

# Three Essays on Asset Pricing

A Thesis Submitted to the University of Oxford  
for the Degree of Doctor of Philosophy

Wentao Li

*Saïd Business School and St Peter's College*  
*University of Oxford*

Trinity 2024

Dedicated to my wife and my parents

# Acknowledgements

I am profoundly grateful to those who have contributed to my doctoral journey and the completion of this thesis. Without their unwavering support, this achievement would not have been possible.

First and foremost, my deepest gratitude goes to my supervisors, Martin Schmalz and Mungo Wilson. Their mentorship, encouragement, and academic insights have been invaluable in shaping me as a researcher. I am eternally indebted to them for their support and guidance. I also want to express special thanks to Jiri Knesl and Bige Kahraman for their tremendous help in refining my thesis, and to Marianne Andries for her firm support in the job market.

In addition, I am very grateful to the academic community at the University of Oxford, particularly to Renée Adams, Thomas Noe, Tim Jenkinson, Joel Shapiro, and İrem Güçeri, whose constructive feedback and guidance have significantly contributed to my academic development. I also want to thank those who offered generous support during my visit to the University of California, Los Angeles, including Ivo Welch, Mark Grinblatt, William Zame, and many other distinguished scholars.

Moreover, I extend my sincere appreciation to my friends and fellow doctoral students in Oxford and Los Angeles, who have not only helped me with my research but also brought joy and balance to my academic journey.

I also want to acknowledge the financial support provided by the Saïd Business School Foundation and St Peter's College, which have enabled me to pursue my doctoral studies with focus and dedication.

Lastly, I want to express my love and gratitude to my wife, Yuqi Gong, and my parents, Caiming Li and Haiyan Wang. Their unconditional love, encouragement, and understanding have been my source of strength throughout this journey. I also extend my thanks to other family members and friends for their unwavering support and belief in me.

# Contents

Abstract	1
I Overview	3
1 Introduction	4
2 Literature Review	15
II Original Research Essays	39
3 Term Structure of Equity Return Volatility	40
4 Short-Duration Equity Return Puzzle	88
5 Beating the Index with ETFs	120
III Remark and Conclusion	179
6 Remark and Conclusion	180

# Contents

<b>Abstract</b>	<b>1</b>
<b>I Overview</b>	<b>3</b>
<b>1 Introduction</b>	<b>4</b>
1.1 Asset Pricing . . . . .	4
1.1.1 Consumption-Based Models . . . . .	5
1.1.2 Cross-Sectional Returns . . . . .	6
1.2 Topics of Interest . . . . .	8
1.2.1 Term Structure of Equity Returns . . . . .	8
1.2.2 Tax-Loss Harvesting with ETFs . . . . .	9
<b>2 Literature Review</b>	<b>15</b>
2.1 Equity Market . . . . .	15
2.1.1 Equity Premium . . . . .	15
2.1.2 Term Structure of Equity Returns . . . . .	16
2.1.3 Volatility . . . . .	18
2.1.4 Trading Volume . . . . .	19
2.1.5 Subjective Belief . . . . .	20
2.2 Mutual Fund . . . . .	21
2.2.1 Fund Flows . . . . .	22
2.2.2 The Growth of Exchange-Traded Funds . . . . .	23
2.2.3 Exchange-Traded Funds and Market Efficiency . . . . .	25
2.3 Capital Gain Taxation . . . . .	26
2.3.1 Market Impact . . . . .	26

2.3.2	Tax-Loss Harvesting . . . . .	27
<b>II</b>	<b>Original Research Essays</b>	<b>39</b>
<b>3</b>	<b>Term Structure of Equity Return Volatility</b>	<b>40</b>
3.1	Introduction . . . . .	41
3.2	Literature Review . . . . .	43
3.2.1	Volatility Feedback Effect . . . . .	44
3.2.2	Term Structure of Equity Returns . . . . .	45
3.3	Volatility Feedback Effect . . . . .	46
3.3.1	Dividend Strip . . . . .	46
3.3.2	Volatility . . . . .	47
3.3.3	Data . . . . .	48
3.3.4	Empirical Results . . . . .	48
3.3.5	Expected and Unexpected Volatility . . . . .	51
3.3.6	Liquidity . . . . .	53
3.4	Cyclicalilty of Equity Return Structure . . . . .	55
3.4.1	Measurement of Equity Return Structure . . . . .	55
3.4.2	Cyclicalilty . . . . .	57
3.4.3	Decomposition of Cyclicalilty . . . . .	61
3.4.4	Good Volatility, Bad Volatility . . . . .	65
3.5	Simulation . . . . .	68
3.5.1	Preferred habit Model . . . . .	69
3.5.2	Long-run Risk Model . . . . .	69
3.5.3	Rare Disaster Model . . . . .	70
3.5.4	Simulation Results . . . . .	71
3.6	Robustness Check . . . . .	73
3.6.1	Sample Length . . . . .	74
3.6.2	Forward Equity Yield . . . . .	75
3.6.3	VIX Square . . . . .	76
3.7	Conclusion . . . . .	78
	Appendix . . . . .	84

A	Standardised Regressions . . . . .	84
B	Details of Simulations . . . . .	85
<b>4</b>	<b>Short-Duration Equity Return Puzzle</b>	<b>88</b>
4.1	Introduction . . . . .	89
4.2	Literature Review . . . . .	90
4.3	Methodology . . . . .	92
4.3.1	Required Rate of Return . . . . .	92
4.3.2	Conditional Sharpe Ratio . . . . .	94
4.4	Empirical Analysis . . . . .	97
4.4.1	Data . . . . .	97
4.4.2	Validity of Expectation . . . . .	97
4.4.3	Main Result . . . . .	101
4.5	Robustness . . . . .	105
4.5.1	Measurement Error . . . . .	105
4.5.2	Transaction Cost . . . . .	107
4.6	Simulation . . . . .	108
4.6.1	Preferred habit Model . . . . .	108
4.6.2	Long-run Risk Model . . . . .	109
4.6.3	Rare Disaster Model . . . . .	110
4.6.4	Term Structure Model . . . . .	111
4.6.5	Simulation Results . . . . .	112
4.6.6	Discussion . . . . .	114
4.7	Conclusion . . . . .	115
<b>5</b>	<b>Beating the Index with ETFs</b>	<b>120</b>
5.1	Introduction . . . . .	121
5.2	Related Literature . . . . .	126
5.2.1	ETF and Mutual Fund . . . . .	126
5.2.2	Capital Gain Tax . . . . .	128
5.2.3	Trading Volume . . . . .	129
5.3	Empirical Analysis . . . . .	130
5.3.1	Variable Construction . . . . .	130

5.3.2	Identify Tax-Loss Harvesting . . . . .	132
5.3.3	Tax-Loss Harvesting and Price Movement . . . . .	136
5.3.4	Heterogeneous Analysis . . . . .	142
5.3.5	Market Incentive . . . . .	144
5.3.6	Robustness Check . . . . .	146
5.4	Tax-Loss Harvesting Model . . . . .	153
5.4.1	Model Setup . . . . .	154
5.4.2	Simulation . . . . .	157
5.4.3	Tax Revenue Loss . . . . .	163
5.5	Conclusion . . . . .	166
Appendix	. . . . .	172
A	ETFs Tracking the Same Index . . . . .	172
B	Correlation Threshold . . . . .	174
C	Additional Predictions . . . . .	176
D	Tax Alpha and Tax Revenue Loss . . . . .	177

### **III Remark and Conclusion 179**

#### **6 Remark and Conclusion 180**

6.1	Summary of Essays . . . . .	180
6.2	Research Implications . . . . .	181
6.2.1	Academic Implications . . . . .	182
6.2.2	Policy Implications . . . . .	183
6.3	Conclusion . . . . .	183

# Abstract

This thesis encompasses three original research studies in asset pricing, accompanied by an introduction, a comprehensive literature review, and concluding remarks. The first two studies delve into the term structure of equity returns, while the third study centres on tax-loss harvesting with ETFs.

