tractability. To address these issues, we compare several econometric or machine-learning methods and examine their performance in (1) tests of asset pricing models, (2) tests of market anomalies, and (3) out-of-sample return forecasts. Our results show that the new conditional betas improve the power of the asset pricing tests; in particular, models featuring conditional betas estimated using firm-level LASSO deliver significant risk prices and insignificant mispricing. When testing market anomalies, models with conditional betas estimated using panel regressions address more anomalies, such as value, turnover, and momentum. Finally, CAPM featuring our conditional betas outperform the unconditional CAPM in out-of-sample return forecasts. (JEL: C22; G12; G14; G17)

Keywords: conditioning information; conditional betas; asset pricing models; anamolies; return predictability; performance evaluation

2.1 Introduction

Asset pricing models are the foundation of finance. Researchers and practitioners use asset pricing models to obtain expected returns of assets. Portfolio managers use these expected returns to determine asset allocations to enhance portfolio performance. Clients of managed portfolios use asset pricing models to evaluate money managers' performance and determine their compensation. Corporate finance managers use asset pricing models to obtain cost of capital estimates that assist them to arrive capital budgeting decisions.

An asset pricing model is commonly expressed in the following factor form (see, e.g., Merton (1973) [1]):

$$E_t(R_{i,t+1}) = R_{f,t} + \sum_{k=1}^{K} \beta_{i,t}^k \lambda_t^k, \tag{2.1}$$

where $E_t(\cdot)$ is the conditional expectation given time-t information set, $R_{i,t+1}$ is asset i's return at time $t+1$, $R_{f,t}$ is the risk-free rate known at time t, $\beta_{i,t}^k$ is asset i's beta with the kth factor, known at time t, and λ_t^k is the factor premium of the kth factor, also known at time t. Such an asset pricing model can be motivated by the intertemporal choices of investors with conditioning information to optimize portfolio risk-return trade-off and hedging demand for state variables affecting consumption and investment opportunity sets.

Among the three components of asset pricing models, including a risk-free rate, betas, and factor premiums, betas are the only component that drives the cross sectional variation in expected returns. Arbitrage pricing theory (Ross (1976) [2]) suggests that betas represent the "systematic" portion of the asset risk because they capture the co-movement between asset returns and systematic factors.[1] Modern portfolio theory indicates that betas should be time-varying because they reflect investors' time-varying information sets. Intertemporal asset pricing models have time-varying betas, too, because consumption or investment opportunity

[1] An equivalent representation is the stochastic discount factor representation that postulates that the product of the stochastic discount factor and an asset excess return is zero in expectation conditionally, where the stochastic discount factor is linear in the factors with proportionality determined by conditioning information.

sets change over time. Hence, the conventional implementation that assumes constant betas is likely misleading because unconditional models are misspecified models. When a model holds conditionally, its unconditional counterpart does not hold. Clearly,

$$E[E_t(R_{i,t+1}) - R_{f,t}] = \sum_{k=1}^{K} E(\beta_{i,t}^k \lambda_t^k) \tag{2.2}$$

$$= \sum_{k=1}^{K} [E(\beta_{i,t}^k)E(\lambda_t^k) + \mathbf{cov}(\beta_{i,t}^k, \lambda_t^k)] \neq \sum_{k=1}^{K} E(\beta_{i,t}^k)E(\lambda_t^k). \tag{2.3}$$

The goal of this paper is to implement econometric and machine-learning methods to estimate the conditional betas which are functions of both firm characteristics and macroeconomic variables, so that they reflect firms' operational and financial status under changing economic conditions. We examine these novel conditional betas' performance in (1) tests of asset pricing models, (2) tests of asset pricing anomalies, and (3) return forecasts. Estimating betas has been a central issue in financial economics. While many researchers or practitioners assume unconditional betas, they are likely to be misspecified because, as described above, when a model holds conditionally, its naive fixed-beta version does not hold. To find conditional betas, researchers or practitioners may use some sort of filters to approximate conditional betas. For example, starting with Fama Macbeth(1973) [3], researchers use rolling estimates of betas to approximate the conditional betas, and such a method remains popular today. In practice, financial media such as Google, Yahoo, Value Line, etc., report rolling betas using a 36-month or 52-week rolling window. There are also other parametric or nonparametric filters, such as multivariate GARCH by Bollerslev, Engle and Wooldridge (1988) [4] and Markov Chain Monte Carlo and Gibbs sampling estimation method by Ang and Chen (2007) [5]. While these filters are certainly improvements over the unconditional counterpart, they typically do not have direct link to the firms' operations or financials, or to the macroeconomic conditions.

Starting in 1990s, betas begin to link to the real economy by researchers. Ferson and Harvey (1991) [6] find that risk premiums are related to the macroeconomic variables. This finding motivates the formulation of betas as *linear* functions of macroeconomic variables, such as level of interest rates, the spread between long-term and short-term bonds, the spread between low-quality and high-quality bonds, the dividend yields, among others. Such a linearity can be obtained from a first-order Taylor-series expansion. Ferson and Harvey (1999) [7] use conditional betas to perform asset pricing test and find the unconditional and conditional versions of Fama and French (1993) [8] three-factor model are strongly rejected. Ferson and Schadt (1996) [9], Christopherson, Ferson, and Glassman (1998) [10], and Buttimer , Chen, and Chiang (2012) [11] use conditional betas to capture the performance of mutual funds, pension funds, and real estate investment trusts, respectively, and obtain very different performance measures than the unconditional counterparts. Simin (2008) [12] uses conditional CAPM and Fama-French three-factor model to predict stock returns. All of the above models show that macroeconomic variables may matter in modeling conditional betas, but they ignore that the nature of the firms can be yet another channel for the time-variation in betas.

It is intuitive that firm characteristics should affect the stock beta of a company, when its operations, product lines, and product markets change over time. For example, when Nokia switched its focus from forestry to telecommunication, its beta must change accordingly. Theory and empirical analysis show that firm characteristics may drive the time-series and cross-sectional variation in betas. For example, Hamada (1972) [13] shows that firm betas in CAPM are linearly increasing in leverage. Gomes, Kogan, and Zhang (2003) [14] theoretically show that betas are functions of firm size and book-to-market ratio. Lin and Zhang (2003) [15] empirically show that betas are dominated by firm characteristics. Chiang (2016) [16] shows that firm characteristics, such as size, book-to-market, and momentum, affect stock betas. These findings reinforce the need to include not only macroeconomic variables but also firm-specific characteristics in conditional betas.

Avramov and Chordia (2006) [17] make the first attempt to model conditional betas as both firm characteristics and macroeconomic variables. Their scope is rather limited however:

it includes only two firm characteristic variables (size and book-to-market ratio) and one macroeconomic variable (default spread). With such a reduced-form formulation, Avramov and Chordia (2006) [17] find that only under a specific setting that conditional betas may help address some of the asset pricing anomalies. With the rapid growth of new drivers of the risk premiums in the past decades, the Avramov and Chordia (2006) [17] setting may need substantial extensions. However, including more firm attributes to beta modelling increases the number of parameters to be estimated. For example, if we want to include K factors, L macroeconomic variables, and M firm characteristics, we need to estimate $K \times L \times M$ variables, i.e., a curse of dimensionality issue.

The curse of dimensionality issue amplifies and compounds with yet another well-known issue in beta estimation: estimation errors. Researchers such as Fama and MacBeth (1973) [3] and Fama and French (1993) [8], among numerous others, form characteristic sorted portfolios and use them as test assets. While forming portfolios may mitigate estimation errors within a given portfolio, this method significantly reduces the power of the asset pricing tests because the number of observations is greatly reduced. Another method is to correct the biases directly, but these bias adjustments, e.g. Shanken (1992) [18], do not allow for conditional betas.

As a summary, the current state of the finance literature calls for conditional betas that (1) include both firm characteristics and macroeconomic information, (2) deliver rich and interpretable economic implications, (3) overcome the curse of dimensionality, and (4) mitigate estimation errors. With these objectives in mind, we propose to use panel regression or machine learning methods, see below for more details, to estimate the conditional betas and compare their performance. To the best of our knowledge, we are the first to use these dimension-reducing methods to estimate conditional betas whose variations are jointly driven by firm characteristics and macroeconomic variables. Such betas are useful in portfolio management, asset pricing, portfolio performance evaluation, and capital budgeting.

Regarding tests of asset pricing models, our results show that all unconditional models are rejected in tests of asset pricing models, due to the significant average excess return on zero-beta portfolio and to the lack of significant risk prices. Among the conditional models, models with conditional betas estimated using firm-level LASSO deliver more robust results, especially for asset pricing models with fewer factors.

For tests of asset pricing anomalies, conditional asset pricing models with betas conditional on the full set of firm characteristics and macroeconomic variables help address market anomalies, especially those with betas estimated using a panel regression and a "panel LASSO" method. However, size and short-term momentum effects remain pervasive. We further test anomalies considering conditional models with alpha conditional on macroeconomic variables. Our results show that "kitchen sink" and "individual LASSO" help address market anomalies including short-term momentum, while the size effects remain pervasive.

We also find promising return forecasting evidence using our conditional betas.

2.2 Methodology

2.2.1 Formulation of Conditional Betas

Following Avramov and Chordia (2006) [17], we assume betas are linear in L macroeconomic variables Z_t, in M firm attributes $X_{i,t}$, and in their interactions, where both Z and X vectors have the first element being 1. The beta of firm i with the kth factor is

$$\beta_{i,t}^{k} = (X_{i,t} \otimes Z_t)^{\top} b_i^{k}, \tag{2.4}$$

where $\otimes$ is the Kronecker product operator. Note that not only firm attributes and macroeconomic variables drive the time-series and/or cross-sectional variations in betas, but their interactions also drive the variations. Hence, we can ask whether the sensitivity of an attribute to beta may change over different economic conditions, e.g., "do small firms have higher betas when the aggregate default risk is higher?". Note this linear structure does not prohibit nonlinear transformations. For example, to include squared terms to capture concavity or convexity with respect to Z, one can simply augment Z to include both Z and

Z^2; to capture seasonality with respect to X, one can include $\sin(X)$ or $\cos(X)$.

We estimate the following K-factor model

$$R_{i,t+1} - R_{f,t} = a_i + \sum_{k=1}^{K} \beta_{i,t}^k f_{t+1}^k + \epsilon_{i,t}, t = 1, \ldots, T, \tag{2.5}$$

where f_k is the kth factor, and T is the number of observations.

2.2.2 Estimation Methods for Conditional Betas

When estimating Equation (2.5) without imposing any parameter restrictions or employing any dimensionality reduction method, we call it a "kitchen sink" model. While it is the full-fledged model, the kitchen sink model may be subject to substantial estimation errors. A common practice to reduce estimation error is to form portfolios, but it reduces the size of cross section substantially and lowers the power of statistical tests. In this paper, to reduce the dimensionality of the estimation problem and mitigate estimation errors, while preserving the size of the cross section without forming portfolios, we consider the following estimation methods:

1. Machine learning methods, such as LASSO regressions (Tibshirani (1996) [19]) that optimize

   $$\min_{a_i, b_i} \sum_{t=1}^{T} \left(R_{i,t+1} - R_{f,t} - a_i - \sum_{k=1}^{K} \beta_{i,t}^k f_{t+1}^k \right)^2, \tag{2.6}$$

   where $\beta_{i,t}^k = (X_{i,t} \otimes Z_t)^\top b_i^k$, subject to

   $$|b_i^k|^\top \mathbf{1} \leq \theta_i \tag{2.7}$$

   where θ_i is the penalty parameter, which is determined by a cross validation. We call this method "LASSO."

2. Panel regressions (i.e., time-series and cross-sectional pooled regressions): We assume

all b_i's are equal, i.e., $b_i = b_j \ \forall i \neq j$. Although this restriction is rather strong, the efficiency of the estimates is also significantly improved at the same time.

3. A mixture of the panel regression and the machine-learning techniques: run the LASSO pooled regressions. We term this method as "Panel-LASSO."

4. Instrumental Principal Component Analysis (IPCA): follow Kelly (2019) [20], we choose different numbers of factors in IPCA exercise. And then get respective latent factors and coefficients accordingly. This method is named 'IPCA' then.

As a comparison, Avramov and Chordia (2006) [17] run separate time-series regression (3.1) for each individual stock, without using panel regressions and/or machine learning methods to reduce the dimensions. Hence, Avramov and Chordia (2006) [17] need to employ a much smaller set of firm attributes and macroeconomic variables. Specifically, Avramov and Chordia (2006) [17] set $X_{i,t}$ to $(1, \text{SIZE}, \text{BM})$, and Z_t to $(1, \text{DEF})$.

We also include the unconditional betas in our analysis as a benchmark setting. Following the terminology of Avramov and Chordia (2006) [17], we call them "unscaled" betas.

2.2.3 Empirical Setup

We use monthly stock trading data from CRSP, and company accounting information from COMPUSTAT, of NYSE, AMEX, and Nasdaq-listed firms, for the period July 1963 to December 2019, to construct our dataset. We consider five predominant models in finance:

1. Sharpe (1964) [21] and Lintner (1965) [22] CAPM: the sole factor is the market excess return (MKT).

2. Fama and French (1993) [8] three-factor model: the factors are MKT, small-minus big (SMB, to capture the size effect), and high-minus-low (HML, to capture the value effect).

3. Carhart (1997)[23]: the factors are MKT, SMB, HML, and the momentum factor (MOM).

4. Fama and French (2015)[24] five-factor model: the factors are MKT, SMB, HML, robust-minus-weak (RMW, to capture operational profitability), and conservative-minus-aggressive (CMA, to capture investment effects).

5. q^5 model by Hou et al., (2019)[25] and Hou et al.(2021)[26]: the factors are MKT, SMB, spread of ROE-sorted portfolios (OP, to capture operating profitability), spread of investment-to-asset sorted portfolios (IA, to capture investment), and spread of expected growth sorted portfolios (EG). This model nests the q-factor model by Hou, Xue, and Zhang(2015)[27].

MKT, SMB, HML, MOM, RMW, and CMA are available from Kenneth French's website at Dartmouth College. OP, IA, and EG are available from Kewei Hou, Chen Xue, and Lu Zhang's website. Note that, while both Fama and French (2015)[24] five-factor model and q^5 model intend to capture profitability and investment effects, they use different measures for profitability (RMW vs. OP) and investment (CMA vs. IA).

IPCA approach could apply a comprehensive predictors affecting the stock returns, and then achieve dimension reduction by using pre-specified number of latent factors and coefficients. This approach could also cover substantial factors to predict stock returns. Following Ferson and Harvey (1999)[7], we consider the following macroeconomic variables: interest rate (one-month T-bill rate), dividend yield, default spread (yield spread between BBB and AAA-rated bonds), and term spread (yield spread between 10-year and one-year Treasury bonds). We also follow Goyal and Welch (2007)[28] and add aggregate book-to-market ratio, long-term yield, net equity expansion, overall stock variance, and inflation rate to the set of macroeconomic variables. The macroeconomic variables are available from Amit Goyal's website.

For the firm characteristics, motivated by Stambaugh, Yu and Yuan (2012)[29], we consider the following variables: size, book-to-market (the ratio of book equity to market equity), return on assets (as a profitability measure), momentum, and incremental investments scaled by assets (as an investment measure). We use Compustat and CRSP data to construct these firm characteristics; see the Appendix for details. Our selection of variables, compared with Avramov and Chordia (2006)[17] who only use two firm characteristics and one macroeconomic variable, results in a much larger set of variables that might potentially drive the time-series and cross-sectional variation in conditional betas.

2.2.4 Empirical Strategy

We consider four empirical exercises that exploit the conditional betas:

1. Cross-sectional regression tests of asset pricing models: We run cross sectional regressions of excess returns on conditional beta estimates to obtain estimates of λ, the "risk prices." We then test whether the (averages of) risk prices significantly different from zero. If a beta is with a risk factor (instead of being with a hedge), we expect the risk price to be significantly positive. This exercise extends Ferson and Harvey (1999) [7].