The term structure of equity return volatility exhibits temporal variability, affecting the term structure of equity returns through the volatility feedback effect and explaining the cyclical nature of the equity return term structure. By analysing the dividend strip futures, the first study finds that the volatility feedback effects of dividend strips decrease with the horizon. Using realised and implied volatilities as business cycle indicators, the study substantiates the pro-cyclical nature of the term structure of equity returns. The decomposition of cyclical nature shows that the pro-cyclical nature comes from the high relative sensitivity of short-duration volatility. Notably, the predictable cyclical nature presents a novel criterion for testing macro-finance models, ultimately leading to the rejection of the rare disaster model proposed by Gabaix (2012).

The subsequent original research uncovers a puzzling phenomenon termed the "short-duration equity return puzzle"—short-duration dividend strips exhibit high conditional Sharpe ratios during crises, surpassing theoretical upper bounds. Using dividend prices and forecasts, this study calculates the required rate of return and conditional Sharpe ratio for dividend strips from 2002 to 2021. Notably, during crisis periods, the required rate of return for 1-year dividends peaks at 55%, accompanied by a conditional Sharpe ratio exceeding 14, far surpassing theoretical predictions by mainstream macrofinance models. This anomaly persists across various horizons and remains robust to measurement errors and transaction costs.

The third study uncovers a new source of tax efficiency for ETFs—using highly correlated ETFs to harvest capital losses without violating the wash-sale rule. By exploiting the

tax loophole, investors can potentially earn a better return than the index. This research reveals that the tax-loss trading volume of highly correlated ETFs accounts for 20.7% of their total trading volume. Tax-loss harvesting is negatively related to past returns, especially for recent and negative ones. ETFs with high past volatility have higher tax-loss trading volumes, while smaller and less liquid ones have lower tax-loss trading volumes. A parsimonious model is developed to elucidate the relationship between tax-loss harvesting and past price movements. Simulations with the model predict an annual tax revenue loss of 0.52% of assets under management for highly correlated ETFs, equivalent to approximately 25 billion USD in 2021.

# Part I

## Overview

# 1

## Introduction

### 1.1 Asset Pricing

Asset pricing is essential to finance research, even as early as the 19<sup>th</sup> century. *The Journal of Finance*, an early-age finance journal published between 1895 and 1899 in London, has articles about the movement of gold price when considering international arbitrage opportunities ([Jamieson 1897](#)) and the yield of redeemable securities ([Van Oss 1897](#)).

According to the JEL classification<sup>1</sup>, the narrowly defined "asset pricing" is classified into Category G12 "Asset Pricing; Trading Volume; Bond Interest Rates", which "covers studies about issues related to asset pricing and returns that deal with equities, bonds, and other fixed income securities. It also covers theoretical studies about asset pricing based on economic models, which should be cross-classified here and under other appropriate economic categories."

When compared with "corporate finance" and "financial institution", "asset pricing" may refer to a broader research area of finance. Under the JEL classification, the broadly defined "asset pricing" is classified into Category G1 "General Financial Markets", which "covers studies about general issues related to securities markets." The category includes G10 "General", G11 "Portfolio Choice; Investment Decisions", G12 "Asset Pricing; Trading Volume; Bond Interest Rates", G13 "Contingent Pricing; Futures Pricing", G14 "Information and Market Efficiency; Event Studies; Insider Trading", G15 "International Financial Markets", G17 "Financial Forecasting and Simulation", G18 "Government Policy and Regulation", and G19 "Other".

---

<sup>1</sup><https://www.aeaweb.org/jel/guide/jel.php>

Using publication records of the *Journal of Financial Economics (JFE)*, [Schwert \(2021\)](#) shows that the percentage of JEL category G1 used in *JFE* papers was 66% in the 1970s. The percentage dropped to 34% in the 2010s but remained the largest category of papers published in *JFE*.

This thesis mainly focuses on the narrowly defined "asset pricing", but it is also related to other areas under the JEL classification G1. Chapter 1 presents a general introduction to the research areas covered by the thesis, which includes discussions of basic asset pricing concepts and specific topics of interest. Then, Chapter 2 provides a thorough review of the literature that is closely related to the original research chapters. Chapters 3 and 4 are two original research articles about the term structure of equity returns, and Chapter 6 is an article on tax-loss harvesting with ETFs. These three original research chapters do not depend on others and can be viewed individually. Finally, Chapter 6 summarises the thesis and discusses policy implications.

## A. Consumption-Based Models

Consumption-based models constitute an essential framework of asset pricing theories, which links the pricing kernel to the preference. [Rubinstein \(1976\)](#) introduces one of the earliest consumption-based asset pricing models, which prices uncertain income streams while considering rational risk aversion. [Lucas \(1978\)](#) uses a time-separable utility function of a representative agent to derive the pricing kernel, which is usually referred to as the "standard" or "Lucas-type" asset pricing model. [Breedon \(1979\)](#) and [Grossman and Shiller \(1982\)](#) also contribute to early-stage consumption-based models.

Although early-stage consumption-based models endow asset pricing with a micro foundation, they do not align with empirical evidence. For instance, standard models cannot justify the volatility of asset returns, known as the "excess volatility puzzle". ([Grossman and Shiller 1980](#), [Shiller 1982](#)) [Mehra and Prescott \(1985\)](#) find that the equity premium is too high to be justified by the coefficient of risk aversion in the microeconomics literature.

The standard asset pricing model does not distinguish the elasticity of intertemporal substitution (EIS) and the coefficient of relative risk aversion. [Epstein and Zin \(1991\)](#) and [Weil \(1989\)](#) use a recursive utility function to separate the EIS and the coefficient of relative risk aversion, showing that the empirical equity premium can be consistent with consumption-based models. However, the time non-separable model is too complicated to

solve analytically.

Scaled consumption-based models avoid the difficulty. The general idea is that the pricing kernel is state-dependent, relating to consumption growth and possibly other fundamental variables. Motivated by [Cochrane \(1996\)](#) and [Lettau and Ludvigson \(2001\)](#), [Campbell and Cochrane \(1999\)](#) develop a preferred habit model in which the stochastic discount factor depends on relative consumption.

Another way to explain the puzzles is to introduce the long-run cash flow risk in asset pricing models. [Bansal and Yaron \(2004\)](#) include a small but persistent common component in the time-series process of consumption, which raises the required rate of return by a large scale. However, [Epstein et al. \(2014\)](#) find that the long-run risk model predicts a large premium for the early resolution of uncertain consumption processes, known as the "timing premium puzzle". [Andries et al. \(2019\)](#) propose a horizon-dependent risk aversion model to obtain a reasonable timing premium.

[Gabaix \(2008\)](#) argues that the possibility of rare disasters explains many asset pricing puzzles, including the equity premium puzzle and the excess volatility puzzle. The stochastic discount factor can be extremely low during disaster periods. Using the distribution of historical disasters in [Barro and Ursúa \(2008\)](#), [Gabaix \(2012\)](#) shows that the asset pricing puzzles can be solved with a coefficient of risk aversion as low as 4.

## B. Cross-Sectional Returns

In addition to the time series of asset prices, which is the main focus of consumption-based models, the cross-section of security returns is also essential to asset pricing research.

**Efficient Frontier** The efficient frontier proposed by [Markowitz \(1952\)](#) lays the foundation for the modern portfolio theory. Connecting portfolios with the lowest standard deviation for any given returns, the Markowitz frontier is the set of efficient portfolios for risk-averse investors.

After the introduction of the concept of the stochastic discount factor (SDF) by [Beja \(1971\)](#), [Hansen and Jagannathan \(1991\)](#) find the duality of the mean-standard deviation frontier for the intertemporal marginal rate of substitution (IMRS) and the mean-standard deviation frontier for the asset payoff. Within the Markowitz mean-variance framework, there must be a maximum conditional Sharpe ratio for all assets, which is a function of the

SDF.

**Capital Asset Pricing Model** The efficient frontier inspires the Capital Asset Pricing Model (CAPM). (Sharpe 1964, Lintner 1965) According to CAPM, there is only one efficient portfolio if some assumptions apply, including risk-free lending and borrowing. The only efficient portfolio is the market portfolio, and other assets are priced based on their market beta (covariance with the market portfolio divided by variance).

The Sharpe-Lintner CAPM has only one period. To overcome this issue, Merton (1973) develops an Intertemporal CAPM, in which current asset demands are affected by the possibility of uncertain changes in future investment opportunities. The Intertemporal CAPM allows returns higher than the risk-free rate for securities with no systematic risk.

**Factor Models** Empirical research on the stock market questions the validity of CAPM. Fama and French (1992) find that the market capitalisation and the book-to-market ratio explain cross-sectional stock returns. In contrast, the price of the market beta in CAPM is insignificant in Fama and MacBeth (1973) regressions. Based on this observation, Fama and French (1993) develop the first factor model in asset pricing, the Fama-French three-factor model. The three factors are the CAPM market beta, the small-minus-big (SMB) premium, and the value-minus-growth (HML) premium.

There is a stream of literature investigating the underlying mechanism of factor models. For example, Campbell and Vuolteenaho (2004) explain the SMB and HML factors with a two-beta model, which prices cash flow and discount rate news separately. Petkova (2006) argues that innovations of investment opportunity variables subsume the SMB and HML factors. Abarbanell and Bushee (1998) document that a significant portion of the abnormal factor return is generated around subsequent earnings announcements.

The success of factor models inspires further discovery of risk factors. Many factors other than SMB and HML are also found significant in determining cross-sectional returns, including momentum (Carhart 1997), profitability (Novy-Marx 2013), and originality of innovations (Hirshleifer et al. 2018). Fama and French (2015) claim that profitability and investment factors are also priced in the cross-section, in addition to the three factors in Fama and French (1993). Hou et al. (2020) document 452 factors in the literature and argue that the actual number of factors is even higher.

The large number of "risk factors" is concerning because many may result from data

mining. This problem is known as the "factor zoo". Using samples either forward or backwards in time, [Linnainmaa and Roberts \(2018\)](#) find that most factors fail to explain returns out-of-sample. [Hou et al. \(2020\)](#) show that the  $t$ -statistic of 65% of factors cannot clear the single test hurdle. [Harvey et al. \(2016\)](#) argue that the significance threshold should increase with the number of discovered factors.

The "factor zoo" emphasises the importance of the question: "Which factors should be included in the pricing model?" The traditional approach is to conduct a multivariate test on the factors in regressions. ([Gibbons et al. 1989](#)) [Barillas and Shanken \(2017\)](#) argue that a set of factors should be used if they explain not only asset returns but also other factors. [Barillas and Shanken \(2018\)](#) further develop a Bayesian test method to compare different sets of factors. [Kozak et al. \(2018\)](#) advocate using principal components of returns to replace factors as long as some arbitrageurs are present.

## 1.2 Topics of Interest

This thesis focuses on two specific topics in asset pricing—the term structure of equity returns, which will be covered by Chapters 3 and 4, and tax-loss harvesting with ETFs, which will be covered by Chapter 5.

### A. Term Structure of Equity Returns

Equity refers to the residual interest of a firm held by shareholders. In the financial market, it is usually traded as a stock or an index-based security. The equity value can be viewed as a combination of dividends with different durations, provided that there is no bubble in the long run. Dividends are relatively low compared with equity. The average dividend yield of the S&P 500 index between 1991 and 2021 is only 1.95%, so most of the equity value is contained in dividends with long durations.

Traditionally, financial economists focus on topics related to equity itself because stocks and stock indexes are widely traded in the market. Dividends are not well investigated due to data availability. However, the dividend market has grown fast in recent years. The trading volume of one-year S&P 500 dividend strip futures in December 2020 is almost four times as high as that in December 2015.

The term structure of equity returns refers to the relationship between durations and

returns of dividend strips. Using S&P 500 options data, [Van Binsbergen et al. \(2012\)](#) first document the different returns of dividend strips. They find that returns and Sharpe ratios of holding one- and two-year dividend strips are much higher than those of the S&P 500 index. Some other papers also find that the dividend return is decreasing with duration. ([Cassella et al. 2023](#), [Van Binsbergen and Koijen 2017](#)) [Giglio and Kelly \(2018\)](#) discover an unconditional upward-sloping term structure of equity return volatility with variance swap data.

The high return and Sharpe ratio of short-duration dividends are puzzling because, intuitively, short-duration dividends should have lower risks than the stock index. Many companies provide forward guidance on the time and level of future dividends. They tend to avoid deviating from the guidance due to the signalling effect ([Woolridge 1983](#)), which makes the short-duration dividend similar to a debt. Even if companies do not provide forward guidance, dividend announcements can still be predicted with high precision. ([Kalay and Loewenstein 1986](#), [DeAngelo et al. 2000](#))

The term structure of equity return volatility is the relationship between durations and realised volatility of dividend strips, which has not yet been well investigated. Chapter 3 asks the question: *"Does the term structure of equity return volatility explain the variation of equity return term structure?"* To the best of my knowledge, this research is the first article that connects the term structure of equity return volatility and equity returns.

There are issues with the current finding of the term structure of equity returns. Most papers in the literature measure dividend returns with realised returns, which are inherently backwards-looking. Besides, the term structure of dividend strip returns is cyclical ([Van Binsbergen et al. 2013](#), [Gormsen 2021](#)), which makes the average return vary with sample periods ([Bansal et al. 2021](#)). Chapter 4 in this thesis introduces the "short-duration equity return puzzle" by asking the following question: *"Can mainstream asset pricing models explain conditional Sharpe ratios of dividend strips over time?"* The research question avoids the drawbacks because it does not rely on the average value, and the expected Sharpe ratio is forward-looking.

## B. Tax-Loss Harvesting with ETFs

In the US, investors must pay tax for their realised capital gains over a certain threshold. If investors expect to pay the capital gain tax and hold a losing security, they can replace

that security with another to realise a capital loss. This action is called tax-loss harvesting, which reduces their portfolios' total realised capital gains and, consequently, their capital gains tax burden.

However, the Internal Revenue Service (IRS) has implemented a wash-sale rule, which disallows tax deductions for purchasing and selling "substantially identical" securities within 30 days. ([Internal Revenue Service 2022](#)) As a result, the replacement security is hardly the same as the old one, making it challenging for investors to realise losses without making substantial changes to their portfolios.

Exchange-traded funds (ETFs) offer an ideal solution to sidestep the wash-sale rule. According to IRS *Publication 550*, "Ordinarily, stocks or securities of one corporation are not considered substantially identical to stocks or securities of another corporation." ETFs, structured as trusts, limited partnerships, or open-end funds, represent distinct investment entities and should be regarded as different "corporations". Despite this distinction, ETFs can exhibit high correlations. Buying and selling highly correlated ETFs change the portfolio little. The IRS provides no specific guidelines or tests about tax-loss harvesting with ETFs. ([Fischer 2010](#), [Bouchey et al. 2016](#))

Chapter 5 in this thesis is the first to document the empirical evidence of using highly correlated ETFs to harvest capital loss. With this trading strategy, investors can enjoy a better return than the index. Chapter 5 also builds a parsimonious model to explain the connection between tax-loss harvesting and past price movements, which can be used to estimate the tax revenue loss due to tax-loss harvesting with ETFs.

## References

- Abarbanell, J. S. and Bushee, B. J. (1998), 'Abnormal returns to a fundamental analysis strategy', *Accounting Review* pp. 19–45.
- Andries, M., Eisenbach, T. M. and Schmalz, M. C. (2019), 'Horizon-dependent risk aversion and the timing and pricing of uncertainty', *FRB of New York Staff Report* .
- Bansal, R., Miller, S., Song, D. and Yaron, A. (2021), 'The term structure of equity risk premia', *Journal of Financial Economics* .