2. Cross-sectional tests of asset pricing anomalies: We run the "risk-adjusted returns" $(= R_{i,t+1} - R_{f,t} - \hat{a}_i - \sum_{k=1}^{K} \hat{\beta}_{i,k,t} f_{k,t+1})$, on the anomalies variables in Avramov and Chordia (2006) [17], including a Nasdaq dummy, book to market ratio, size, Nasdaq turnover, NYSE turnover, short term momentum, intermediate term momentum, and long-term momentum. This exercise extends Avramov and Chordia (2006) [17].

3. Out-of-sample return forecast: We begin with an "in-sample" period to estimate $\beta_{i,k,t}$ and forecast $f_{k,t+1}$, to obtain out-of-sample forecasts of $R_{i,t+1} - R_{f,t}$. Then we use the forecasts to construct portfolios and, after reiterating for many periods, assess the economic value—in terms of performance measures or equivalent management fees—of modeling a set of more comprehensive conditional betas. This exercise extends the return forecasting exercise in Simin (2008) [12]. Following Simin (2008) [12], in this study, we also implement the out-of-sample forecast using CAPM. Due to the high demand for computation, this exercise is implemented on CAPM which is the approach requires lowest computing power. Further study would be done in future research.

In the following sections, we exploit conditional betas in three applications.

2.3 Test of Asset Pricing Models

In this section, we conduct tests of asset pricing models: We run cross sectional regressions of excess returns on conditional beta estimates to obtain estimates of λ_t^k, the "risk prices." We then test whether the (averages of) risk prices are significantly different from zero. If a beta is with a risk factor (instead of being with a hedge), we expect its risk price to be significantly positive. This exercise extends the analysis of Ferson and Harvey (1999)[7].

Table 2.1 compares the results for Sharpe (1964)[21] and Lintner (1965)[22] CAPM. The average of risk price for the unconditional MKT factor beta is 0.3% with a t-ratio of 1.66, while the zero-beta asset has a significantly positive average excess return (0.5% per month with a t-ratio of 4.16). Throughout, the t-ratios are based on Fama and Macbeth (1973)[3] standard errors. Clearly, the unconditional CAPM is rejected. The average risk prices for the conditional MKT factor beta, modeled by Avramov and Chordia (2006) [17], is 0.7% per month (with a t-ratio of 4.24). The average risk price for the conditional MKT factor beta, estimated using the full set of firm characteristics and macroeconomic variables, is 0.5% per month with a t-ratio of 3.62. Employing LASSO selection in the conditional beta estimation helps to reduce the dimensionality of the estimation problem for the "kitchen sink" method. The average risk prices for the conditional MKT factor beta is 0.9% per month with a t-ratio of 4.28. And the zero-beta asset cannot generate abnormal return under our "Individual LASSO" method. The risk prices for MKT beta under panel regression and panel with LASSO are insignificantly different from zero. Overall speaking, conditional CAPM performs better than unconditional CAPM. Among them, models with Avramov and Chordia (2006)[17] and "individual LASSO" betas perform the best, because they have statistically significant risk prices and insignificant zero-beta premium.

Figure 2.1 plots the cross-sectional averages of CAPM betas throughout the sample period. The average of unscaled betas is smooth over time,[2] and does not vary with economic conditions. Avramov and Chordia (2006) [17] betas have more fluctuations, but they still mostly

[2] It is not constant over time due to the changing number of stocks over time.

around 1. Conditional betas by panel regression display different patterns from Avramov and Chordia (2006)[17]. Interestingly, when LASSO technique is employed, betas become less volatile. Figure 2.2 plots the the cross-sectional standard deviation throughout the sample period. Across all conditional methods, the volatility of MKT beta estimating under "kitchen sink" is the highest. Panel regression reduces the dispersion, and LASSO methods further reduce the volatilities. Figure 2.3 plots the risk prices of MKT beta for CAPM. They have larger ranges when conditional betas are estimated using firm-level LASSO, or panel-related methods.

Table 2.2 compares the results for Fama and French (1993)[8] three-factor models. None of the three factors has a significantly positive risk price for unconditional betas, and the zero-beta portfolio has an abnormal return of 0.4% per month (with a t-ratio of 2.41). Again, the unconditional Fama and French (1993)[8] three-factor model is rejected. Now turn to the conditional Fama and French (1993)[8] three-factor models. The average risk prices for SMB beta is significantly positive under the formulation of Avramov and Chordia (2006)[17], while the zero-beta asset can still generate significantly positive abnormal return of 0.6% per month. The risk price for all three factor betas (MKT, SMB, HML) are significantly positive when we use "kitchen sink" or "individual LASSO" to estimate conditional betas for each factor, showing that all of these conditional factor betas can drive the cross-sectional variation in stocks' expected returns. And the zero-beta asset yields a significantly negative return under these two beta-estimation methods (-0.17% with a t-ratio of -3.61 for "kitchen sink" and -0.10% with a t-ratio of -3.96 for individual LASSO). The average risk prices for SMB and HML beta are significantly positive using panel regression or panel regression with LASSO, showing that these two conditional betas can drive the cross-sectional variation in expected returns. The zero-beta asset does not generate abnormal returns under the conditional beta models estimating through panel or "panel LASSO" approaches. Overall speaking, none of the conditional Fama and French (1993)[8] three-factor model is supported by the data, although kitchen-sink and individual LASSO models work better.

The results for Carhart (1997)[23] four-factor model are in Table 2.3. Even adding the momentum factor (MOM), the unscaled four-factor asset pricing model cannot explain the cross-sectional variation in stocks' expected returns except for SMB with a marginally significant risk price of 0.3% per month with a t-ratio of 2.00. The zero-beta return is 0.4% per month with a t-ratio of 2.93. Once again the unconditional model is rejected. The conditional asset pricing model introduced by Avramov and Chordia (2006)[17] improves the explanatory power of factors a lot by including a limited set of firm characteristics and macroeconomic variables. All four conditional betas can positively explain the cross-sectional variation in expected returns. And the zero-beta return is insignificantly negative. Including more firm characteristics and macroeconomic variables, all four factors in the "kitchen sink" model can significantly drive the cross-sectional variation in stock's returns. And the abnormal return for the zero-beta asset is -0.02% with t-ratio of -0.93. Individual LASSO also performs well, although the risk price associated with size is merely marginally significant at 0.2% per month. Panel and Panel with LASSO methods show some improvements, too.

Similarly, the results of Fama and French (2015)[24] five-factor model show that the unconditional five factor betas do not drive the stock returns' cross-sectional variation (Table 2.2). The zero-beta asset can still generate abnormal return of 0.60% per month with a t-ratio of 6.68. Four factor betas except for MKT beta in the model of Avramov and Chordia (2006)[17] model can positively explain the cross-sectional variation, while the zero-beta return remains significantly positive (0.6% per month with a t-ratio of 2.99). Including more firm characteristics and macroeconomic variables do help: All factor betas except for SMB can positively explain the cross-sectional stock return variation under the "kitchen-sink" model. The zero-beta asset has a significantly negative return of -0.1% per month with a t-ratio of -7.08, however. Employing LASSO in the "kitchen sink" method, MKT and SMB betas have significantly positive risk prices, but the HML, CMA, and RMW betas do not deliver significant risk prices. The zero-beta asset cannot generate any abnormal return under this individual LASSO method. On the other hand, in panel regression and panel regression with LASSO, CMA beta has a significantly positive risk price, and the zero-beta asset does not generate significantly positive return. Overall speaking, conditional betas show some improvements.

Table 2.5 presents the results for the q^5 model by Hou, Mo, Xue, and Zhang (2019)[25] and Hou, Mo, Xue, and Zhang (2021)[26]. None of the five unconditional factor betas can significantly drive the cross-sectional stock return variation, while zero-beta asset can generate abnormal return with 0.6% per month. All five factors have significantly positive risk prices under "kitchen sink" and "individual LASSO". The OP beta has a significant risk price of 1.00% per month under panel regression. The risk prices for both OP beta and EG beta are significantly positive under "panel-LASSO" method. The zero-beta asset cannot generate abnormal return under panel regression and "panel LASSO" method.

We also perform the IPCA approach following Kelly, Pruitt, and Su (2019) [20], we test the IPCA asset pricing models of 1-, 2-, 3-, 4-, 5- factors,respectively. The 1-factor model achieves an average price of risk is 0.8% with a t-ratio of 3.54. And the intercept is insignificant with a t-ratio of -1.41. The three-factor model also has a insignificant intercept of 0.0112 with a t-ratio of 1.20. One out of three factor is priced with a t-ratio of 2.20. For 3- and 4-factor models, there are also factors priced with a t-ratio higher than 2. From the results, we can get to the conclusion that the IPCA conditional model also performs better than the unconditional models in the perspective of asset pricing model tests.

In summary, our cross-sectional regressions show that the unconditional models are uniformly rejected, due to the significant average excess return on zero-beta portfolio and to the lack of significant risk prices. Among the conditional formulations considered, individual LASSO delivers more robust results, especially for asset pricing models with fewer factors.

2.4 Test of Anomalies

2.4.1 Models with Time-invariant Alphas

In this section, we perform cross-sectional tests of asset pricing anomalies. We regress the "risk-adjusted returns" ($= R_{i,t+1} - R_{f,t} - \sum_{k=1}^{K} \hat{\beta}_{i,t}^{k} f_{t+1}^{k} = \hat{a}_i + \hat{\epsilon}_{i,t}$) on the anomaly variables in Brennan et al.,(1998)[30] and Avramov and Chordia (2006)[17], including a Nasdaq dummy, size, book to market ratio, NYSE turnover, Nasdaq turnover, short-term momentum, intermediate term momentum, and long-term momentum. This exercise extends Brennan, Chordia, and Subrahmanyam (1998)[30] and Avramov and Chordia (2006)[17].

We run cross-sectional regressions of monthly individual stock excess returns on anomalies in Table 2.13. Among these anomaly variables, coefficients on book-to-market, short-term momentum, intermediate term momentum, and long-term momentum are significantly positive; and the coefficient on size is significantly negative. We further study whether our conditional asset pricing model help address these anomalies.

Table 2.8 summarizes how Sharpe (1964)[21] and Lintner (1965) [22] CAPM works in addressing asset pricing anomalies. Avramov and Chordia (2006)[17] documents that unconditional and conditional versions of CAPM do not capture any of the size, book-to-market, turnover, and momentum effects in stock returns. We find the similar results with Avramov and Chordia (2006) [17]. Across all methods, the coefficients on size and NYSE turnover are significantly negative; and the coefficients on book-to-market, short term momentum, intermediate term momentum, and long-term momentum are significantly positive. Adding more sets of firm characteristic and macroeconomic variables when estimating the MKT beta, conditional CAPM can capture some impact of firm characteristics on risk-adjusted returns: The coefficient on book-to-market is insignificantly different from zero under "kitchen sink" method. The NYSE turnover is insignificantly different from zero under panel regression and panel with LASSO methods.

Table 2.9 summarizes that the unscaled Fama-French three-factor model performs similarly with the unscaled single-factor CAPM. Most coefficients are significant. When the Fama-French three factors are scaled by two firm characteristic variables and one macroeconomic variables, as shown in Avramov and Chordia (2006) [17], the coefficient on book-to-market is no longer significant. When factors are scaled by a larger set of firm characteristic and macroeconomic variables, the coefficient on book-to-market is no longer significant using "kitchen sink" and "individual LASSO", and the coefficients on book-to-market and NYSE turnover are not significant under panel regression and "panel LASSO" method.

The time-series averages of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on the lagged firm attributes for Carhart (1997) [23] four-factor model are in Table 2.10. Adding additional factors for unscaled model has limited improvement in addressing market anomalies, only the coefficient on NYSE turnovers is insignificant under Carhart (1997) [23] four-factor model. Note that the coefficient on NYSE turnovers is insignificant under single-factor CAPM using our panel regression and "panel LASSO" method (see Table 2.8), adding more variables when estimating conditional betas helps in improving the performance of asset pricing models with fewer factors. Regarding the conditional Carhart (1997) [23] four-factor model, the coefficient on book-to-market is insignificant for "individual LASSO"; the coefficients on book-to-market, NYSE turnover, and long-term momentum are no longer significant under panel regression and "panel LASSO" method.

The results for Fama and French (2015) [24] five-factor model reported in Table 2.11 show that adding two more factors (CMA and RMW) under the unscaled Fama-French three-factor model does not help in mitigating market anomalies. All of the firm characteristics, except for the Nasdaq dummy, are associated with significant coefficients in the unconditional Carhart (1997) [23] model. The coefficient associated with book-to-market become insignificant if we use the Avramov and Chordia (2006) [17] method to estimate the conditional betas, compared with the single-factor CAPM. Among the conditional beta models including the full set of firm characteristic and macroeconomic variables, the coefficient of book-to-market becomes insignificant under "kitchen sink" method; the coefficients of book-to-market, NYSE turnover, and intermediate term momentum become insignificant under "individual LASSO"; the coefficients of book-to-market, NYSE turnover become insignificant for panel regression and "panel LASSO" method.

Table 2.12 reports the results for q^5 model by Hou, Mo, Xue, and Zhang (2019) [25] and Hou, Mo, Xue, and Zhang (2021) [26]. Under unscaled q^5 model, the NYSE turnover becomes insignificant. The coefficient on book-to-market is insignificant under "individual LASSO", and the coefficients on book-to-market, NYSE turnover, intermediate term momentum, and long term momentum are insignificant under panel regression and "panel LASSO" method.

We perform the IPCA conditional models using 1-,2-,3-,4-,and 5- factors. The one-factor model cannot capture most of the market anomalies including firm size, book to market ratio, and return momentum. Conditional beta models with 3 factors have an insignificant coefficients of size, and return momentum. And 4-factor model can also effectively capture return momentum, i.e., the coefficients of RET23, RET46 and RET712 are insignificant. The 5-factor model can capture size, BM ratio, and return momentum effectively, with t-ratios lower than 2 constantly.

Overall, the conditional models with more factors help address market anomalies. Especially the 5-factor model can capture the famous market anomalies. Another finding is that the adjusted R-square decreases with the number of factors included. The 5-factor model achieves a 3.85% R-square and shows that conditional models can capture anomalies effectively. Overall, conditional-beta asset pricing models scaling with more firm characteristics and macroeconomic variables help address market anomalies, especially in our panel regression and "panel LASSO" methods. However, size and short-term momentum effects remain pervasive.

2.4.2 Models with Conditional Alphas

Avramov and Chordia (2006) [17] find that, when allowing alpha to be time-varying with macroeconomic variables, conditional asset pricing models may mitigate market anomalies further. Thus, we reconsider the asset pricing models, while allowing for conditional alphas, and then test the market anomalies.

We first run the following time-series regression

$$R_{i,t+1} - R_{f,t} = a_i^\top Z_t + \sum_{k=1}^{K} \beta_{i,t}^k f_{t+1}^k + \epsilon_{i,t}, t = 1, \dots, T, \tag{2.8}$$

where $\beta_{i,t}^k = (X_{i,t} \otimes Z_t)^\top b_i^k$, and a_i is an $L \times 1$ vector with the first element being a_{i1}. Then we run cross-sectional regression of $\hat{a}_{i1} + \hat{\epsilon}_{i,t}$ on a constant and the anomalies, and compute the average each slope coefficient and its standard error. For comparison, we consider models with constant betas (unscaled) or time-varying betas, which are estimated using Avramov and Chordia (2006) [17], kitchen-sink, LASSO, panel, and panel LASSO methods.

Table 2.14 summarizes how Sharpe (1964) [21] and Lintner (1965) [22] CAPM with conditional alpha works in addressing asset pricing anomalies. When taking into account of time-varying alphas without modeling time-variation in betas, coefficients on NYSE turnover and long-term momentum are insignificant. Under Avramov and Chordia (2006) [17] formulation, where betas are scaled by firm size, book-to-market, and default spread, the coefficients on book-to-market and NYSE turnover are insignificant. "Kitchen sink" CAPM mitigates short-term, intermediate-term, and long-term momentum, where coefficients on these anomalies become insignificant. "Individual LASSO" performs similarly with "kitchen sink", except that the coefficient on long-term momentum is significant. Panel regression and "panel LASSO" do not help address any of the asset pricing anomalies. Again, CAPM with conditional alpha and beta cannot explain all market anomalies. Among them, "kitchen sink" CAPM with conditional alpha and beta works best, which mitigate the short-term, intermediate-term, and long-term momentum.

Table 2.15 presents results for the Fama and French (1993) [8] three-factor model with a time-varying alpha. The coefficient on book-to-market becomes insignificant for unscaled Fama and French (1993) [8] three-factor model, compared with the unscaled single-factor CAPM. Among all methods with conditional betas, coefficients on size, book-to-market, and NYSE turnover are insignificant for Avramov and Chordia (2006) [17] framework; coefficients on size, short-term momentum, and long-term momentum are insignificant for "kitchen sink" and "individual LASSO"; coefficients on book-to-market are insignificant for panel regression and "panel LASSO."

The time-series averages of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on the lagged firm attributes for Carhart (1997) [23] four-factor model are in Table 2.16. Adding momentum factor in the Fama and French (1993) [8] three-factor model with conditional alpha and beta does not help mitigate market anomalies. Among all methods, "kitchen sink" and "individual LASSO" perform better, where coefficients associated with book-to-market and short-term momentum are insignificant.

The results for Fama and French (2015) [24] five-factor model with conditional alpha, reported in Table 2.17, are very similar with the Fama-French three factor model. Coefficients associated with book-to-market, NYSE turnover, and long-term momentum are insignificant for unscaled model; coefficients on NYSE turnover, intermediate-term momentum, and long-term momentum are insignificant. Both "kitchen sink" and "individual LASSO" help in mitigating book-to-market, turnover, and momentum anomalies, while the coefficient on size anomaly is still significant. Panel regression and "panel LASSO" help in mitigating book-to-market and turnover anomalies, while coefficients on size and momentum are still significant.

Table 2.18 reports the results for q^5 model by Hou, Mo, Xue, and Zhang (2019) [25] and Hou, Mo, Xue, and Zhang (2021) [26]. Similarly to unscaled CAPM, coefficients on book-to-market, NYSE turnover, and long-term momentum are insignificant for unscaled q^5 model. Coefficients on Nasdaq turnover and intermediate-term momentum become insignificant under Avramov and Chordia (2006) [17] method. Among our four methods, "kitchen sink" and "individual LASSO" can help in mitigating all anomalies except for size. Panel regression and "panel LASSO" can help in mitigating book-to-market and turnover anomalies, while coefficients on size and momentum remain significant.

Overall, conditional-beta asset pricing models scaling with more firm characteristics and macroeconomic variables help address market anomalies, especially in our "kitchen sink" and "individual LASSO" method. However, size effects remain pervasive.

2.5 Return Forecasts

In this section, we form out-of-sample return forecast: We begin with a 60-month "in-sample" period to estimate $\beta_{i,t}^k$ and forecast f_{t+1}^k, to obtain step-ahead out-of-sample forecasts of $R_{i,t+1}-R_{f,t}$. We use recursive estimation window forward to obtain more step-ahead excess return forecasts. Then we compute the root mean square forecast error (RMSFE) for each stock to assess each method's performance. This exercise extends the return forecasting exercise of Simin (2008) [12], who finds that unconditional models have better forecasting performance than conditional models.

Our results are in sharp contrast to the results of Simin (2008) [12]. Table 2.19 reports the RMSFEs of return forecasts produced by unconditional CAPM, and conditional CAPM featuring betas estimated by Avramov and Chordia (2006) [17] model, kitchen sink model, individual LASSO, panel regression, and a mix of panel regression and LASSO. The cross-sectional mean (median) of RMSFEs by the unconditional CAPM is 9.38% (8.54%), with a standard deviation of 3.19%. The Avramov and Chordia (2006) [17] model is outperformed by the unconditional model: its cross-sectional mean (median) of RMSFEs is 9.58% (8.59%), with a larger dispersion of 3.32%. However, our conditional betas improve the forecasting performance: all versions of our conditional models produce lower average RMSFEs with smaller dispersions. For example, when we use panel-LASSO approach to beta estimation, the average RMSFE is 8.85%, with a standard deviation of 3.08%.

2.6 Conclusion

We implement econometric and machine-learning methods to estimate the conditional betas which are functions of both firm characteristics and macroeconomic variables and compare how these conditional betas perform in tests of asset pricing models, tests of asset pricing anomalies, and return forecasts. We consider four estimation methods: "kitchen sink" method includes seven firm characteristic variables and eight macroeconomic variables to estimate conditional betas. "Individual LASSO", which helps reduce the dimensionality of the estimation problem for "kitchen sink" method, selects the optimal set of variables for each firm when estimating conditional betas. Panel regression can significantly improve the efficiency of the estimates. And lastly "panel LASSO" is a combination of panel regression and the machine-learning techniques. We compare their performance with unscaled asset pricing model and Avramov and Chordia (2006) [17] model.

Our results show that unconditional models are strongly rejected, due to the significant average excess return on zero-beta portfolio and to the lack of significant risk prices. Conditional models work better, especially models with betas estimated using "individual LASSO".

Regarding tests of market anomalies, unconditional models with fewer factors cannot address size, value, momentum, and turnover effects. Value and turnover effects are gone when there are five factors in the asset pricing models. On the other hand, conditional models help address the anomalies, especially the panel-related methods, except that size and short-term momentum are still pervasive.

Considering time-varying alpha when estimating conditional betas may also help address anomalies. Both "kitchen sink" and "individual LASSO" mitigate most market anomalies except for the size effect when there are five factors in the asset pricing model.

Finally, in terms of out-of-sample return predictive power, while the unconditional CAPM outperforms the Avramov and Chordia (2006) [17] version of conditional CAPM, all versions of our conditional CAPM uniformly outperform the unconditional CAPM, with lower RMSFEs. This evidence is in sharp contrast with existing evidence, e.g., Simin (2008) [12].

In sum, our new conditional betas improve the power of the asset pricing tests, mitigate asset pricing anomalies, and form better return forecasts.

2.7 Appendix

2.7.1 Variable Construction

We form the following firm characteristic variables, using data available from CRSP and Compustat:

- Firm size: the natural logarithm of market value of equity measured in billions of dollars. The market value of equity is calculated as the product of share price and the number of shares outstanding.

- BM ratio: the natural logarithm of the book-to-market ratio. We use book and market value of equity in the end of year $t-1$ to calculate the book-to-market ratio for July of year t to June of year $t+1$. If the book-to-market ratio is greater than the 0.995 fractile or less than the 0.005 fractile, we set them equal to the 0.995 and 0.005 fractile value.

- Turnover: the natural logarithm of monthly turnover. The monthly turnover is trading volume divided by the number of shares outstanding.

- ROE: the natural logarithm of one plus return on equity, which is measured as the income before extraordinary items divided by one-year lagged book value of equity. The ROE from July of year t to June of year $t+1$ is computed from ROE at the end of year t.

- ROA: the natural logarithm of one plus return on assets, which is measured as the income before extraordinary items divided by one-year lagged total assets. The ROA from July of year t to June of year $t+1$ is computed from ROA at the end of year t.

- Momentum: the natural logarithm of prior cumulative stock return from $t-12$ to $t-2$ month.

- Investments: the natural logarithm of the change of total assets divided by one-year lagged total assets. The Investments from July of year t to June of year $t+1$ is computed from Investments at the end of year t.

We employ the following macroeconomic variables using data available from Amit Goyal's website:

- Aggregate dividend yield: the dividend yield for the S&P 500 index is measured as 12 month moving sum of dividends divided by the price for S&P 500.
- Aggregate book-to-market ratio: the book-to-market ratio for the Dow Jones Industrial Average.
- Treasury bills: 3-month Treasury-bill rates.
- Default spread: the difference between AAA-rated and BAA-rated corporate bonds yields.
- Term spread: the spread between 10-year Treasury yield and one-month Treasury bill rate.
- Aggregate net equity expansion: the ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year capitalization of NYSE stocks.
- Aggregate stock variance: the sum of squared daily returns on S&P 500.
- Inflation rate: the growth rate of CPI.

REFERENCES

[1] R. C. Merton, "An intertemporal capital asset pricing model," *Econometrica*, vol. 41, pp. 867–887, 1973.

[2] S. A. Ross, "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, vol. 13, pp. 341–360, 1976.

[3] E. F. Fama and J. D. MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, vol. 81, pp. 607–636, 1973.

[4] T. Bollerslev, R. Engle, and J. Wooldridge, "A capital asset pricing model with time-varying covariances," *Journal of Political Economy*, vol. 96, pp. 116–131, 1988.

[5] A. Ang and J. Chen, "CAPM over the long run: 1926â2001," *Journal of Empirical Finance*, vol. 14, pp. 1–40, 2007.

[6] W. E. Ferson and C. R. Harvey, "The variation of economic risk premiums," *Journal of Political Economy*, vol. 99, pp. 385–415, 1991.

[7] W. E. Ferson and C. R. Harvey, "Conditioning variables and the cross section of stock returns," *Journal of Finance*, vol. 54, pp. 1325–1360, 1999.

[8] E. F. Fama and K. R. French, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, vol. 33, pp. 3–56, 1993.

[9] W. E. Ferson and R. Schadt, "Measuring fund strategy and performance in changing economic conditions," *Journal of Finance*, vol. 51, pp. 425–462, 1996.

[10] J. A. Christopherson, W. E. Ferson, and D. A. Glassman, "Conditioning manager alphas on economic information: Another look at the persistence of performance," *Review of Financial Studies*, vol. 11, pp. 111–142, 1998.

[11] R. J. Buttimer, J. Chen, and I.-H. E. Chiang, "REIT performance and market timing ability," *Managerial Finance*, vol. 38, pp. 249–279, 2012.

[12] T. Simin, "The poor predictive performance of asset pricing models," *Journal of Financial and Quantitative Analysis*, vol. 43, pp. 355–380, 2008.

[13] R. S. Hamada, "The effect of the firm's capital structure on the systematic risk of common stocks," *Journal of Finance*, vol. 27, pp. 435–486, 1972.

[14] J. Gomes, L. Kogan, and L. Zhang, "Equilibrium cross section of returns," *Journal of Political Economy*, vol. 111, pp. 639–732, 2003.

[15] X. Lin and L. Zhang, "The investment manifesto," *Journal of Monetary Economy*, vol. 60, pp. 351–366, 2013.

[16] I. E. Chiang, "Skewness and coskewness in bond returns," *Journal of Financial Research*, vol. 39, pp. 145–178, 2016.

[17] D. Avramov and T. Chordia, "Asset pricing models and financial market anomalies," *Review of Financial Studies*, vol. 19, pp. 1001–1040, 2006.

[18] J. Shanken, "On the estimation of beta-pricing models," *Review of Financial Studies*, vol. 5, pp. 1–33, 1992.

[19] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, pp. 267–288, 1996.

[20] B. T. Kelly, S. Pruitt, and Y. Su, "Characteristics are covariances: A unified model of risk and return," *Journal of Financial Economics*, vol. 134, no. 3, pp. 501–524, 2019.

[21] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance*, vol. 19, pp. 425–442, 1964.

[22] J. Lintner, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Statistics*, vol. 47, pp. 13–37, 1965.

[23] M. M. Carhart, "On persistence in mutual fund performance," *Journal of Finance*, vol. 52, pp. 57–82, 1997.

[24] E. F. Fama and K. R. French, "A five-factor asset pricing model," *Journal of Financial Economics*, vol. 116, pp. 1–22, 2015.

[25] K. Hou, H. Mo, C. Xue, and L. Zhang, "Which factors?," *Review of Finance*, vol. 23, pp. 1–35, 2019.

[26] K. Hou, H. Mo, C. Xue, and L. Zhang, "An augmented q-factor model with expected growth," *Review of Finance*, vol. 25, pp. 1–41, 2021.

[27] K. Hou, C. Xue, and L. Zhang, "Digesting anomalies: An investment approach," *Review of Financial Studies*, vol. 28, pp. 650–705, 2015.

[28] A. Goyal and I. Welch, "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, vol. 21, pp. 1455–1508, 2007.

[29] R. F. Stambaugh, J. Yu, and Y. Yuan, "The short of it: Invest or sentiment and anomalies," *Journal of Financial Economics*, vol. 104, pp. 288–302, 2012.

[30] M. J. Brennan, T. Chordia, and A. Subrahmanyam, "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns," *Journal of Financial Economics*, vol. 49, pp. 345–373, 1998.

Table 2.1: Test of Asset Pricing Model: CAPM

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the beta associated with the market (MKT) factor. t-ratios, based on Fama and Macbeth (1973) [3] standard errors, are reported below the averages and in parentheses. Betas are estimated using different versions of time-series regressions. "Unscaled" betas are time-invariant, or unconditional, betas. Five versions of conditional betas, that are functions of firm characteristics, macroeconomic variables, and their interactions, are considered. "Avramov and Chordia" betas consider only two characteristics (size and book-to-market) and one macroeconomic variable (default premium). "Kitchen Sink" betas include all characteristics and macroeconomic variables. "Individual LASSO" betas are estimated using a LASSO with cross-validation for each stock. "Panel" assumes that all time-series regression coefficients being equal across all stocks. "Panel with LASSO" uses a LASSO with cross-validation when running the panel regression to obtain betas.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Constant	0.51	0.15	0.24	0.06	0.46	0.97
	(4.16)	(1.32)	(2.06)	(0.51)	(0.72)	(1.90)
MKT	0.32	0.71	0.53	0.92	0.38	0.14
	(1.66)	(4.24)	(3.62)	(4.28)	(0.91)	(0.23)

Table 2.2: Test of Asset Pricing Model: Fama and French (1993) [8] Three-factor Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the beta associated with market (MKT), size (SMB), and value (HML) factors. See Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Constant	0.41	0.61	-0.17	-0.14	-0.06	-0.12
	(2.41)	(2.91)	(-3.61)	(-3.96)	(-0.11)	(-0.20)
MKT	0.41	0.01	0.51	0.53	0.34	0.41
	(1.34)	(1.55)	(3.46)	(2.83)	(0.07)	(0.01)
SMB	0.22	0.33	0.25	0.22	1.65	1.63
	(1.84)	(4.07)	(2.22)	(2.16)	(3.67)	(3.75)
HML	-2.69	-0.00	0.26	0.35	0.74	0.71
	(-0.95)	(-0.30)	(2.55)	(2.70)	(4.80)	(4.79)

Table 2.3: Test of Asset Pricing Model: Carhart (1997) [23] Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the beta associated with market (MKT), size (SMB), value (HML), and momentum (MOM) factors. See Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Constant	0.41	-0.15	-0.02	0.00	2.59	2.41
	(2.93)	(-1.84)	(-0.93)	(0.37)	(3.17)	(3.26)
MKT	0.33	0.66	0.55	0.54	2.36	2.22
	(1.62)	(3.93)	(3.03)	(2.83)	(3.06)	(3.09)
SMB	0.24	0.41	0.26	0.24	0.51	0.51
	(2.00)	(2.86)	(2.03)	(1.84)	(1.72)	(1.48)
HML	-0.16	0.25	0.37	0.34	0.36	0.36
	(-1.06)	(1.98)	(2.59)	(2.67)	(1.93)	(1.87)
MOM	0.39	0.64	0.64	0.69	1.47	1.12
	(1.29)	(3.66)	(3.76)	(3.85)	(5.59)	(5.52)

Table 2.4: Test of Asset Pricing Model: Fama and French (2015) [24] Five-factor Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the beta associated with market (MKT), size (SMB), value (HML), operating profitability (RMW), and investment (CMA) factors. See Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Constant	0.61	0.63	-0.14	0.17	1.57	0.21
	(6.68)	(2.99)	(-7.08)	(0.82)	(0.59)	(0.06)
MKT	0.22	0.01	0.55	0.44	0.91	0.20
	(0.99)	(1.04)	(2.96)	(5.54)	(0.52)	(0.10)
SMB	0.23	0.34	0.26	0.23	0.74	0.66
	(1.69)	(4.15)	(1.83)	(3.54)	(0.77)	(0.79)
HML	-0.17	0.00	0.38	0.06	1.42	0.90
	(-0.66)	(0.53)	(2.71)	(0.01)	(0.81)	(1.25)
CMA	-0.12	0.12	0.22	-0.00	1.46	1.42
	(-1.46)	(2.13)	(3.45)	(-0.71)	(2.73)	(3.26)
RMW	0.02	0.11	0.24	-0.00	2.32	2.51
	(0.17)	(2.52)	(2.79)	(-0.28)	(1.56)	(2.32)

Table 2.5: Test of Asset Pricing Model: q^5 Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the beta associated with market (MKT), size (SMB), operating profitability (OP), investment (IA), and expected growth (EG) factors. See Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Constant	0.61	0.12	0.12	0.12	1.63	1.71
	(5.99)	(1.84)	(4.60)	(6.74)	(1.19)	(1.37)
MKT	0.22	0.63	0.55	0.51	1.42	0.88
	(1.20)	(3.41)	(3.02)	(2.78)	(1.14)	(0.77)
SMB	0.21	0.34	0.33	0.32	0.81	0.11
	(1.45)	(2.92)	(2.57)	(2.44)	(1.16)	(0.21)
OP	0.18	0.36	0.59	0.55	1.22	1.15
	(1.10)	(3.59)	(4.91)	(4.93)	(2.94)	(3.30)
IA	-0.15	0.39	0.31	0.33	0.61	0.44
	(-1.18)	(3.69)	(4.57)	(4.62)	(1.84)	(1.52)
EG	0.05	0.60	0.74	0.71	1.04	1.34
	(0.60)	(8.70)	(10.36)	(10.27)	(1.76)	(2.23)

Table 2.6: Test of Asset Pricing Model: IPCA model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the betas using IPCA approach. See Table 2.6 for details.