- Bansal, R. and Yaron, A. (2004), ‘Risks for the long run: A potential resolution of asset pricing puzzles’, *The Journal of Finance* **59**(4), 1481–1509.
- Barillas, F. and Shanken, J. (2017), ‘Which alpha?’, *The Review of Financial Studies* **30**(4), 1316–1338.
- Barillas, F. and Shanken, J. (2018), ‘Comparing asset pricing models’, *The Journal of Finance* **73**(2), 715–754.
- Barro, R. J. and Ursúa, J. F. (2008), Macroeconomic crises since 1870, Technical report, National Bureau of Economic Research.
- Beja, A. (1971), ‘The structure of the cost of capital under uncertainty’, *The Review of Economic Studies* **38**(3), 359–368.
- Bouchev, P., Brunel, J. L. and Li, T. (2016), ‘The role of etfs in active tax management’, *The Journal of Wealth Management* **19**(3), 75–86.
- Breeden, D. T. (1979), ‘An intertemporal asset pricing model with stochastic consumption and investment opportunities’, *Journal of financial Economics* **7**(3), 265–296.
- Campbell, J. Y. and Cochrane, J. H. (1999), ‘By force of habit: A consumption-based explanation of aggregate stock market behavior’, *Journal of political Economy* **107**(2), 205–251.
- Campbell, J. Y. and Vuolteenaho, T. (2004), ‘Bad beta, good beta’, *American Economic Review* **94**(5), 1249–1275.
- Carhart, M. M. (1997), ‘On persistence in mutual fund performance’, *The Journal of finance* **52**(1), 57–82.
- Cassella, S., Golez, B., Gulen, H. and Kelly, P. (2023), ‘Horizon bias and the term structure of equity returns’, *The Review of Financial Studies* **36**(3), 1253–1288.
- Cochrane, J. H. (1996), ‘A cross-sectional test of an investment-based asset pricing model’, *Journal of Political Economy* **104**(3), 572–621.
- DeAngelo, H., DeAngelo, L. and Skinner, D. J. (2000), ‘Special dividends and the evolution of dividend signaling’, *Journal of Financial Economics* **57**(3), 309–354.

- Epstein, L. G., Farhi, E. and Strzalecki, T. (2014), ‘How much would you pay to resolve long-run risk?’, *American Economic Review* **104**(9), 2680–2697.
- Epstein, L. G. and Zin, S. E. (1991), ‘Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis’, *Journal of political Economy* **99**(2), 263–286.
- Fama, E. F. and French, K. R. (1993), ‘Common risk factors in the returns on stocks and bonds’, *Journal of financial economics* **33**(1), 3–56.
- Fama, E. F. and French, K. R. (2015), ‘A five-factor asset pricing model’, *Journal of financial economics* **116**(1), 1–22.
- Fama, E. F. and MacBeth, J. D. (1973), ‘Risk, return, and equilibrium: Empirical tests’, *Journal of political economy* **81**(3), 607–636.
- Fama, E. and French, K. (1992), ‘The cross-section of expected stock returns’, *Journal of Finance* **47**(2), 427–465.
- Fischer, G. M. (2010), ‘New twists on an old plot: investors look to avoid the wash sale rule by harvesting tax losses with exchange-traded funds’, *Wash. UL Rev.* **88**, 229.
- Gabaix, X. (2008), ‘Variable rare disasters: A tractable theory of ten puzzles in macro-finance’, *American Economic Review* **98**(2), 64–67.
- Gabaix, X. (2012), ‘Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance’, *The Quarterly journal of economics* **127**(2), 645–700.
- Gibbons, M. R., Ross, S. A. and Shanken, J. (1989), ‘A test of the efficiency of a given portfolio’, *Econometrica: Journal of the Econometric Society* pp. 1121–1152.
- Giglio, S. and Kelly, B. (2018), ‘Excess volatility: Beyond discount rates’, *The Quarterly Journal of Economics* **133**(1), 71–127.
- Gormsen, N. J. (2021), ‘Time variation of the equity term structure’, *The Journal of Finance* **76**(4), 1959–1999.
- Grossman, S. J. and Shiller, R. J. (1980), ‘The determinants of the variability of stock market prices’.

- Grossman, S. J. and Shiller, R. J. (1982), ‘Consumption correlatedness and risk measurement in economies with non-traded assets and heterogeneous information’, *Journal of financial Economics* **10**(2), 195–210.
- Hansen, L. P. and Jagannathan, R. (1991), ‘Implications of security market data for models of dynamic economies’, *Journal of political economy* **99**(2), 225–262.
- Harvey, C. R., Liu, Y. and Zhu, H. (2016), ‘... and the cross-section of expected returns’, *The Review of Financial Studies* **29**(1), 5–68.
- Hirshleifer, D., Hsu, P.-H. and Li, D. (2018), ‘Innovative originality, profitability, and stock returns’, *The Review of Financial Studies* **31**(7), 2553–2605.
- Hou, K., Xue, C. and Zhang, L. (2020), ‘Replicating anomalies’, *The Review of financial studies* **33**(5), 2019–2133.
- Internal Revenue Service (2022), Publication 550: Investment income and expenses (including capital gains and losses), Technical report, Department of the Treasury.
- Jamieson, G. (1897), ‘Eastern exchanges and gold prices’, *Journal of Finance* **1**(1), 44–51.
- Kalay, A. and Loewenstein, U. (1986), ‘The informational content of the timing of dividend announcements’, *Journal of Financial Economics* **16**(3), 373–388.
- Kozak, S., Nagel, S. and Santosh, S. (2018), ‘Interpreting factor models’, *The Journal of Finance* **73**(3), 1183–1223.
- Lettau, M. and Ludvigson, S. (2001), ‘Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying’, *Journal of political economy* **109**(6), 1238–1287.
- Linnainmaa, J. T. and Roberts, M. R. (2018), ‘The history of the cross-section of stock returns’, *The Review of Financial Studies* **31**(7), 2606–2649.
- Lintner, J. (1965), ‘Security prices, risk, and maximal gains from diversification’, *The journal of finance* **20**(4), 587–615.
- Lucas, R. E. (1978), ‘Asset prices in an exchange economy’, *Econometrica: journal of the Econometric Society* pp. 1429–1445.
- Markowitz, H. M. (1952), ‘Portfolio selection’, *Journal of finance* **7**(1), 71–91.

- Mehra, R. and Prescott, E. C. (1985), ‘The equity premium: A puzzle’, *Journal of monetary Economics* **15**(2), 145–161.
- Merton, R. C. (1973), ‘An intertemporal capital asset pricing model’, *Econometrica: Journal of the Econometric Society* pp. 867–887.
- Novy-Marx, R. (2013), ‘The other side of value: The gross profitability premium’, *Journal of financial economics* **108**(1), 1–28.
- Petkova, R. (2006), ‘Do the fama–french factors proxy for innovations in predictive variables?’, *The Journal of Finance* **61**(2), 581–612.
- Rubinstein, M. (1976), ‘The valuation of uncertain income streams and the pricing of options’, *The Bell Journal of Economics* pp. 407–425.
- Schwert, G. W. (2021), ‘The remarkable growth in financial economics, 1974–2020’.
- Sharpe, W. F. (1964), ‘Capital asset prices: A theory of market equilibrium under conditions of risk’, *The journal of finance* **19**(3), 425–442.
- Shiller, R. J. (1982), Consumption, asset markets and macroeconomic fluctuations, in ‘Carnegie-Rochester Conference Series on Public Policy’, Vol. 17, Elsevier, pp. 203–238.
- Van Binsbergen, J., Brandt, M. and Koijen, R. (2012), ‘On the timing and pricing of dividends’, *American Economic Review* **102**(4), 1596–1618.
- Van Binsbergen, J. H. and Koijen, R. S. (2017), ‘The term structure of returns: Facts and theory’, *Journal of Financial Economics* **124**(1), 1–21.
- Van Binsbergen, J., Hueskes, W., Koijen, R. and Vrugt, E. (2013), ‘Equity yields’, *Journal of Financial Economics* **110**(3), 503–519.
- Van Oss, S. L. (1897), ‘The net yield of redeemable stocks’, *Journal of Finance* **2**(5), 44–47.
- Weil, P. (1989), ‘The equity premium puzzle and the risk-free rate puzzle’, *Journal of monetary economics* **24**(3), 401–421.
- Woolridge, J. R. (1983), ‘Dividend changes and security prices’, *The Journal of Finance* **38**(5), 1607–1615.