	1factor	2factor	3factor	4factor	5factor
beta1	0.862***	-0.1828	0.000254	0.105***	0.100*
	(3.54)	(-0.16)	(0.00)	(3.96)	(2.50)
beta2		-0.6683	0.137	-0.0154	0.00271
		(-0.51)	(1.13)	(-0.59)	(0.10)
beta3			0.392*	0.0153	0.0279*
			(2.20)	(0.81)	(2.15)
beta4				0.0451**	0.00265
				(2.89)	(0.13)
beta5					0.0618***
					(7.11)
cons	-0.00584	0.0265	0.0112	0.0136*	0.0140**
	(-1.41)	(0.61)	(1.20)	(2.40)	(2.75)

Table 2.7: Cross-sectional regression of excess returns on anomalies

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock excess returns on a constant and the lagged firm attributes, including firm size (SIZE), book-to-market ratio (BM), NYSE turnover (NYTURN), Nasdaq Turnover (NASDTURN), return over $t-3$ to $t-2$ months, return over $t-6$ to $t-4$ months, and return over $t-12$ to $t-6$ months. t-ratios, based on Fama and Macbeth(1973)[3] standard errors, are reported below the averages and in parentheses.

	Estimates
Intercept	4.72
	(8.06)
Nasd	0.61
	(1.26)
SIZE	-0.33
	(-7.60)
BM	0.18
	(4.16)
NYTURN	-0.03
	(-0.59)
NASDTURN	-0.92
	(-1.17)
RET2-3	0.91
	(3.86)
RET4-6	0.92
	(4.55)
RET7-12	0.67
	(4.76)
R^2	5.47

Table 2.8: Test of Market Anomalies: CAPM

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes, including firm size (SIZE), book-to-market ratio (BM), NYSE turnover (NYTURN), Nasdaq Turnover (NASDTURN), return over $t-3$ to $t-2$ months, return over $t-6$ to $t-4$ months, and return over $t-12$ to $t-6$ months. t-ratios, based on Fama and Macbeth (1973) [3] standard errors, are reported below the averages and in parentheses. Risk-adjustment is based on CAPM, where betas are estimated using various methods; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	4.35	4.10	2.01	3.87	4.29	4.31
	(7.93)	(7.65)	(7.35)	(7.61)	(7.67)	(7.79)
Nasd	0.42	0.41	0.48	0.52	0.46	0.45
	(0.84)	(0.84)	(1.02)	(1.04)	(0.91)	(0.90)
SIZE	-0.33	-0.30	-0.14	-0.27	-0.34	-0.34
	(-7.62)	(-7.23)	(-6.73)	(-6.97)	(-7.85)	(-7.78)
BM	0.18	0.15	-0.05	0.17	0.15	0.18
	(3.92)	(3.44)	(-1.43)	(4.01)	(3.07)	(3.88)
NYTURN	-0.12	-0.11	-0.11	-0.11	-0.04	-0.05
	(-2.46)	(-2.42)	(-3.67)	(-2.43)	(-0.72)	(-0.82)
NASDTURN	-0.64	-0.59	-0.71	-0.72	-0.64	-0.65
	(-0.71)	(-0.66)	(-0.86)	(-0.86)	(-0.71)	(-0.72)
RET2-3	1.04	1.10	1.55	1.13	1.23	1.13
	(4.75)	(5.22)	(7.87)	(5.58)	(5.56)	(5.16)
RET4-6	0.82	0.95	1.16	0.90	1.05	0.96
	(4.42)	(5.40)	(8.64)	(5.30)	(5.39)	(4.97)
RET7-12	0.75	0.82	0.87	0.86	0.92	0.81
	(5.87)	(6.79)	(8.66)	(7.49)	(6.89)	(6.09)
R^2	6.06	5.85	4.03	5.83	6.65	6.60

Table 2.9: Test of Market Anomalies: Fama and French (1993) [8] Three-factor Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Fama and French (1993) [8] three-factor model, where betas are estimated using various methods; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	3.57	2.67	-0.91	-0.88	3.29	3.30
	(8.75)	(8.16)	(-9.76)	(-7.52)	(7.67)	(7.68)
Nasd	0.45	0.59	0.22	0.14	0.45	0.45
	(0.92)	(1.22)	(2.38)	(1.99)	(0.91)	(0.91)
SIZE	-0.28	-0.20	0.07	0.07	-0.27	-0.27
	(-8.40)	(-7.52)	(8.66)	(7.95)	(-7.62)	(-7.63)
BM	0.13	0.02	-0.01	0.03	0.07	0.07
	(3.58)	(0.55)	(-0.83)	(1.92)	(1.83)	(1.83)
NYTURN	-0.11	-0.12	-0.14	-0.11	-0.05	-0.05
	(-2.54)	(-2.81)	(-9.67)	(-6.80)	(-0.83)	(-0.83)
NASDTURN	-0.61	-1.46	-0.20	-0.14	-0.64	-0.64
	(-0.69)	(-1.45)	(-1.31)	(-0.86)	(-0.72)	(-0.72)
RET2-3	1.03	1.07	1.25	1.14	1.41	1.41
	(5.33)	(5.43)	(12.67)	(10.31)	(6.43)	(6.42)
RET4-6	0.94	1.04	1.12	1.08	1.24	1.24
	(5.68)	(7.09)	(14.64)	(15.21)	(6.43)	(6.43)
RET7-12	0.71	0.88	0.84	0.82	1.09	1.09
	(5.88)	(8.36)	(16.19)	(17.65)	(7.60)	(7.60)
R^2	4.66	4.38	2.78	2.87	5.76	5.76

Table 2.10: Test of Market Anomalies: Carhart (1997) [23] Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Carhart(1997)[23] model, where betas are estimated using various methods; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	3.54	2.32	-1.31	-1.01	3.16	3.17
	(8.88)	(7.49)	(-12.30)	(-10.43)	(7.41)	(7.44)
Nasd	0.44	0.31	-0.00	0.03	0.45	0.45
	(0.92)	(0.74)	(-0.03)	(0.58)	(0.92)	(0.92)
SIZE	-0.27	-0.16	0.11	0.09	-0.24	-0.24
	(-8.33)	(-6.60)	(12.85)	(11.35)	(-6.81)	(-6.84)
BM	0.14	0.01	0.03	0.01	0.06	0.06
	(3.90)	(0.32)	(2.92)	(0.39)	(1.51)	(1.51)
NYTURN	-0.08	-0.13	-0.13	-0.06	-0.04	-0.04
	(-1.95)	(-3.22)	(-12.07)	(-5.30)	(-0.76)	(-0.77)
NASDTURN	-0.55	-0.69	0.17	0.18	-0.64	-0.64
	(-0.64)	(-1.02)	(0.90)	(1.18)	(-0.72)	(-0.72)
RET2-3	1.00	0.98	0.92	0.93	0.49	0.49
	(5.41)	(5.27)	(11.54)	(10.67)	(2.78)	(2.79)
RET4-6	0.91	0.92	0.75	0.88	0.31	0.31
	(5.80)	(6.72)	(12.08)	(16.11)	(2.21)	(2.22)
RET7-12	0.72	0.80	0.56	0.68	0.16	0.17
	(6.29)	(8.85)	(11.92)	(17.90)	(1.53)	(1.54)
R^2	4.55	4.11	3.02	3.67	4.72	4.71

Table 2.11: Test of Market Anomalies: Fama and French (2015) [24] Five-factor Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Fama and French(2015)[24] five-factor model, where betas are estimated using various methods; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	3.98	2.47	-1.64	-0.62	3.12	3.13
	(10.11)	(8.35)	(-26.63)	(-0.82)	(7.63)	(7.66)
Nasd	0.46	0.38	-0.10	2.83	0.46	0.46
	(0.94)	(0.85)	(-1.88)	(1.46)	(0.92)	(0.92)
SIZE	-0.32	-0.19	0.13	0.01	-0.25	-0.25
	(-10.10)	(-8.04)	(23.80)	(0.17)	(-7.43)	(-7.46)
BM	0.08	-0.06	0.00	-0.01	0.00	0.00
	(2.14)	(-1.84)	(0.10)	(-0.06)	(0.01)	(0.02)
NYTURN	-0.09	-0.08	-0.17	-0.07	-0.04	-0.04
	(-2.11)	(-2.06)	(-19.68)	(-0.57)	(-0.69)	(-0.68)
NASDTURN	-0.61	-0.65	0.13	-0.65	-0.65	-0.65
	(-0.69)	(-0.86)	(1.24)	(-0.57)	(-0.73)	(-0.73)
RET2-3	1.00	1.05	1.01	2.92	1.08	1.08
	(5.33)	(5.50)	(16.60)	(2.18)	(5.04)	(5.05)
RET4-6	0.82	0.89	0.87	1.05	0.90	0.90
	(5.08)	(6.47)	(18.33)	(1.38)	(4.86)	(4.87)
RET7-12	0.70	0.78	0.67	0.79	0.76	0.76
	(5.96)	(7.88)	(21.95)	(1.93)	(5.54)	(5.55)
R^2	4.47	4.08	3.49	3.98	5.49	5.48

Table 2.12: Test of Market Anomalies: q^5 Model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on the q^5 model, where betas are estimated using various methods; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	4.44	2.64	-0.86	-0.80	3.29	3.30
	(10.80)	(8.86)	(-12.64)	(-12.00)	(7.79)	(7.79)
Nasd	0.47	0.33	0.21	0.18	0.48	0.48
	(0.93)	(0.87)	(2.54)	(2.84)	(0.93)	(0.93)
SIZE	-0.34	-0.18	0.09	0.08	-0.24	-0.24
	(-10.26)	(-7.71)	(14.69)	(14.02)	(-7.05)	(-7.06)
BM	0.13	-0.01	0.02	-0.00	0.02	0.02
	(3.70)	(-0.36)	(2.38)	(-0.00)	(0.55)	(0.54)
NYTURN	-0.04	-0.08	-0.04	-0.04	-0.05	-0.05
	(-1.04)	(-2.30)	(-4.25)	(-4.51)	(-0.89)	(-0.89)
NASDTURN	-0.59	-0.78	-0.10	-0.13	-0.66	-0.66
	(-0.65)	(-1.22)	(-0.97)	(-2.61)	(-0.71)	(-0.71)
RET2-3	0.99	1.02	0.96	0.94	0.49	0.49
	(5.33)	(5.69)	(12.67)	(13.29)	(2.40)	(2.40)
RET4-6	0.76	0.80	0.83	0.85	0.29	0.29
	(4.60)	(5.72)	(13.71)	(16.21)	(1.65)	(1.66)
RET7-12	0.59	0.72	0.66	0.71	0.10	0.10
	(5.25)	(7.67)	(16.35)	(20.14)	(0.82)	(0.83)
R^2	2.78	2.28	1.57	1.84	3.49	3.48

Table 2.13: Test of Asset Pricing Model: IPCA model

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; *t*-ratios are reported below in parentheses.

	1factor	2factor	3factor	4factor	5factor
Nasd	0.518	0.51	0.511	0.516	0.513
	(1.12)	(1.11)	(1.12)	(1.13)	(1.12)
SIZE	-0.240	-0.25	-0.00557	-0.0338	-0.0175
	(-7.01)	(-9.50)	(-0.33)	(-2.15)	(-1.15)
BM	0.265	0.18	0.171	-0.0747	-0.0362
	(5.45)	(4.22)	(4.31)	(-3.48)	(-1.97)
NYTURN	-0.000626	-0.01	-0.00370	0.00281	0.0141
	(-0.01)	(-0.10)	(-0.07)	(0.05)	(0.26)
NASDTURN	-0.581	-0.58	-0.581	-0.586	-0.574
	(-0.69)	(-0.69)	(-0.71)	(-0.71)	(-0.69)
RET2-3	1.053	1.10	0.260	0.302	0.315
	(4.83)	(5.88)	(1.59)	(1.88)	(1.97)
RET4-6	0.910	0.96	0.120	0.158	0.140
	(4.79)	(6.59)	(1.00)	(1.35)	(1.22)
RET7-12	0.748	0.80	-0.0292	-0.00519	-0.0159
	(5.77)	(7.81)	(-0.33)	(-0.06)	(-0.19)
Intercept	2.741	2.83	0.157	0.317	0.122
	(7.03)	(9.33)	(0.88)	(1.91)	(0.80)
R^2	7.19	5.72	4.51	3.99	3.85

Table 2.14: Test of Market Anomalies: CAPM with Conditional Alpha

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes, including firm size (SIZE), book-to-market ratio (BM), NYSE turnover (NYTURN), Nasdaq Turnover (NASDTURN), return over $t-3$ to $t-2$ months, return over $t-6$ to $t-4$ months, and return over $t-12$ to $t-6$ months. t-ratios, based on Fama and Macbeth(1973) [3] standard errors, are reported below the averages and in parentheses. Risk-adjustment is based on CAPM, where betas are estimated using various methods and alphas are conditional to macroeconomic variables; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	10.92	3.10	11.80	10.06	4.45	4.34
	(15.09)	(7.22)	(11.92)	(11.65)	(8.08)	(7.87)
Nasd	0.18	0.52	3.97	1.39	0.46	0.46
	(0.39)	(1.04)	(5.71)	(3.18)	(0.92)	(0.92)
SIZE	-0.85	-0.22	-0.99	-0.77	-0.35	-0.34
	(-14.53)	(-6.66)	(-13.35)	(-12.25)	(-8.03)	(-7.82)
BM	-0.31	0.07	-0.78	-0.43	0.17	0.19
	(-3.45)	(1.75)	(-5.83)	(-5.18)	(3.39)	(3.88)
NYTURN	-0.56	-0.67	-2.27	-1.19	-0.64	-0.65
	(-0.74)	(-0.81)	(-3.19)	(-2.25)	(-0.71)	(-0.72)
NASDTURN	0.28	-0.13	0.24	0.14	-0.03	-0.03
	(3.93)	(-3.11)	(2.32)	(2.04)	(-0.56)	(-0.61)
RET2-3	-2.25	1.36	1.48	-0.90	1.08	1.06
	(-4.25)	(6.49)	(1.73)	(-1.34)	(4.79)	(4.74)
RET4-6	-1.56	1.10	1.01	-0.48	0.91	0.88
	(-3.81)	(7.12)	(1.54)	(-1.07)	(4.53)	(4.44)
RET7-12	-0.48	0.98	0.09	-0.60	0.78	0.75
	(-1.69)	(9.39)	(0.22)	(-2.20)	(5.53)	(5.49)
R^2	4.21	5.49	2.41	3.49	6.75	6.71

Table 2.15: Test of Market Anomalies: Fama and French(1993) [8] Three-factor Model with Conditional Alpha

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Fama and French(1993) [8] three-factor model, where betas are estimated using various methods and alphas are conditional to macroeconomic variables; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	5.69	5.24	2.52	2.98	3.35	3.69
	(9.88)	(-6.88)	(3.46)	(5.54)	(7.86)	(5.01)
Nasd	-0.23	1.23	-1.73	-0.88	0.46	0.46
	(-0.47)	(0.60)	(-2.64)	(-2.14)	(0.93)	(-0.88)
SIZE	-0.45	-0.12	-0.07	-0.05	-0.27	-0.23
	(-9.52)	(-1.51)	(-1.24)	(-1.01)	(-7.83)	(-6.98)
BM	-0.03	-0.01	1.16	0.99	0.06	0.05
	(-0.36)	(-0.15)	(11.20)	(6.33)	(1.51)	(-1.34)
NYTURN	-0.48	-1.00	-1.39	-1.19	-0.65	-0.62
	(-0.53)	(-0.86)	(-1.24)	(-1.02)	(-0.72)	(-0.59)
NASDTURN	0.27	0.18	0.06	0.02	-0.04	-0.03
	(4.07)	(3.07)	(0.59)	(0.03)	(-0.63)	(-0.55)
RET2-3	-2.24	-1.89	-1.19	-1.11	1.23	1.10
	(-4.40)	(-2.40)	(-1.34)	(-1.01)	(5.56)	(5.19)
RET4-6	-1.31	-1.58	-1.02	-1.13	1.05	1.98
	(-3.23)	(-2.23)	(-1.47)	(-1.90)	(5.45)	(5.22)
RET7-12	-0.30	-0.28	-1.15	-1.87	0.93	0.59
	(-1.11)	(-1.95)	(-2.17)	(-2.79)	(6.75)	(3.26)
R^2	3.53	3.23	2.01	1.69	5.71	5.46

Table 2.16: Test of Market Anomalies: Carhart(1997) [23] Model with Conditional Alpha

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Carhart(1997) [23] model, where betas are estimated using various methods and alphas are conditional to macroeconomic variables; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	5.95	3.16	-2.09	-1.80	3.13	3.14
	(10.88)	(5.88)	(5.10)	(-4.61)	(7.42)	(7.44)
Nasd	-0.49	-0.59	0.50	0.70	0.45	0.45
	(-1.01)	(-1.30)	(2.59)	(2.37)	(0.92)	(0.92)
SIZE	-0.50	-0.69	0.11	0.15	-0.24	-0.24
	(-11.29)	(-13.27)	(5.28)	(4.57)	(-6.80)	(-6.82)
BM	-0.48	0.30	-0.01	-0.01	0.05	0.05
	(-4.99)	(1.99)	(-0.20)	(-0.09)	(1.48)	(1.48)
NYTURN	-0.09	0.07	-1.66	-1.59	-0.64	-0.64
	(-0.11)	(1.10)	(-1.99)	(-1.42)	(-0.72)	(-0.72)
NASDTURN	0.31	0.29	0.08	0.08	-0.04	-0.04
	(4.54)	(2.55)	(1.59)	(1.52)	(-0.75)	(-0.75)
RET2-3	-1.66	-0.90	0.39	0.24	0.49	0.49
	(-3.23)	(-1.17)	(0.69)	(0.60)	(2.78)	(2.79)
RET4-6	-0.86	-0.95	0.90	0.72	0.31	0.31
	(-2.12)	(-2.26)	(2.56)	(2.43)	(2.21)	(2.22)
RET7-12	0.16	-0.90	1.13	0.68	0.16	0.17
	(0.60)	(-2.46)	(3.97)	(3.25)	(1.53)	(1.55)
R^2	3.45	3.54	2.91	2.97	4.72	4.61

Table 2.17: Test of Market Anomalies: Fama and French(2015) [24] Five-factor Model with Conditional Alpha

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on Fama and French(2015) [24] five-factor model, where betas are estimated using various methods and alphas are conditional to macroeconomic variables; see Table 2.1 for the details.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	4.25	6.00	-5.68	-8.33	3.14	5.42
	(7.49)	(8.74)	(-12.17)	(-8.17)	(7.73)	(9.60)
Nasd	-0.09	0.85	0.05	0.03	0.46	0.42
	(-0.18)	(1.03)	(0.12)	(0.10)	(0.92)	(1.18)
SIZE	-0.35	-0.55	0.46	0.65	-0.25	-0.29
	(-7.59)	(-8.26)	(11.55)	(13.16)	(-7.42)	(-8.41)
BM	0.12	0.55	-0.15	0.25	-0.00	-0.00
	(1.35)	(2.01)	(-1.57)	(0.09)	(-0.02)	(-0.01)
NYTURN	-0.51	-1.04	-0.80	-0.76	-0.65	-0.96
	(-0.58)	(-0.95)	(-1.22)	(-1.03)	(-0.73)	(-0.64)
NASDTURN	0.26	-0.20	-0.11	-0.21	-0.04	-0.05
	(3.74)	(-1.21)	(-1.42)	(-1.90)	(-0.67)	(-1.24)
RET2-3	-1.63	1.21	0.92	0.88	1.08	1.59
	(-3.18)	(2.16)	(1.64)	(1.54)	(5.04)	(6.10)
RET4-6	-1.08	0.96	0.68	0.42	0.90	0.85
	(-2.67)	(1.19)	(1.57)	(1.33)	(4.86)	(2.97)
RET7-12	0.11	0.90	0.27	-0.12	0.76	0.41
	(0.41)	(1.07)	(0.83)	(-0.04)	(5.54)	(2.25)
R^2	3.28	4.21	2.56	2.54	5.49	5.28

Table 2.18: Test of Market Anomalies: q^5 Model with Conditional Alpha

This table reports the time-series averages (in percentage) of estimates from cross-sectional regressions of monthly individual stock risk-adjusted returns on a constant and the lagged firm attributes; see Table 2.8 for the description of the attributes. t-ratios are reported below the averages and in parentheses. Risk-adjustment is based on the q^5 model, where betas are estimated using various methods and alphas are conditional to macroeconomic variables; see Table 2.1 for the details.

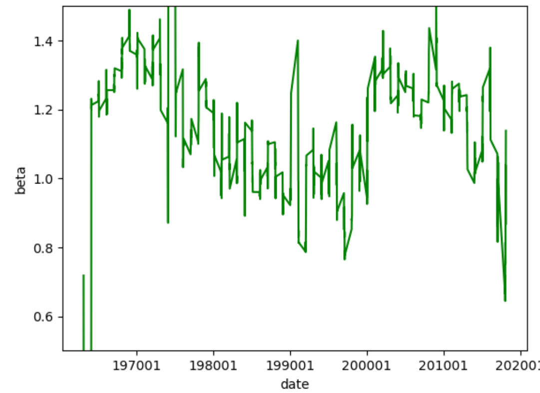
	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Intercept	4.13	4.76	3.51	5.51	4.53	6.55
	(7.04)	(7.99)	(7.25)	(8.66)	(8.70)	(10.18)
Nasd	-0.53	-0.23	-0.34	-0.12	0.48	0.22
	(-0.96)	(-0.40)	(-1.15)	(-1.36)	(0.90)	(0.50)
SIZE	-0.33	-0.75	-0.23	-0.19	-0.35	-0.46
	(-6.96)	(-10.96)	(-6.06)	(-2.14)	(-7.32)	(-8.99)
BM	0.28	0.08	0.02	0.04	0.21	0.10
	(2.88)	(1.20)	(0.32)	(1.84)	(4.64)	(2.31)
NYTURN	-0.17	-0.12	0.39	0.98	-0.69	-0.20
	(-0.22)	(-0.02)	(0.93)	(1.95)	(-0.72)	(-0.67)
NASDTURN	0.25	0.14	-0.02	-0.04	-0.06	-0.06
	(3.85)	(2.58)	(-0.41)	(-0.95)	(-1.09)	(-0.99)
RET2-3	-1.69	-1.99	0.38	0.99	1.14	1.02
	(-2.77)	(-3.29)	(0.61)	(2.23)	(4.88)	(2.13)
RET4-6	-1.05	-0.86	0.44	0.60	0.96	0.90
	(-2.37)	(-1.77)	(0.89)	(0.67)	(4.70)	(3.90)
RET7-12	-0.03	-0.00	0.57	0.05	0.76	0.59
	(-0.09)	(-0.02)	(1.57)	(0.45)	(5.72)	(4.59)
R^2	1.42	1.22	1.01	0.99	5.02	4.99

Table 2.19: Return Forecast: CAPM

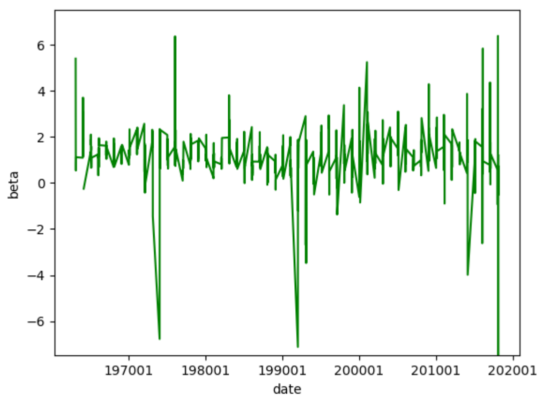
This table reports the cross-sectional averages, median, and standard deviation, of root mean squared forecast errors (in percentage) of individual stock excess return forecasts based on various versions (see Table 2.1 for the details) of CAPM.

	Unscaled	Avramov and Chordia	Kitchen Sink	Individual LASSO	Panel	Panel with LASSO
Mean	9.38	9.58	8.92	8.89	8.86	8.85
Median	8.54	8.59	8.42	8.41	8.28	8.31
Stddev	3.19	3.32	3.01	3.02	3.13	3.08

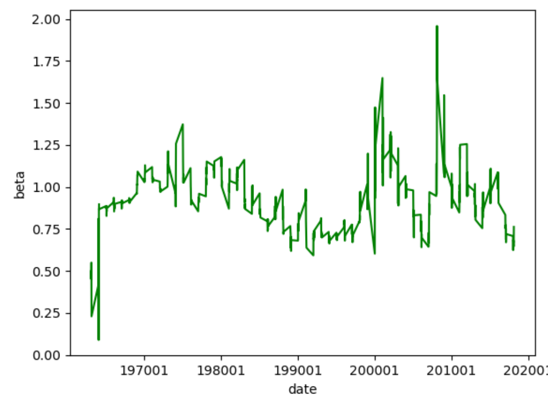
(a) Unscaled



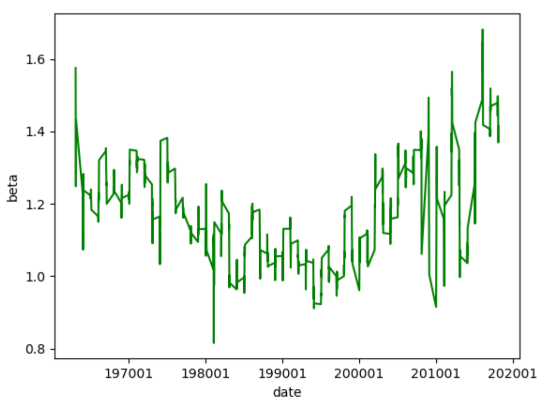
(b) Avramov and Chordia



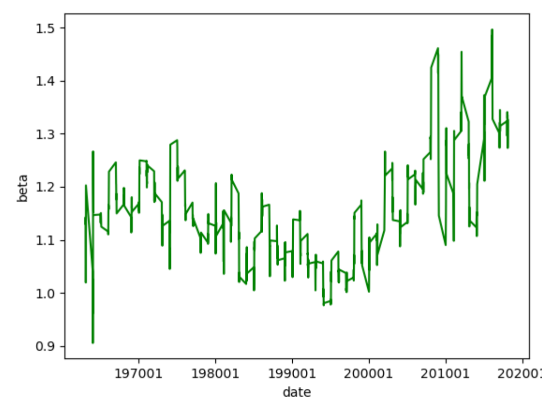
(c) Kitchen sink



(d) Individual LASSO



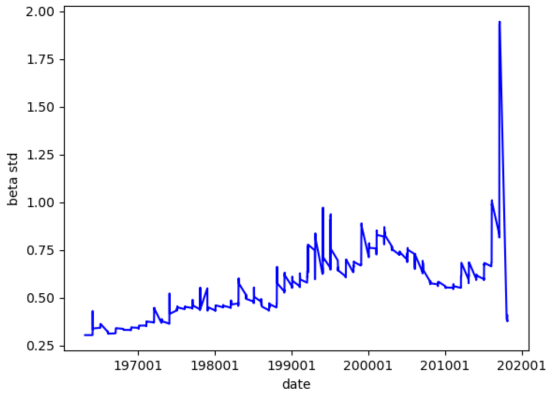
(e) Panel regression



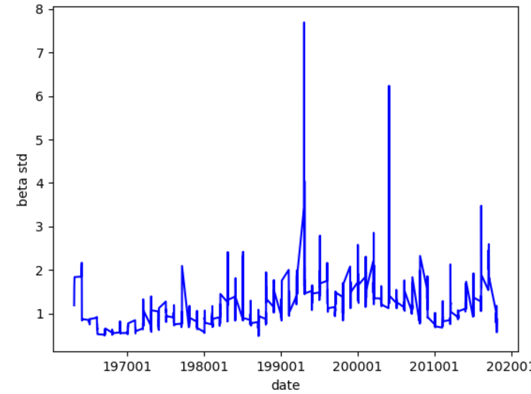
(f) Panel LASSO

Figure 2.1: Cross-sectional Averages of CAPM Betas

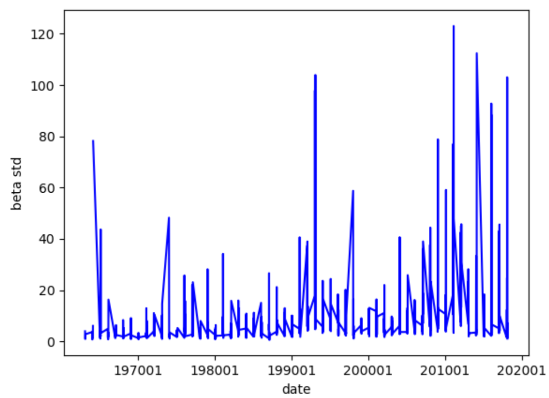
This figure plots the cross-sectional averages of (a) unscaled betas, (b) Avramov and Chordia conditional betas, (c) conditional betas estimated using kitchen sink, (d) conditional betas estimated using individual LASSO, (e) conditional betas estimated using panel regression, and (f) conditional betas estimated using panel LASSO, of the CAPM.