## 2

# Literature Review

## 2.1 Equity Market

With its nature of uncertainty and the accessibility to rich data, the equity market is central to asset pricing research. As discussed in Chapter 1, many asset pricing models aim to explain the equity market's stylised features. The discovery of new stylised facts and puzzles, in turn, stimulates the development of asset pricing theories. Famous empirical findings on the equity market include the excess volatility puzzle ([Shiller et al. 1981](#)), the equity premium puzzle ([Mehra and Prescott 1985](#)), the risk-free rate puzzle ([Weil 1989](#)), the return predictability of stocks ([Campbell and Shiller 1988](#)), the characteristics puzzle ([Daniel and Titman 1997](#)), and stock market anomalies attributed to factors ([Fama and French 1993, 2015](#), [Hou et al. 2015](#)).

### A. Equity Premium

The equity premium is the difference between the equity return and the risk-free rate. Since the discovery of the equity premium puzzle, which states that the equity premium is too high to be explained by the Lucas-type "standard" asset pricing model, this topic has inspired many empirical and theoretical works.

Some papers argue that the equity premium is not as puzzling as [Mehra and Prescott \(1985\)](#) present. Using the classical Gordon model with dividend growth, [Jagannathan et al. \(2001\)](#) find that the average equity premium between 1971 and 1999 was only 0.7%, much lower than between 1926 and 1970 (7%). [Dimson et al. \(2006\)](#) also document a smaller equity premium using data from 17 countries. The average global equity premium was

estimated to be 3–3.5% on a geometric mean basis or 4.5–5% on an arithmetic basis. [Fama and French \(2002\)](#) argue that a substantial portion of previously discovered equity premium was attributed to the capital gain, resulting from the declining discount rate. After accounting for the capital gain, the equity premium between 1951 and 2000 was 2.55–4.32%. In response to the challenging evidence against the equity premium puzzle, [Mehra \(2003\)](#) collects the data between 1802 and 2000 for the US market and between 1947 and 1998 for international markets, showing that the equity premium puzzle has been consistent for nearly two centuries.

The equity premium puzzle is a popular test criterion for asset pricing models. [Benartzi and Thaler \(1995\)](#) argue that the prospect theory can explain the equity premium puzzle by combining loss aversion and frequent portfolio revisions. Consumption-based asset pricing models, such as the preferred habit model ([Campbell and Cochrane 1999](#)), the long-run risk model ([Bansal and Yaron 2004](#)), and the rare disaster model ([Gabaix 2012](#)), set the empirical equity premium as a benchmark for parameter tuning. More recently, [Bai and Zhang \(2022\)](#) derive a DSGE model with recursive utility, search frictions, and capital accumulation to explain the equity premium. Their model predicts an equity premium of 4.27%.

## B. Term Structure of Equity Returns

The term structure of equity returns measures the relationship between dividend strip return and duration. The concept is relatively new due to data availability. [Van Binsbergen et al. \(2012\)](#) first show that the unconditional term structure of equity returns is downward sloping, using S&P 500 options data (Long-Term Equity Anticipation Securities, LEAPS) ranging from 1996 to 2009. Proprietary datasets help researchers explore dividend strip returns with longer horizons, hence the term structure. Using private data from investment banks, [Van Binsbergen and Koijen \(2017\)](#) show that short dividend strips have smaller betas but higher returns and Sharpe ratios.

Recent literature pays much attention to the cyclical nature of equity return term structure. [Van Binsbergen et al. \(2013\)](#) decompose forward dividend yields, which are closely related to expected hold-to-maturity dividend strip returns. They show that the term structure of equity returns is upward-sloping in the expansion (pro-cyclical), while the term structure of expected dividend growth is upward-sloping in recession (counter-cyclical). [Bansal et al.](#)

(2021) develop a regime-switching model with Bayesian methods, showing that the term structure of equity returns is pro-cyclical. Golez and Jackwerth (2020) extend the observations of Van Binsbergen et al. (2012) to a more extended period (1995–2019). They find that the 2-year dividend strip outperforms the index in recessions and underperforms in normal periods. Chabi-Yo and Loudis (2020) focus on shorter horizons (less than one year), showing that the short-end term structure is pro-cyclical. On the contrary, Gormsen (2021) argue that the term structure is counter-cyclical with business cycle measured by equity yield. Gonçalves (2021a) develops an ICAPM model to support this finding.

However, most papers in the literature analyse realised returns instead of required returns, so they are naturally backwards-looking. As a result, the term structure of equity returns is not comparable to the term structure of interest rates.<sup>1</sup> To solve this problem, Bansal et al. (2021) construct the expected equity return by estimating a conditional dividend growth rate with historical data. Gormsen (2021) uses predictive regression to estimate the expected return with the realised return. Giglio, Kelly and Kozak (2021) employ a factor model to estimate the expected return without dividend strips data. These papers do not directly observe the term structure of expected returns.

Many papers try to build theoretical models to explain the term structure of equity returns. Based on the long-run risk model proposed by Bansal and Yaron (2004), Ai et al. (2018) introduce a general equilibrium framework to model the dynamics of discount factor and cash flow. They show that the model matches some stylised facts and helps to extend the term structure to a longer period. Andries et al. (2019) introduce risk aversions that decrease with the horizon into the long-run risk model. The model implies that liquidity is critical for the term structure slope. Wu (2020) incorporates the dividend recovery feature into the variable disaster model of Gabaix (2012). The model generates a lot of stylised features.

Equity duration is closely related to the term structure of equity returns, but it focuses on stocks rather than dividend strips. Equity duration literature cares about dividends further into the future because equity duration is usually very long. Weber (2018) discovers that short-duration equities have higher returns than long-duration equities, which risk factors cannot explain. Gonçalves (2021b) finds a similar result and shows that short-duration stocks are exposed to expected return variation risk. Gormsen and Lazarus (2021)

---

<sup>1</sup>See Moench and Siavash (2021) for a review.

provide a duration-based explanation for the premia on major equity factors. Some papers apply the concept of duration in different contexts. [Chen \(2020\)](#) proposes that the effective equity duration accounts for both cash flow and discount rate changes. [Dechow et al. \(2021\)](#) suggest using the implied equity duration to measure the pandemic shutdown risk.

## C. Volatility

Early research on the volatility feedback effect can be traced back to [Pindyck \(1984\)](#). [French et al. \(1987\)](#) first document the effect systemically, finding that the predicted volatility explains the required return positively. Using the QGARCH model, [Campbell and Hentschel \(1992\)](#) find that the volatility feedback effect is asymmetric—high volatility periods have high feedback effects. [Martin \(2017\)](#) derives the lower bound of equity premium with implied volatility calculated from option prices, showing that the lower bound is tight and predicts the realised equity premium. [Chabi-Yo and Loudis \(2020\)](#) further derive the upper bound with high-order moments.

Volatility matters not only in time series but also in cross-sections. [Herskovic et al. \(2016\)](#) find that the common idiosyncratic volatility (CIV) predicts dividend growth and explains cross-sectional equity returns. It can also explain the value factor. [Chabi-Yo et al. \(2021\)](#) derive an extended ICAPM model with equity yield and realised volatility as risk factors. They argue that the volatility factor conveys a negative price and cannot be explained by other prevalent factors. [Martin and Wagner \(2019\)](#) derive risk neutral variance from stock options, showing that high variance is associated with small and value stocks. Moreover, their model explains over 30% of equity returns. [Vasquez \(2017\)](#) shows implied volatility is positively associated with straddle returns in cross-section, which prevalent risk factors cannot explain.