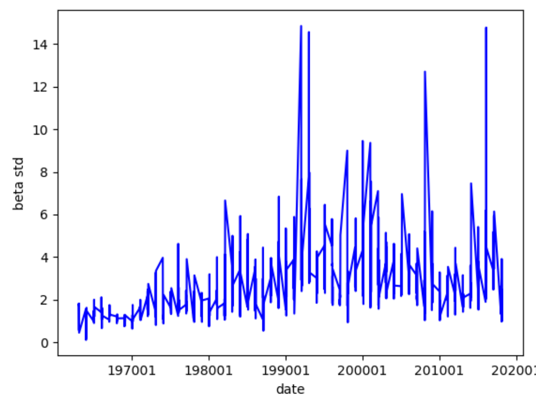
(a) Unscaled



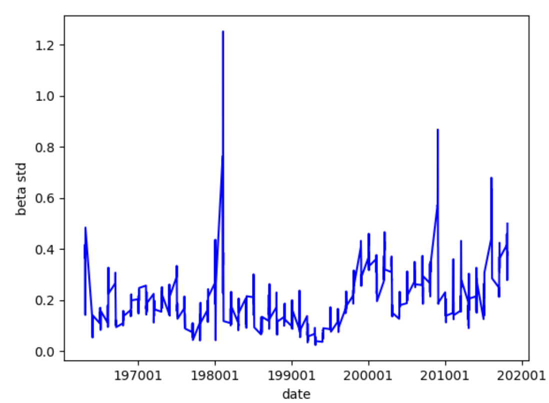
(b) Avramov and Chordia



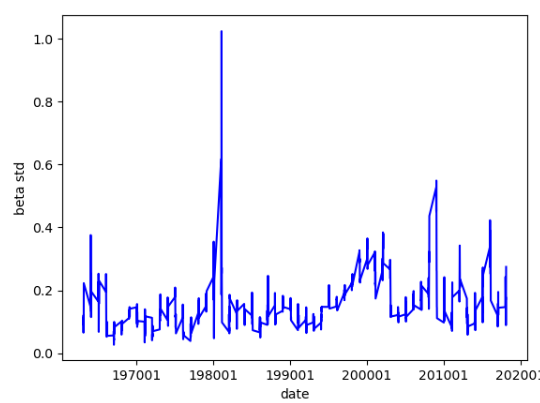
(c) Kitchen sink



(d) Individual LASSO



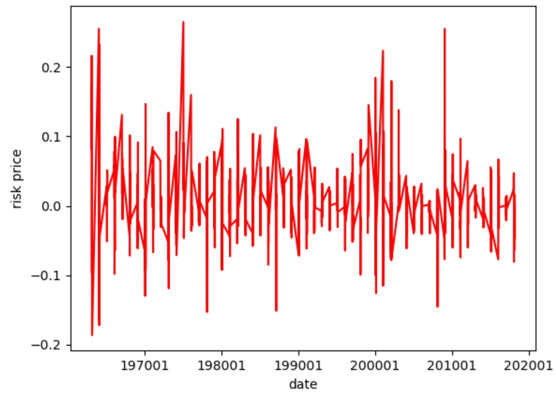
(e) Panel regression



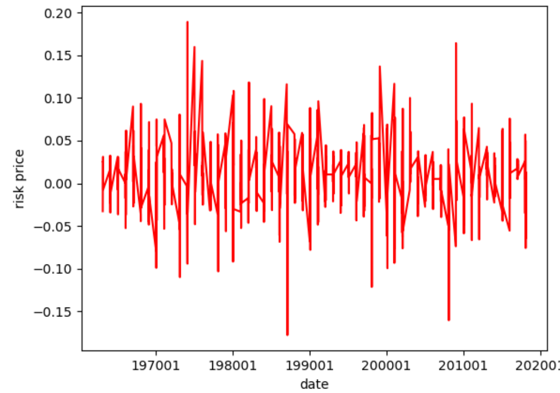
(f) Panel LASSO

Figure 2.2: Cross-sectional Standard Deviations of CAPM Betas

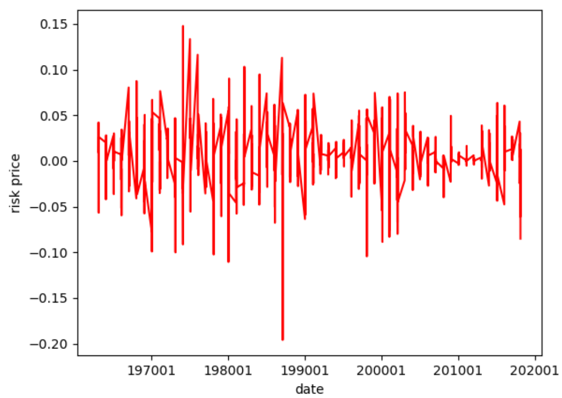
This figure plots the cross-sectional standard deviations of (a) unscaled betas, (b) Avramov and Chordia conditional betas, (c) conditional betas estimated using kitchen sink, (d) conditional betas estimated using individual LASSO, (e) conditional betas estimated using panel regression, and (f) conditional betas estimated using panel LASSO, of the CAPM.



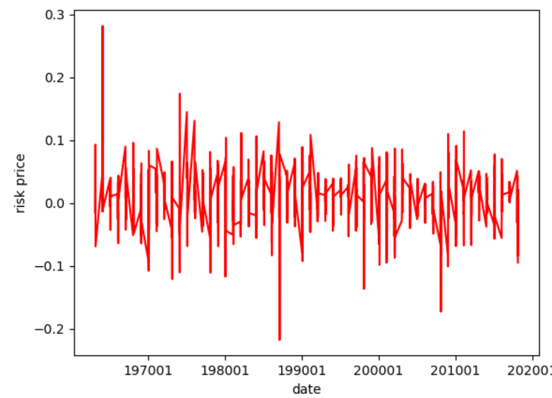
(a) Unscaled



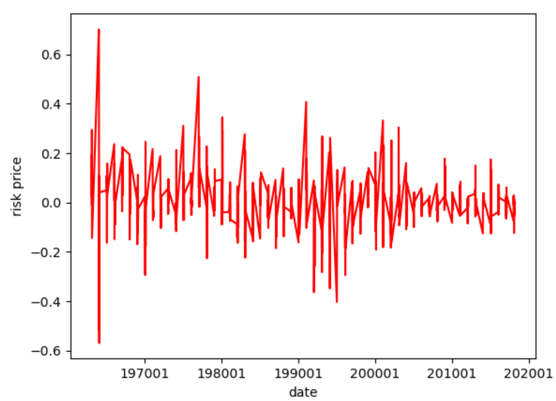
(b) Avramov and Chordia



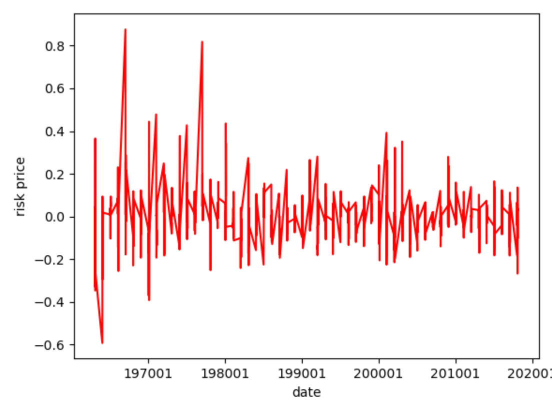
(c) Kitchen sink



(d) Individual LASSO



(e) Panel regression



(f) Panel LASSO

Figure 2.3: Risk Prices for CAPM Betas

This figure plots the risk prices for (a) unscaled beta, (b) Avramov and Chordia conditional beta, (c) conditional beta estimated using kitchen sink, (d) conditional beta estimated using individual LASSO, (e) conditional beta estimated using panel regression, and (f) conditional beta estimated using panel LASSO, of the CAPM.

CHAPTER 3: REIT Conditional betas: a Machine Learning Approach

3.1 Abstract

Portfolio diversification is the primary reason given by financial advisors for allocating a portion of client portfolios into Real Estate Investment Trusts (REITs). Beta is perhaps the most commonly utilized proxy to measure such potential diversification. Accordingly, betas receive substantial attention within REIT studies, industry reports, and financial market websites. The current state of the REIT literature demonstrates the benefits of various time-varying betas versus static betas. The purpose of this study is to move this literature forward in a new direction by demonstrating how incorporating more information and machine learning can be combined to improve beta estimates. More specifically, this study utilizes macroeconomic, firm performance, and REIT-specific characteristics within a straightforward machine learning technique to estimate beta on a monthly basis for nearly three decades. Results from this study demonstrate improved performance relative to time-varying and static betas in regards to out-of-sample return forecasts, and pricing of market anomalies. Going forward, future studies have opportunities to utilize more sophisticated techniques to improve on this inceptive effort.

3.2 Introduction

A recent Chatham Partners survey found that 83% of financial advisors allocate 4% to 12% of their client's investment portfolio in Real Estate Investment Trusts (REITs), with portfolio diversification being the most common attribute referenced for this allocation.[1] With so many investors relying on the diversification aspect of REITs to help provide financial security during retirement, it is critical to reevaluate the diversification benefits when new tools and techniques emerge to ensure this reliance is justified.

Modern portfolio theory provides the conceptual framework that motivates the focus on diversification within an investment portfolio. Within this framework, sufficiently diversified investors mitigate idiosyncratic (diversifiable) risk so that the remaining risk is the systematic risk (non-divesifiable) created by an asset's relationship with the overall stock market. Fur-

[1] A summary of this analysis is available at: https://www.reit.com/news/blog/market-commentary/morningstar-analysis-shows-importance-meaningful-reit-allocations

thermore, modern portfolio theory considers well-diversified investment portfolios to include allocations to all assets in the market basket, including real estate. A recent Morningstar Associates analysis (sponsored by Nareit) found that the optimal portfolio allocation to REITs ranges between 5% and 18% depending on the risk-aversion of the investor.[2] [Cite original allocation suggested by MPT??]

As an intuitive measure of an asset's systematic risk, the market beta ("beta" hereafter) is perhaps the most commonly utilized proxy to measure such potential diversification. Accordingly, betas receive substantial attention within REIT studies, industry reports, and financial market websites. For example, NAREIT (National Association of Real Estate Investment Trusts) has a number of research articles focused on this topic.

The current state of the REIT literature on this topic demonstrates the benefits of various time-varying betas versus static betas. These studies use various GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) models in efforts to examine which type of models produce the best performing (i.e.: lowest mean squared error) time varying estimate of beta.

The purpose of this study is to embrace the findings of this prior literature while demonstrating how incorporating more information and machine learning can be combined to improve time varying estimates of beta. More specifically, this study utilizes macroeconomic, firm performance, and REIT-specific characteristics within a straightforward machine learning technique to estimate beta on a monthly basis for nearly three decades. Results from this study demonstrate improved performance relative to time-varying and static betas in regards to 1) asset pricing models, 2) out-of-sample return forecasts, and 3) pricing of market anomalies. Furthermore, results reveal substantial hetorgeneity within REIT betas according to property type, up and down markets, and pre and post crisis.

The remainder of the article is organized as follows. In the next section, we briefly outline the extant literature. The third section presents our machine learning model, while the fourth section provides a discussion of the data. In the fifth section, we provide results from our empirical tests. The article concludes with final remarks and a summary of key findings.

[2] Available at: https://www.morningstar.com/funds/role-real-estate-investments-portfolio

3.3 LITERATURE REVIEW

Ferson and Harvey(1999) [1] first mentioned the importance of conditional variables in the stock returns.Avramov and Chordia(2006) [2] includes more information in beta estimation in stock market. Najand, Yan and Fitzgerald(2006)[3] proposed the conditional beta in equity REITs. They raised a conditional CAPM methodology in REITs index using ARCH and GARCH model and found that the time-varying model outperformed the stock market with an abnormal return of 2.25%. In this study, we apply an unique set of information variables in REITs beta estimation, including more comprehensive information to estimate the REITs conditional betas, not only firm financial performance and macroeconomic trend in real estate market, but also REITs unique features. Glascock er al.(2018) [4] talks about the asymetry of conditional betas in REITs. Particularly, they first raised 'semi-beta' in REITs industry and compared the betas under different conditions. This application is intuitive in the REITs which performs diversification purpose for investors, we follow this intuition and dig deeper on how the conditional beta performs under different market conditions. Aloy et.al (2021) [5] uses daily REITs indeices to test conditional CAPM model and the out-of-sample return forecast. While different from their study, we use monthly REITs returns instead of REITs index to capture more information for each REITs firms. To better capture the unique feature of each firm, we include a more comprehensive information including REITs information, the accounting performance and macroeconomic variables to estimate the conditional beta. Zhou(2013) [6] compares different time-varying models in REITs using daily REITs index, and found that the time-varying beta shows improvment on return forecast. However, few literature applies REITs firms returns and ignores unique property REITs types. While, we also follow Tibshirani (1996)[7] and applies lasso regression to achieve dimension deductibility, which we believe the first study using Machine Learning techinque to estimate conditional REITs beta. Besides testing the return predictability of conditional model, we follow Stambaugh, Yu, Yuan (2012)[8] to test the popular market anomalies. We also follows Chervachidze and Wheaton(2013)[9] to investigate the performance of conditional beta under financial crisis. We have also put a review matrix table to show our contribution to conditional beta literature in REITs.

Figure 3.1: Literature review

Study	Machine Learning	Anomaly Pricing	Out of Sample	Property Betas	Semibetas	REIT & Macro Variables	Data	Time Period	Sample
Aloy et al. (2021)	No	No	Yes	No	No	No	Indices	2009-2019	Daily
Najand (2020)	No	No	No	No	No	No	Indices	1995-2003	Daily
Glascock et al. (2018)	No	No	No	No	Yes	No	Firms	1992-2014	Monthly
Zhou (2013)	No	No	Yes	No	No	No	Indices	1999-2011	Daily
Current study	Yes	Yes	Yes	Yes	Yes	Yes	Firms	1992-2021	Monthly

Notes: This table summarizes the extant REIT asset pricing literature. Key attributes of the studies are contrasted with those from the current study. "Anomaly Pricing" refers to whether the estimated betas were subsequently tested to price some of the known price anamolies. Simarly, "Out of Sample" refers to whether such forecasts were subsequently estimated. No prior study reports esimates by property type ("Property Betas") and only one prior study reports estimates by up/down markets ("Semibetas"). Current study sample time period is January 1992 - January 2021.

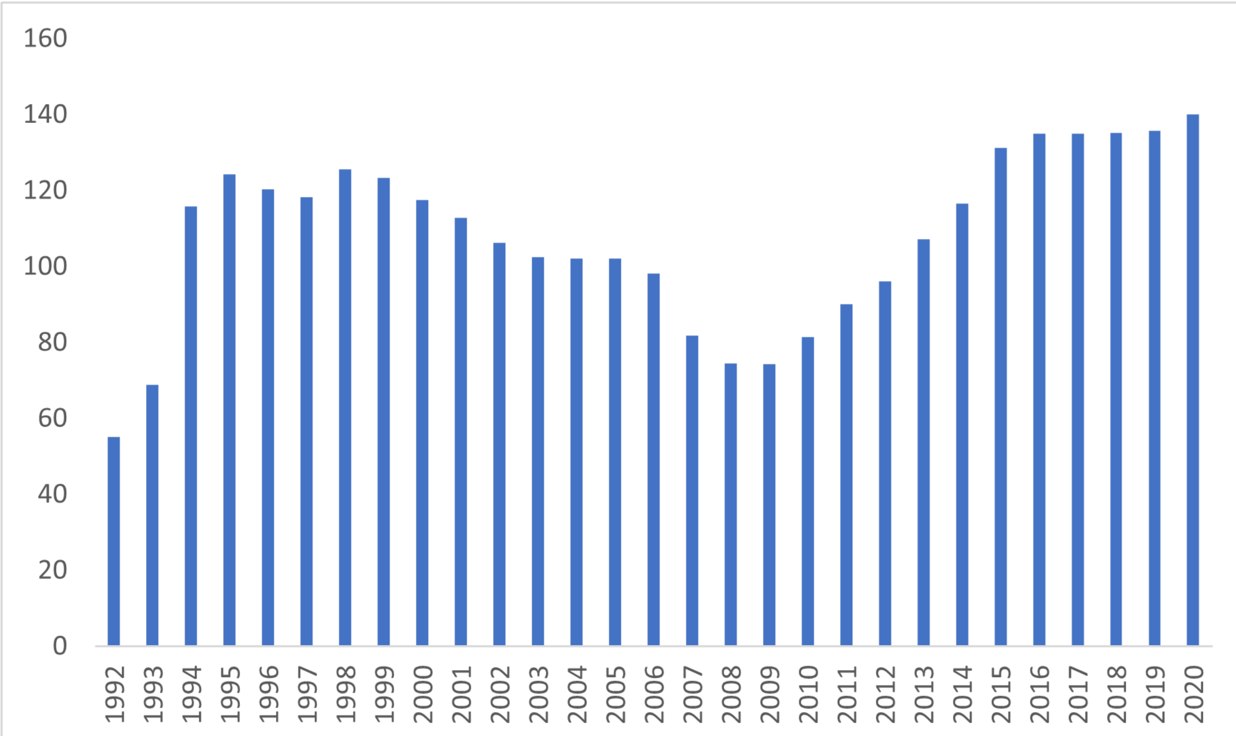
In this study, we also follow Shanken(1992) [10] to implement the asset-pricing model test. Simin(2008) [11]finds that the estimation error of unconditional model always performs better than conditional model, while in our study, we find the out-of-sample predictability following Rapach, Strauss, and Zhou(2010) [12]does a better job to dominate unconditional CAPM. This study also shows the efficiency of our model on involving more information to model conditional betas.

3.4 DATA

We consider the REITs observations starting from 1992 to 2021. Our sample is considered as a comprehensive sample of observations including the *Modern REIT Era*. The year of 1992 has been regarded as the beginning of the new era of REIT, since the REITs has been grown from a small group of companies to more developed publicly - traded equity REITs, including different property types of REITs as well. The initial list of REITs are from S&P Global. This sample includes all of US domestic firms, which have elected REITs' status, property types, locations and essential information.

The REITs returns are from CRSP and the firm characteristics from Compustat. The market excess return is the difference between the S&P500 index and the 3-month Treasury

Figure 3.2: REITs sample by year



Notes: The x-axis is calendar years and the y-axis is the count of REITs in the final sample each year. The sample is an unbalanced panel, as some firms enter and exit the sample in diffent years.

bill rate. The REITs unique characteristics include Property types, Locations(Northeast, Northwest, Southeast, and Southwest), Self-managed, Self-advised, and Triple net leased. Most of the dropped firms were lost due to financial data availablity in Compustat.The REITs macroeconomic variables used in our research are defined as follows.

1. Risk-free rate: 3-month treasury bill rate

2. Corporate risk premium: the spread between the Moody's AAA yield and the 10 year Treasury bond yield.

3. Debt availability: the annual growth rate in the total debt issued divided by the GDP at the same year.

The REITs characteristics are constructed as below. We provide several variables to depict REITs purposes.

1. Self-management: used as a dummy variable, is equal to 1 if the REITs is used for

self-management purpose.

2. Self-advised: used as a dummy variable, is equal to 1 if the REITs is used for self-advised purpose.

3. Triple Net Leased: used as a dummy variable, is equal to 1 if the REITs is Triple Net Leased.

4. Ownership by OP voting: used as a dummy variable, is equal to 1 if the REITs' ownership is decided by OP voting.

5. Operating partnership: used as a dummy variable, is equal to 1 if the property owners allows the property for exchange for operating partnership.

6. The rest of the REITs characteristics are defined as dummy variables if the property types are retail, industrial, office, and health care.

Table 3.1
Summary statistics

Table 1.1 Summary statistics.

Variable	N	Mean	Std. Dev.	Min	Max	p-value
Panel A: Macro (monthly)						
Corporate Risk Premium	349	0.014	0.004	0.006	0.027	<0.0001
Risk Free Rate	349	0.018	0.020	0.001	0.062	<0.0001
Debt to GDP	349	0.776	0.209	0.54	1.359	<0.0001
Panel B: Firm (firm-month)						
Stock Return	34,616	0.011	0.085	-0.724	1.80	<0.0001
Book to Market	34,616	0.672	0.583	-8.471	16.755	<0.0001
Size	25,576	7.094	1.443	2.256	11.263	<0.0001
Momentum	23,461	0.111	0.248	-2.247	2.746	<0.0001
Asset Growth	25,508	25.67	329	-3,792	22,397	<0.0001
Lottery	23,896	0.131	0.104	0.016	1.80	<0.0001
Skewness	23,896	-0.052	0.793	-3.192	3.17	<0.0001
Panel C: REIT (firm-month)						
Self-Managed	34,616	0.831	0.375	0	1	<0.0001
Self-Advised	34,616	0.953	0.211	0	1	<0.0001
Triple Net Leased	34,616	0.233	0.422	0	1	<0.0001
Multifamily	34,616	0.156	0.363	0	1	<0.0001
Retail	34,616	0.214	0.410	0	1	<0.0001
Industrial	34,616	0.088	0.283	0	1	<0.0001
Health Care	34,616	0.102	0.302	0	1	<0.0001
Office	34,616	0.167	0.373	0	1	<0.0001
Other	34,616	0.178	0.356	0	1	<0.0001
Geographic Region	yes					

Notes: This table provides descriptive statistics for the variables utilized in this study. The panels organize the variables into three main categories of information. Macro variables are monthly. Firm and REIT variables are at the firm-month level. Six geographic regions (dummies) are utilized. The reported p-values are based on the results of a t-test of zero means.

3.5 EMPIRICAL METHODOLOGY

Conditional betas construction

In this study, we apply the CAPM to construct the conditional betas. We estimate the following conditional CAPM

$$R_{i,t+1} - R_{f,t} = a_i + \beta_{i,t} MKT_{t+1} + \epsilon_{i,t}, t = 1, \dots, T, \tag{3.1}$$

whereR_i is the return for REITs i, R_f is the risk free rate, MKT is the market excess returns, and T is the number of observations.

Estimation of conditional betas

1. Lasso regressions, propose

$$\min_{a_i, b_i} \sum_{t=1}^{T} (R_{i,t+1} - R_{f,t} - a_i - \beta_{i,t} MKT_{t+1})^2 \tag{3.2}$$

, where the $\beta_{i,t}$ is a linear function in macroeconomic variables, firm accounting performance, REITs unique features, and their interaction terms. And thus the beta of a REITs firm can be shown as

$$\beta_{i,t} = (X_{i,t} \otimes Z_t)^{\top} b_i, \tag{3.3}$$

where $X_{i,t}$ includes firms' financial performance and REITs features, Z_t refers to the macroeconomic variables affecting the REITs industry.

2. Panel Lasso regression

We assume that the b_i is the same for every firm for each month. And we run lasso regression with the penalty parameters at each month for every firm.

Mean Squared Forecast Error(MSFE)

Following Simin (2008)[11] we evaluate the out-of-sample predictability using the mean of the squared forecast error(MSFE). More precisely, the quadratic function of the difference between the expected returns with the true returns.

$$MSFE(\hat{r_t}) = E[(r_t - \hat{r_t})^2]$$

where the $\hat{r_t}$ is the forecast of the return r_t at time t, thus the mean of the squared forecast error (MSFE) can provide an intuitive measure of accuracy of the out-of-sample predictability. The basic idea of Lasso regression is to introduce a little bias so that the variance can be substantially reduced, which leads to a lower overall MSFE.

3.6 Time varying beta

3.6.1 Semi-betas

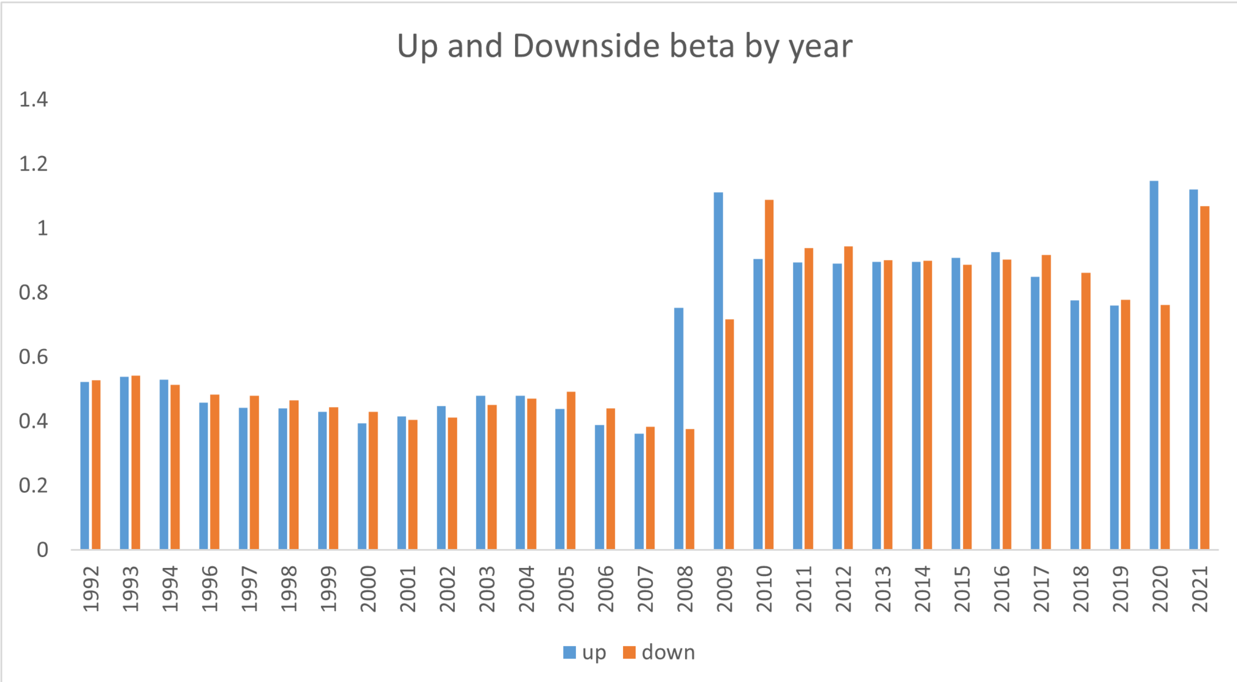
Since the year of 1993, which has been widely recognized that, it started the new era of REITs, which means that the small group of REITs has become a large, publicly-traded market. More and more investors have recognized a great investment opportunity in the market. According to a Morningstar survey, an increasing number of investors have started to allocate more to the REITs. To achieve better diversification purpose, REITs have been more and more popular. As an alternative investment, REITs always play a role as defensive investment, which could help to diversification for hedge fund managers, since sometimes even move to the opposite way of the market overall.

Glascock (2018)[4] discussed the importance of evaluating conditional betas of REITs under different market conditions. They measure the performance under different market conditions of the REITs. The comovement of REITs with the overall financial market are then supposed to be addressed, specially for both hedge fund manager and for academia. It is thus intuitively interesting to explore the comovement with market, i.e., the beta of REITs. Especially, the time-varying beta with changing market conditions. We model the beta of REITs under a time-varying basis, especially, we take a further look at the conditional beta and observe the defensive role of REITs. We then implement a sub-sample analysis, seeing the market excess returns larger or smaller than 0. Particularly, we define the Up- and Down- market on the market index performance. Using this setup, we investigate how the conditional beta changes with different market conditions change, so that we can explore the trend of diversification of REITs and thus provide an intuitive picture for the whole industry.

The conditional beta trend results are shown below. The left figure dipicts the conditional beta trend when the market excess return is larger than 0, while the right figure shows the conditional beta when the market excess return is lower than 0. We can find that among the whole sample period, the REITs play a role of defensive investment, that the time-varying beta is always below or lower than 1. These two plots also verify the conclusion that the REITs can always have diversification effect on the stock portfolio performance no matter how the market is good or bad. When the stock market goes above 0, the conditional beta has an average of 0.80. From the New Era of REITs to the 2008 financial crisis, the beta level is near 0, and even negative before the 2008 Financial Crisis, which proves that the defensive role in the stock market. The diversification function of REITs is still there even in the great crisis, the beta is still lower than 1. In other words, when the market excess returns are larger than 0 even when the market portfolio overall stucks in the crisis, the REITs have played the defensive role in the stock portfolio all the time. The REITs can still well diversify the risk in the 2020 Pandemic, the conditional beta goes up but still much lower than 1. When the market risk premium is lower than 0, the pattern of great diversification exhibits all over the time. Except for the 2008 Financial crisis higher than 1, the conditional beta is much lower than 1 in the rest of the sample period. The average conditional beta level as of 0.81 shows that the strong diversification function of REITs investment. Overall, we can find that no matter how the market portfolio performs, the defensive role of REITs is always there. And even in the down market, the REITs can well diversify the risk within a stock portfolio. The empirical results support that applying the Machine-learning technique can accurately capture the comovement of REITs with market portfolio.

The graph demonstrates the beta in up- and downside market conditions. From the comparison, we can clearly find that two different trends have been shown in the Pre- and Post- crisis, the conditional beta increased larger after the 2008-2009 financial crisis. Before the financial crisis, the down-market beta is larger than the up-market beta. The REITs could play the role of diversification.

Figure 3.3: Conditional Semi-betas



Notes: The x-axis is years and the the y-axis is magnitude of the condtional betas estimated by the fully specified LASSO model using macroeconomic variables, firm financial performance, and REITs attributes. Up and down markets are defined as if the monthly excess return is larger than 0 or not.

Table 3.2

Conditional beta estimates from LASSO models. This table reports various beta estimates from the LASSO model using all three categories of variables (Macro+REIT+Firm) as reported in Panel D, Model 3. Panel A reports estimates for different property types. Panel B reports estimates separately for up and down markets and by pre- and post-financial crisis eras. Panel C reports the conditional betas for up and down markets within the pre- and post-financial crisis eras.

Panel A. Property Type Betas							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Multifamily	Retail	Industrial	Health Care	Office	Other
Mean	0.753	0.481	0.634	0.582	0.493	0.701	0.658
Medians	0.498	0.394	0.455	0.571	0.491	0.580	0.755
Std Dev	0.402	0.512	0.398	0.455	0.489	0.344	0.578
N	34,616	5,252	7,172	2,995	2,766	5,276	11,155

Panel B. Semi Betas and Financial Crisis Eras				
	(1)	(2)	(3)	(4)
	Up Markets	Down Markets	Pre-Crisis	Post-Crisis
Mean	0.677	0.655	0.458	0.909
Medians	0.539	0.529	0.451	0.878
Std Dev	0.250	0.232	0.317	0.390
N	15,499	19,117	25,576	9,040

Panel C. Semi Betas within Pre- and Post-Financial Crisis Eras				
	(1)	(2)	(3)	(4)
	Pre-Crisis		Post-Crisis	
	Up Markets	Down Markets	Up Markets	Down Markets
Mean	0.452	0.463	0.917	0.897
Medians	0.443	0.466	0.861	0.901
Std Dev	0.266	0.387	0.412	0.398
N	11,378	14,198	4,121	4,919

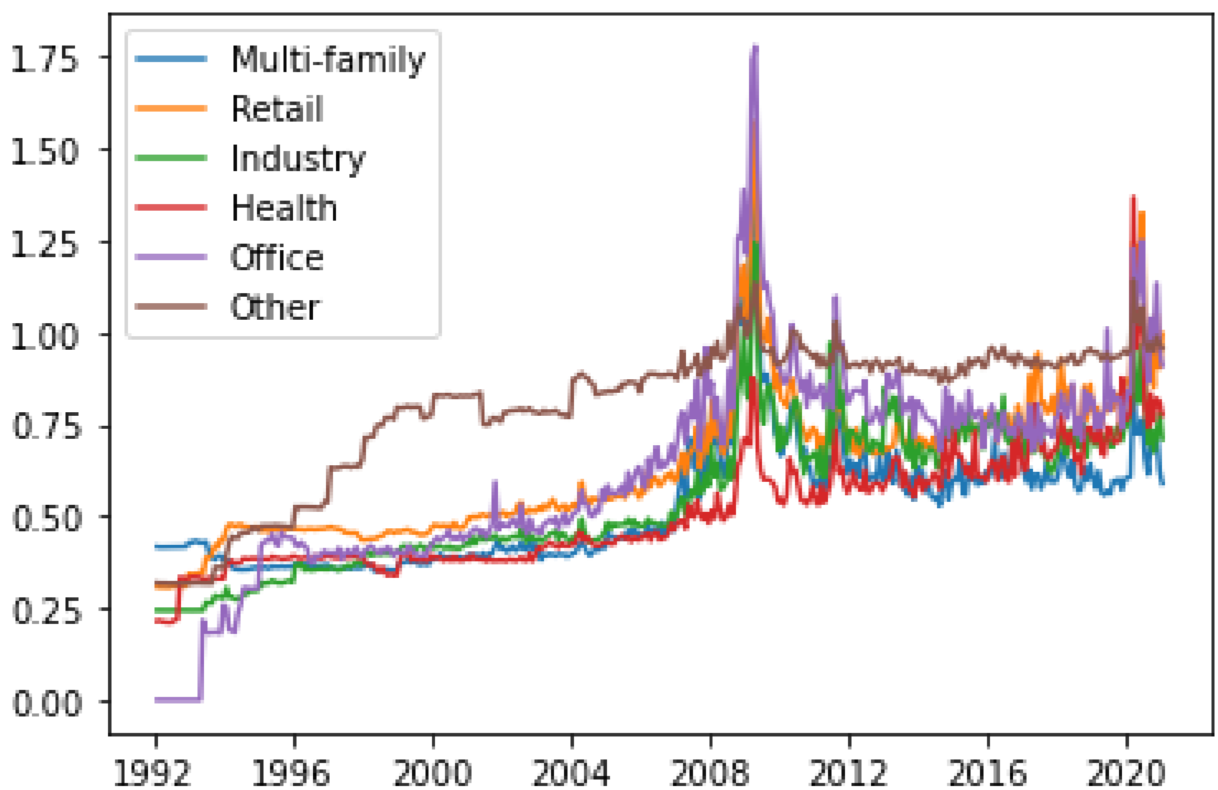
Panel D. Betas by Variable Types			
	(1)	(2)	(3)
	Macro	Macro+REIT	Macro+REIT+Firm
Mean	0.647	0.638	0.655
Medians	0.717	0.664	0.515
Std Dev	0.355	0.387	0.412
N	34,616	34,616	34,511

3.6.2 Conditional betas of different property types

In the REITs investment, it is always interesting to investigate different property types. Because of the special feature of REITs investment, that those firms own or operate different types of real estate, and thus the main cash flow of REITs investment comes from kinds of real estate. The property types of the REITs have important implications to the investors.