Recent research recognises that the difference between risk-neutral and physically expected volatility, called variance risk premium (VRP), is crucial in explaining asset returns. [Bollerslev et al. \(2009\)](#) build a model based on the long-run risk model with time-varying volatility of volatility (VoV). With options data from 1990 to 2007, they find that high VRP predicts high returns, especially at the quarterly horizon. [Drechsler and Yaron \(2011\)](#) construct VRP as the difference between the squared VIX index and the expected realised volatility, making a similar conclusion. [Dew-Becker et al. \(2017\)](#) use variance swap to estimate the prices of variance risk in different horizons, finding that only very-short-term

variance (up to 2 months) is priced. With option prices, [Andries et al. \(2015\)](#) find that the prices of variance risk are negative and decrease with the horizon in absolute terms. They also show that high volatility is associated with a steep price structure. VRP predicts both return ([Bekaert and Hoerova 2014](#)) and dividend growth ([Bollerslev et al. 2015](#)). [Bekaert et al. \(2019\)](#) argue that time-varying risk aversion can explain changes in VRP. Instead of using VRP, [Cheng \(2019\)](#) considers the VIX premium, which is defined as the difference between the VIX futures index and the expected VIX forecasted by ARMA. The author shows that the VIX premium predicts future returns.

## D. Trading Volume

A long piece of literature focuses on why the trading volume exists. Intuitively, no trade occurs if all agents agree on the same price as new information comes. By allowing agents to be heterogeneous, [Karpoff \(1986\)](#) finds that there can be a normal trading volume with random meetings between buyers and sellers. Besides, the trading volume persists. [Holthausen and Verrecchia \(1990\)](#) classify the trading volume in response to information into two effects—informedness and consensus. They find that high informedness leads to high return volatility and trading volume, while high consensus leads to high volatility and low trading volume. [Easley and O'Hara \(1992\)](#) argue that the trading volume depends on the adverse selection of informed investors. A higher trading volume relates to a faster convergence to the equilibrium. In addition, [Odean \(1998\)](#) and [Scheinkman and Xiong \(2003\)](#) argue that overconfidence results in high trading volumes.

Many papers focus on the connection between security price movement and trading volume. [Karpoff \(1987\)](#) surveys the literature and finds that the trading volume reacts positively to both return and the absolute value of return, raising the asymmetric volume-price hypothesis. [Kim and Verrecchia \(1991\)](#) build a 2-period rational expectation model to investigate the return and volume in response to announcements. They show that price changes proportionally to the importance of information, and volume changes proportionally to the absolute value of return. The preciseness of information is positively associated with both volume and return, while the preciseness of pre-announcement information decreases them. [Wang \(1994\)](#) also proposes a model with asymmetric information, arguing that volume positively correlates to absolute return and dividend. The paper also shows that the trading volume contains no information on price. Borrowing Wang's idea, [Llorente](#)

[et al. \(2002\)](#) show that the asymmetric information model can explain the empirical findings of momentum and volume. They find that hedging trades reverse while speculative ones persist following high-volume days, especially for small and high-spread firms. [Chen et al. \(2001\)](#) conduct empirical research on nine markets, finding that absolute return is positively associated with trading volume. Moreover, return and volume Granger cause each other. Return volatility persists after controlling for the trading volume.

## E. Subjective Belief

Under Lucas's "standard" asset pricing framework, investors are rational, and there is no difference between the model-based expected return and investors' forecast. However, survey-based research shows that people do not always have rational expectations, and subjective beliefs matter. ([Manski 2018](#))

Subjective beliefs are important in asset pricing. Using six different datasets covering the period between 1963 and 2011, [Greenwood and Shleifer \(2014\)](#) discover that the model-based expected return positively predicts the realised return. However, investors' return expectation negatively predicts the realised return. The inconsistency between the model-based expected return and investors' return expectation challenges the rational expectation model. [Manela and Moreira \(2017\)](#) construct a news-based measure of subjective belief, news implied volatility (NVIX), which rises before the occurrence of rare disasters.

Recent research pays attention to the distinguishment between return and cash flow expectations. Using a survey of wealthy retail investors and corresponding administrating data of holdings, [Giglio, Maggiori, Stroebel and Utkus \(2021\)](#) show that cash flow and return expectations are positively correlated, and the return expectation and the expected probability of rare disasters are negatively correlated. [De La O and Myers \(2021\)](#) find that the cash flow growth expectation accounts for 93% of the variation of the S&P500 price-dividend ratio and 63% of the S&P500 price-earning ratio. In contrast, the return expectation has low co-movement with price ratios.

Many recent asset pricing models incorporate subjective belief as an essential factor. [Collin-Dufresne et al. \(2017\)](#) include the belief heterogeneity between young and old groups in the asset pricing model with recursive preferences. They show that a small departure from rational expectation can generate high risk premiums, extrapolative forecasts, and substantial and persistent over- or under-valuation. By deriving a "risk-centric" represen-

tation of the New Keynesian model, [Caballero and Simsek \(2020\)](#) show that average beliefs play a vital role in the valuation of assets and the amplification of the demand recession loop, while heterogeneous beliefs result in speculations during the recovery. In addition, [Nagel and Xu \(2022\)](#) include the fading memory of endowment growth in their asset pricing model. On the contrary, [Wachter and Kahana \(2024\)](#) develop the retrieved-context theory, which argues that remote but relevant memories play an important role in asset prices.

The literature on subjective belief heavily relies on surveys.<sup>2</sup> There are questions about whether surveys are accurate and can represent marginal investors. [Adam et al. \(2021\)](#) collect seven different survey datasets, finding that investors' forecasts are predictably optimistic and unconditionally unbiased. Return forecasts in surveys are significantly higher than the risk-free rate. They argue that survey respondents do not confound between belief and preference. [D'acunto et al. \(2023\)](#) discover that the IQ scores of survey respondents play an important role, even after controlling for education, income, and many other characteristics. The research shows that high-IQ men display 50% lower forecast errors with consistent expectations.

## 2.2 Mutual Fund

Mutual funds have a history of over 200 years. The first mutual fund was started in Holland in 1774, and the first in the US was in 1824. ([Elton and Gruber 2013](#)) Mutual funds play a vital role in the financial market. In 2021, the total market capitalisation in the US was about \$27 trillion, more than the nominal GDP of the US. Mutual funds, in terms of assets under management (AUM), are one of the two largest financial intermediaries in the US.

There are four types of mutual funds—the open-end mutual fund, the closed-end mutual fund, the exchange-traded mutual fund (ETF), and the unit investment trust (UIT). Open-end and closed-end mutual funds are traditional mutual funds. Shares of an open-end mutual fund can be purchased from and redeemed to the fund, while shares of a closed-end fund are traded on the market. The ETF is very young compared with traditional mutual funds. The first ETF was an S&P 500 Index ETF (SPY), launched in 1993. ETFs are very similar to closed-end mutual funds, but they allow authorised participants (APs) to exchange ETF shares with a basket of underlying assets or vice versa. As a result, ETFs

---

<sup>2</sup>For example, [Case and Shiller \(2003\)](#), [Piazzesi and Schneider \(2009\)](#)

can track the net asset value (NAV) closely, unlike closed-end mutual funds. UITs only account for a small share of the mutual fund market in terms of AUM.

## A. Fund Flows

The capital flows into and out of mutual funds have attracted many research interests. Gruber (1996) finds that past fund returns and capital flows can predict future fund performance. The author argues that the fund manager's ability is not priced in the NAV, so rational investors invest more in funds with better management. Sirri and Tufano (1998) also discover a strong relationship between past returns and fund flows. They find that the capital inflow is particularly large for funds with outstanding returns in the prior period, which the search cost can explain. They also investigate the impacts of other variables on flows, including the expense ratio and volatility.

Built on the initial discovery of the flow-performance relationship, the following research finds many other variables related to fund flows. Using Morningstar mutual funds data between 1993 and 1999, Bergstresser and Poterba (2002) argue that tax plays an important role in fund flows. They find that after-tax returns predict fund flows better than pre-tax returns. Unrealised gains of mutual funds reduce both inflows and outflows, but the effect is more substantial for capital inflows. Del Guercio and Tkac (2008) conduct event studies with Morningstar rating changes to explore the relationship between ratings and fund flows. They find that the Morningstar rating has independent effects on fund flows, even after controlling for various performance measurements. Using daily flow data of nearly 1000 mutual funds, Goetzmann et al. (2000) discover that equity and money mutual fund flows are negatively correlated, as are equity and metal fund flows. They argue that the negative correlations are due to sentiments about the equity premium, not simply liquidity concerns.