From this plot, we can find that different property types REITs show various roles of diversification in a portfolio. Our estimation of conditional betas can depict the trend of the REITs intuitively. There are clearly two different regimes, there are clearly different features happening before and after the crisis, i.e., the two economic recessions. The REITs beta could show that the defensive role of investment clearly. The comovement of the REITs with the market show different degrees accurately show that Especially after the 2008-2009 financial crisis, although all of the property types REITs show an upward trend betas, the health care type of REITs still keep a low level of betas. Furthermore, in the 2020 Covid pandemic, compared with other types of REITs, health care still associate lower betas than other REITs types. On the opposite, the Retail type of REITs suffer the highest upward change in the 2020 pandemic, it achieves an average beta nearly 2.0. These results can also reflect the business and core operations of REITs styles. Essentially, the health care REITs would always behave as the most defensive investment during the Covid pandemic. While, the retail business would be potentially the business sector that suffers the loss worst.

Figure 3.4: Conditional betas by property types



Notes: The x-axis is calendar years and the y-axis is the magnitude of the condtional betas estimated by the fully specified LASSO model from Table 3, Panel D, Model 3. Property type is per the S&P Global Market Intelligence SNL Real Estate Property dataset.

3.7 EMPIRICAL RESULTS

3.7.1 PRICING MARKET ANOMALIES

In this section, we investigate if our conditional beta models could capture the common market anomalies in the REITs market. Following Avramov and Chordia (2006)[2] and Brennan, Chordia, and Subrahmanyam (1998)[13], we run the cross-sectional regression of risk-adjusted returns, ($= R_{i,t+1} - R_{f,t} - \hat{a}_i - \hat{\beta}_{i,t} MKT_{t+1}$) on market anomalies. Intuitively, the risk adjusted returns are the portion which cannot be explained by the asset pricing models. From the results shown in Table 1.3, we can find that the conditional beta models can capture the size and return momentum anomalies. Furthermore, the conditional beta models with Lasso regression could achieve a lower Adjusted R-square compared with unconditional model, and could price anomalies more effectively.

3.7.2 RETURN PREDICTABILITY

In this section, we investigate the predictability using time-varying model with Machine Learning techniques. Specifically, the out-of-sample REITs returns predictability. We compare the predictability of the conditional model with the traditional CAPM model, i.e., the unconditional model. We also compare with the rolling window of historical returns which has been used as a benchmark in both academic and in the industry. Furthermore, in the REITs literature, Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) has been widely used. We also evaluate the time-varying model performance with GARCH model. Applying the data from 1992 to 2021, we find that the conditional models using Lasso regression can perform better than the other models.

Following Simin (2008)[11] we evaluate the out-of-sample predictability using the mean of the squared forecast error(MSFE). More precisely, the quadratic function of the difference between the expected returns with the true returns.

$$MSFE(\hat{r_t}) = E[(r_t - \hat{r_t})^2]$$

where the $\hat{r_t}$ is the forecast of the return r_t at time t, thus the mean of the squared forecast error (MSFE) can provide an intuitive measure of accuracy of the out-of-sample predictability.

Our empirical results have shown that our conditional models using Lasso regression have shown better out-of-sample predictability. We choose the rolling window basis as 60 months, particularly, the estimation window is fixed at 24 months. Due to the limited observations, we choose this window size to prevent losing more observations. Extensive research has shown the importance of stock returns forecast, Rapach, Strauss, and Zhou[12] illustrated the dramatic difference in the performance of in-sample and out-of-sample analysis, while very little attention has been paid to the role of REITs returns forecast. This study therefore intends to assess the Machine-learning application on the REITs out-of-sample predictability. To apply the Lasso regression, we explore different sets of information on estimating the time-varying betas, specifically, the macroeconomic conditions in the REITs market, the REITs unique characteristics (including the property types, locations, and the business features), and the financial performance of the equity REITs. This comprehensive dataset can not only capture the movement in the REITs market, but also the unique features of equity REITs. The performance of predictability has been shown in tables below, the columns represent the Mean of Forecast Error using our time-varying model, the rolling historical average returns, the traditional unconditional CAPM, and the GARCH model respectively.

Figure 3.5: In sample mean squared forecast error

Panel A. In-sample squared errors

	(1) LASSO Macro	(2) LASSO Macro+REIT	(3) LASSO Macro+REIT+Firm	(4) Historical Average	(5) Unconditional	(6) GARCH
MSE	0.007592	0.007597	0.007601	0.008896	0.008896	0.007899
Median	0.001608	0.001608	0.001611	0.002856	0.002425	0.001901
Std Dev	0.044808	0.044653	0.044708	0.058614	0.045028	0.044452

Notes: This table reports the distribution of squared errors using different models. Panel A presents results from in-sample estimates, while Panel B reports results from out-of-sample forecasts. Mean squared error (MSE) and mean squared forecast error (MSFE) are presented first, followed by medians and standard deviations. Models 1, 2, and 3 utilize Lasso regression. Models 4, 5, and 6 utilize historical or rolling average, unconditional CAPM, and GARCH, respectively. The historical average model (Model 4) in Panel A is switched to a rolling average model in Panel B in order to generate the out-of-sample forecast.

We also do the in-sample analysis to compare the difference between conditional beta using Lasso regression with the actual REITs returns.We apply the whole sample observation to estimate the conditional betas and the expected returns. We then apply the MSFE as performance metrics to compare our time-varying beta with the REITs returns. The results are shown in Figure 1.6, and we can find that the conditional beta still perform better than the unscaled model, i.e., a lower MSFE.

Figure 3.6: Out of sample mean squared forecast errors.

Panel B. Out-of-sample forecast squared errors

	(1) LASSO Macro	(2) LASSO Macro+REIT	(3) LASSO Macro+REIT+Firm	(4) Rolling Average	(5) Unconditional	(6) GARCH
MSFE	0.006006	0.004708	0.004683	0.007889	0.055318	0.007806
Median	0.001536	0.001323	0.001317	0.001900	0.003235	0.001773
Std Dev	0.035875	0.028856	0.028479	0.044430	0.650399	0.043837

Notes: This table reports the distribution of squared errors using different models. Panel A presents results from in-sample estimates, while Panel B reports results from out-of-sample forecasts. Mean squared error (MSE) and mean squared forecast error (MSFE) are presented first, followed by medians and standard deviations. Models 1, 2, and 3 utilize Lasso regression. Models 4, 5, and 6 utilize historical or rolling average, unconditional CAPM, and GARCH, respectively. The historical average model (Model 4) in Panel A is switched to a rolling average model in Panel B in order to generate the out-of-sample forecast.

Column 1 shows that conditional models using only the most relevant macroeconomic conditions in REITs. Conditional on REITs macroeconomic conditions, the time-varying model using Lasso regression shows a lowest out-of-sample forecast error. The average of the MSFE is 0.006, which represents a much stronger predictability than the unconditional CAPM. In most of the recent REITs research,the GARCH model has always been selected as the model to capture the REITs returns movement, which has shown a MSFE of 0.008 approximately.We also report the median and standard deviation of the MSFE, all of the statistics have supported that the better predictability of time-varying model compared with GARCH model. We report the rolling-window historical average as another benchmark, i.e, the past 60 months average returns. This historical average, which has been widely used in both academic and for practitioners, has a average MSFE of 0.008. Through the comparison, the time-varying conditional model has shown better predictability among these four models. We then apply a more comprehensive information variables including macroeconomic condition variables and REITs unique characteristics. The return prediction results have been shown in Table 2. When we apply more conditional variables in our time-varying models, we find that the predictability has improved, i.e., the MSFE decreases even, reaches to 0.005. This result has advanced that our conditional model does show better prediction when we combine different perspectives of the REITs returns. We also involve REITs financial performance in our conditional model construction. The prediction results have been shown in Table 3. When we include firms' book to market ratio and liquidity in the conditional model, a slight decrease in the forecast error can be found. That, the dynamic Lasso regression achieves a MSFE of 0.004, which is better than the other models.

Combining different conditional variables in the time-varying models, we consistently find that the out-of-sample prediction is improved by using Lasso regression. The results have shown that the advantage of the time-varying models in the perspective of return predictability. It also supports that the comprehensive set of information could help capture the variations in the REITs returns. Compared with other models, the time-varying conditional models applying Lasso regression show improvement on the forecast accuracy. Within the set of conditional models, we find that more information included, less prediction error. The MSFE estimation demonstrate that the time-varying model can effectively improve the out-of-sample performance. Conclusively, the Lasso regression could select the more useful information on the return predictability further.

3.8 CONCLUSIONS

Portfolio diversification is the primary reason given by financial advisors for allocating a portion of client portfolios into REITs. Therefore, it's critical to reevaluate the diversification benefits of REITs as new tools and techniques emerge. Traditionally, beta is commonly used as a proxy to measure the potential diversification of an investment in relation to the overall market.

While machine learning has been applied in a wide variety of real estate studies, this study is the first known effort to incorporate machine learning into the beta estimation process.

Machine learning algorithms can analyze vast amounts of data and identify patterns that may not be apparent through traditional methods. By incorporating more information into the beta estimation process within such models, it becomes possible to enhance the accuracy and robustness of these estimates.

Overall, results from this study demonstrate that the integration of machine learning into the beta estimation process represents a promising avenue for improving diversification analysis and providing more accurate estimates of REIT betas.

In regards to accuracy, the results show improved performance compared to time-varying and static betas in asset pricing models, out-of-sample return forecasts, and the pricing of market anomalies.

In regards to the beta estimates, the study reveals a high degree of heterogeneity in REIT betas based on property types, market conditions, and pre- and post-crisis periods. Significantly, the results reveal that many property types have betas less than .60 (the typical beta quoted for REITs in most publications) and that in recent years REIT betas demonstrate asymmetry in different market condition (where REITs have higher betas in up markets than in down markets).

Going forward, with so many maching learnine techniques available today and with future innovations in algorithms and architectures on the way, future studies have opportunities to utilize more sophisticated machine learning techniques to improve on this inceptive effort.

Table 3.3
Pricing anomalies

Table 1.3. Pricing anamolies.
Dependent variable = adjusted returns

	(1)	(2)
	LASSO	Unconditional CAPM
Book to Market	-0.0103***	-0.00949***
	(-5.445)	(-4.91)
Size	0.000309	0.000309
	(0.431)	(0.605)
RET_sum	-0.00188	-0.00132
	(-0.380)	(-0.521)
AG	-6.791	-8.208
	(-0.116)	(-0.1835)
ROE	0.0100*	0.00969*
	(1.914)	(1.928)
Lottery	0.147***	0.150***
	(6.519)	(6.43)
Shewness	-0.00360***	-0.00320**
	(-2.772)	(-2.678)
Constant	-0.0333***	-0.0345***
	(-6.243)	(-6.124)
N	23,871	23,871
Groups	338	338
R2	0.302	0.360

Notes: This table reports the results of an cross-sectional regression of adjusted return (deemed as the portion of return which cannot be explained by asset pricing models) on a list of seven common pricing anamolies from the asset pricing literature. A lower R^2 implies more accurate pricing by the beta, as the pricing anomalies do not contribute to the fit of the model. Model 1 utilizes monthly conditional beta estimates from the fully specified LASSO model (Table 3, Panel D, Model 3) while Model 2 utilizes monthly unconditional beta estimates from a CAPM.

REFERENCES

[1] W. E. Ferson and C. R. Harvey, "Conditioning variables and the cross section of stock returns," *The Journal of Finance*, vol. 54, no. 4, pp. 1325–1360, 1999.

[2] D. Avramov and T. Chordia, "Asset pricing models and financial market anomalies," *The Review of Financial Studies*, vol. 19, no. 3, pp. 1001–1040, 2006.

[3] M. Najand, C. Yan, and E. Fitzgerald, "The conditional capm and time varying risk premium for equity reits," *Journal of Real Estate Portfolio Management*, vol. 12, no. 2, pp. 167–176, 2006.

[4] J. L. Glascock and R. Lu-Andrews, "The asymmetric conditional beta-return relations of reits," *The Journal of Real Estate Finance and Economics*, vol. 57, no. 2, pp. 231–245, 2018.

[5] M. Aloy, F. Laly, S. Laurent, and C. Lecourt, "Modeling time-varying conditional betas. a comparison of methods with application for reits," in *Recent Econometric Techniques for Macroeconomic and Financial Data*, pp. 229–264, Springer, 2021.

[6] J. Zhou, "Conditional market beta for reits: A comparison of modeling techniques," *Economic Modelling*, vol. 30, pp. 196–204, 2013.

[7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[8] R. F. Stambaugh, J. Yu, and Y. Yuan, "The short of it: Investor sentiment and anomalies," *Journal of Financial Economics*, vol. 104, no. 2, pp. 288–302, 2012.

[9] S. Chervachidze and W. Wheaton, "What determined the great cap rate compression of 2000–2007, and the dramatic reversal during the 2008–2009 financial crisis?," *The Journal of Real Estate Finance and Economics*, vol. 46, no. 2, pp. 208–231, 2013.

[10] J. Shanken, "On the estimation of beta-pricing models," *The review of financial studies*, vol. 5, no. 1, pp. 1–33, 1992.

[11] T. Simin, "The poor predictive performance of asset pricing models," *Journal of Financial and Quantitative Analysis*, vol. 43, no. 2, pp. 355–380, 2008.

[12] D. E. Rapach, J. K. Strauss, and G. Zhou, "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy," *The Review of Financial Studies*, vol. 23, no. 2, pp. 821–862, 2010.

[13] M. J. Brennan, T. Chordia, and A. Subrahmanyam, "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns," *Journal of financial Economics*, vol. 49, no. 3, pp. 345–373, 1998.