Ivković and Weisbenner (2009) delve into the fund flows of individual investors. They find that individual investors are subject to the lock-in effect due to the capital gain tax and are sensitive to fund costs. Investors make buying decisions based on the relative performance of the funds but make selling decisions based on absolute returns. Barber et al. (2016) find that investors consider the CAPM alpha when choosing a fund, but size, value and industry factors are ignored. Besides, more sophisticated investors use better benchmarks to choose a fund.

The flow-performance sensitivity (FPS) measures the sensitivity of fund flows in re-

response to past performance, which has attracted many research interests. Using mutual fund data from 28 countries between 2001 and 2007, [Ferreira et al. \(2012\)](#) find considerable heterogeneity of the FPS across countries. Investors in developed countries have lower FPS than those in developing countries, which the sophistication of investors can explain. [Huang et al. \(2007\)](#) develop a rational model with the learning of the manager's ability and the participation cost to explain the FPS. They show that the low participation cost is associated with high FPS to medium returns and low FPS to high returns, which is consistent with empirical findings. [Franzoni and Schmalz \(2017\)](#) find that the FPS is hump-shaped across deciles of the market return, which existing theories cannot explain. They develop a model with Bayesian investors who can learn from both alpha and beta to explain the shape of FPS.

## B. The Growth of Exchange-Traded Funds

Although young, the ETF gained popularity quickly, taking the market share from traditional mutual funds. Using CRSP and N-SAR data between 2000 and 2004, [Agapova \(2011\)](#) investigates how ETF and mutual fund flows interact. The paper shows that the ETF performs better regarding tracking errors. The ETF and the mutual fund are not perfect substitutes. The inflow to the ETF is associated with a 22% outflow from the mutual fund. The substitute effect became higher after the 2003 tax reform, which decreased the top-tier capital gain tax rate from 20% to 15%.

There is a long literature trying to explain the success of the ETF. [Lettau and Madhavan \(2018\)](#) argue that ETFs are usually more liquid than underlying securities. [Kostovetsky \(2003\)](#) builds a 1-period threshold model to compare the ETF and the mutual fund. He finds that the expense ratio, transaction costs, tax considerations, and qualitative factors are important in explaining the switch from mutual funds to ETFs. [Gardner and Welch \(2005\)](#) argue that ETFs offer investors greater flexibility due to their structural characteristics. Like stocks, ETFs are traded during regular trading hours, while mutual fund shares can only be redeemed after the market closes. Consequently, ETFs typically exhibit superior liquidity. Besides, holding a leveraged or short position of ETFs is easy, whereas short-selling mutual fund shares is not viable.

Tax efficiency is a crucial factor in the ETF's success. [Frei and Welsh \(2022\)](#) establish a rank-dependent expected utility model to find equilibrium market shares for ETFs and

mutual funds. They assume agents have constant absolute risk aversion and returns are drawn from a given distribution. Simulation shows that agents will move from ETFs to mutual funds if ETFs become taxable. Besides, authorised participants (APs) can create (destroy) ETF shares by giving (withdrawing) a basket of underlying securities to the ETF sponsor. By delivering securities with the lowest basis to the redeeming AP, the ETF sponsor can circumvent capital gain distributions to investors. This practice, known as "in-kind redemption," is tax-exempt by law. [Colon \(2017\)](#) argues that this tax exemption should be regarded as an "ETF tax subsidy". In cases where there are still positive realised capital gains despite day-to-day redemption, ETFs can collaborate with APs to execute the so-called "heartbeat" trade to eliminate any remaining realised capital gains. A "heartbeat" trade involves a large inflow of securities followed by a similar-size outflow to the same AP days later. [Moussawi et al. \(2022\)](#) find that ETFs and mutual funds realise similar capital gains, but the capital gain distribution of ETFs is almost 0. The tax difference produces an annual tax alpha of 0.92%. The result is robust to industry, value, and size. The paper also finds that the realised capital gain drives heartbeat trades and explains migration from active funds to ETFs. Exploiting the quasi-natural experiment of the 2012 tax reform, the paper finds that high-net-worth investors switch from mutual funds to ETFs when the tax rate increases after 2012.

The holding tax efficiency of the ETF comes from the tax-exempt in-kind redemption. Looking at the returns and tax distributions of SPY and Vanguard Index 500 Fund between 1994 and 2000, [Poterba and Shoven \(2002\)](#) find the two funds have similar returns and dividends, but SPY distributes almost no capital gain due to in-kind redemption. Some papers criticise the exploitation of in-kind redemption as an ETF tax subsidy. ([Colon 2017](#), [Hodaszy 2016](#), [2022](#)) First documented by [Kashner \(2017\)](#), heartbeat trades are found to be used to eliminate realised capital gains of ETFs, enhancing after-tax returns by 6 bps per quarter. [Moussawi et al. \(2022\)](#) find that heartbeat trades can reduce the tax burden by 0.86% annually. [Colon \(2023\)](#) notices that Congress set limits on in-kind redemption in 2012, but it still allows non-pro-rata security distribution due to index rebalancing, which is exploited by heartbeat trades.

Although many papers document the advantages of ETFs over mutual funds, it does not mean there is unanimity that ETFs dominate mutual funds. For example, [Gastineau \(2004\)](#) finds that the Russell 2000 mutual fund performs better than the ETF tracking the

same index in 2001 and 2002. The author argues that mutual funds have better operational efficiency because they do not need to keep the security basket fixed for the whole trading day.

### C. Exchange-Traded Funds and Market Efficiency

The price of an ETF is usually very close to the value of its underlying portfolio because APs can arbitrage in the primary market if the parity does not hold.<sup>3</sup> This specific feature makes the ETF suitable for investigating market efficiency.

[Hasbrouck \(2003\)](#) is among the first to look into the price discovery ability of ETFs. Using NYSE Trade and Quote (TAQ) intraday records of three popular indices, the author finds that most price discovery occurs in the E-mini market. For the index without E-mini futures, price discovery is shared between the regular futures contract and the ETF. [Fang and Sanger \(2011\)](#) claim that a significant portion of ETF trading is motivated by private information, which is important for price discovery.

Several papers delve deep into the arbitrage with the ETF and its underlying assets in the primary market. [Marshall et al. \(2013\)](#) find that liquidity explains the occurrence of arbitrage opportunities, which are quickly eliminated by arbitrages. The medium time of the price correction is 2.27 minutes for SPY and 0.92 minutes for IVV. [Brown et al. \(2021\)](#) show that arbitrage flows predict future returns. The long-short portfolio based on ETF flows has a return of 1.1-2% per month and an annual Sharpe ratio of 0.6-0.99. Using the N-CEN filings from the SEC, [Gorbatikov and Sikorskaya \(2022\)](#) identify the ETF-AP network in the US market. They find that ETF mispricing is negatively correlated with network diversity, and ETFs sharing more APs have more similar mispricing.

The ETF is also connected with the market efficiency of underlying stocks. [Yu \(2005\)](#) shows that ETF contains substantial information on stock prices. Informational efficiency and liquidity of underlying stocks increase after the introduction of ETFs. Exploiting the mechanism of price limits in the Chinese market, [Chen and Strother \(2008\)](#) find that the SSE 50 ETF helps discover underlying stock prices when price limits are reached. Besides, [Brown et al. \(2021\)](#) find that ETF flows predict returns of both ETF and its underlying assets. However, some papers reach different conclusions. [Chen et al. \(2024\)](#) argue that high ETF ownership leads to low market efficiency of underlying stocks due to the trading

---

<sup>3</sup>For detailed illustrations, see [Brown et al. \(2021\)](#) and [Gorbatikov and Sikorskaya \(2022\)](#).

noises passed on from ETFs. Using minute-level data, [Box et al. \(2021\)](#) discover that ETFs have minimal impacts on the returns of underlying assets, while the returns of underlying assets significantly affect the quotes of ETFs.

## 2.3 Capital Gain Taxation

By surveying the taxation literature, [Hanlon and Heitzman \(2010\)](#) summarise four main areas of the taxation research: (1) the informational role of income tax expense reported for financial accounting, (2) corporate tax avoidance, (3) corporate decision-making, and (4) taxes and asset pricing. The research on the fourth area concentrates on two taxes—the dividend tax and the capital gain tax. The capital gain taxation literature is particularly relevant to this thesis.

### A. Market Impact

The existence of the capital gain tax alters investors' behaviour. [Holt and Shelton \(1961\)](#) argue that the capital gain tax encourages investors to defer the realisation of capital gains, known as the lock-in effect. Using a lifetime model, [Holt and Shelton \(1962\)](#) further identify variables that affect lock-in decisions. The lock-in effect distorts the financial market and harms market efficiency, leading to negative macroeconomic impacts. ([David 1964](#)) In addition, [Poterba \(1987\)](#) finds that the capital gain tax rate is negatively associated with taxpayer compliance. A 1% increase in the marginal capital gain tax rate reduces voluntary compliance by 0.5% to 1%.

The lock-in effect means the optimal portfolio choice with the capital gain tax differs from the one without the capital gain tax. [Dammon et al. \(2001\)](#) build a model to investigate the optimal consumption and investment strategies with the capital gain tax. With a random life expectancy, agents maximise the expected utility of lifetime consumption and terminal wealth. There are two assets in the model—risky and risk-free assets. Short sales are not allowed. The model contains many decision-making periods, from 20 to 100 years old. To keep the model tractable, the paper uses the moving average tax basis of the asset instead of the exact tax basis. The model is solved numerically. [Dammon et al. \(2004\)](#) update the model of [Dammon et al. \(2001\)](#) to allow the co-existence of two accounts—taxable and tax-deferral accounts. Short sales are allowed. They find a strong tendency to invest

the equity in the taxable account and the taxable bond in the tax-deferral account. Tax consideration outweighs the liquidity requirement.

[DeMiguel and Uppal \(2005\)](#) use non-linear programming to solve for the optimal portfolio with the exact tax basis. Due to the high dimensionality problem, the paper only simulates the model with up to ten periods and two assets. Short sales are not allowed in the model. Simulations predict a tax burden as low as 1% with the optimal tax-loss harvesting strategy. The paper also compares the welfare loss by approximating the tax basis with the moving average, finding that the certainty equivalence loss is 1%. Furthermore, [Duarte et al. \(2021\)](#) use a machine learning algorithm to solve for the optimal portfolio choice in a lifecycle model that includes many real-life features. Feasible assets include stocks, bonds, and liquid accounts. They estimate that the welfare loss resulting from the age-dependent rule of current target-date/lifecycle funds is 2-3% of total consumption.

The capital gain tax reduces after-tax returns of factors. [Bergstresser and Pontiff \(2013\)](#) collect tax regulations between 1927 and 2009, calculating the effective tax rates and after-tax performance for several strategies. They investigate the effective tax rates for different quantiles of income. The paper shows that the momentum strategy has the best pre-tax performance, but the value strategy has the best after-tax performance. The value and size premia are lower after deducting the tax. [Sialm \(2009\)](#) argues that the tax burden is capitalised. i.e. Investors factor in the tax impact when pricing securities. Sialm calculates effective tax rates in both time series and cross-section. Using a multifactor pricing model, the paper finds that stocks with high dividends tend to have high returns, especially for high-tax periods.

## B. Tax-Loss Harvesting

The capital gain tax encourages tax-loss harvesting behaviour. [Chaudhuri et al. \(2020\)](#) backtest the tax-loss harvesting alpha between 1926 and 2018, estimating an average tax alpha of 1.08% when the wash-sale rule does not apply and 0.82% when the wash-sale rule applies. The tax alpha is higher in recessions than expansions, peaking at 2.13% during the great depression. Besides, the tax alpha increases with return volatility, drawdown, tax rate, and cash contribution.

Some papers find that the turn-of-year effect is associated with tax-loss harvesting. [Poterba and Weisbenner \(2001\)](#) show that the return in the first five trading days in

January is negatively associated with past losses, while past gains have no impact on the return in the new year. The effect is found to be stronger for small firms. The magnitude of the effect depends on tax laws, confirming the existence of tax-loss harvesting behaviour. [Ivković et al. \(2005\)](#) investigate the disposition and lock-in effects with taxable and tax-deferred accounts of 78000 households. They find that tax-loss harvesting happens all year round, while December witnesses above-average (16.9%) loss sales. Positions purchased within six months are more likely to be sold for tax purposes. Capital gains lock-in effect exists in taxable accounts, especially for large and long-term positions. The disposition effect also exists. [Agarwal et al. \(2014\)](#) also find that tax-loss harvesting happens all year round. [Grinblatt and Moskowitz \(2004\)](#) find that tax-loss harvesting is associated with the momentum effect. Using hedged portfolios neutral to value, size, and industry, they find that consistent winner stocks have positive returns. Winner and loser stocks behave in opposite ways around the turn of the year, indicating the tax-loss selling. Small stocks with high turnover and low institutional ownership exhibit more pronounced turn-of-year and seasonal effects.

Using a model similar to [Dammon et al. \(2001\)](#), [Garlappi et al. \(2001\)](#) look at the value of tax-loss harvesting. There are two risky assets and one risk-free asset in the model. They argue that the tax-deferral feature of securities can be viewed as an option, with a value of 5-10% of the wealth. The option value decreases with volatility, risk aversion, and correlation between risk assets.

## References

- Adam, K., Matveev, D. and Nagel, S. (2021), ‘Do survey expectations of stock returns reflect risk adjustments?’, *Journal of Monetary Economics* **117**, 723–740.
- Agapova, A. (2011), ‘Conventional mutual index funds versus exchange-traded funds’, *Journal of Financial Markets* **14**(2), 323–343.
- Agarwal, V., Gay, G. D. and Ling, L. (2014), ‘Window dressing in mutual funds’, *The Review of Financial Studies* **27**(11), 3133–3170.
- Ai, H., Croce, M. M., Diercks, A. M. and Li, K. (2018), ‘News shocks and the production-

- based term structure of equity returns', *The Review of Financial Studies* **31**(7), 2423–2467.
- Andries, M., Eisenbach, T. M. and Schmalz, M. C. (2019), 'Horizon-dependent risk aversion and the timing and pricing of uncertainty', *FRB of New York Staff Report* .
- Andries, M., Eisenbach, T. M., Schmalz, M. C. and Wang, Y. (2015), 'The term structure of the price of variance risk', *FRB of New York Staff Report* .
- Bai, H. and Zhang, L. (2022), 'Searching for the equity premium', *Journal of Financial Economics* **143**(2), 897–926.
- Bansal, R., Miller, S., Song, D. and Yaron, A. (2021), 'The term structure of equity risk premia', *Journal of Financial Economics* .
- Bansal, R. and Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The journal of Finance* **59**(4), 1481–1509.
- Barber, B. M., Huang, X. and Odean, T. (2016), 'Which factors matter to investors? evidence from mutual fund flows', *The Review of Financial Studies* **29**(10), 2600–2642.
- Bekaert, G., Engstrom, E. C. and Xu, N. R. (2019), 'The time variation in risk appetite and uncertainty', *National Bureau of Economic Research Working Paper Series* .
- Bekaert, G. and Hoerova, M. (2014), 'The vix, the variance premium and stock market volatility', *Journal of econometrics* **183**(2), 181–192.
- Benartzi, S. and Thaler, R. H. (1995), 'Myopic loss aversion and the equity', *The Quarterly Journal of Economics* **110**(1), 73–92.
- Bergstresser, D. and Pontiff, J. (2013), 'Investment taxation and portfolio performance', *Journal of Public Economics* **97**, 245–257.
- Bergstresser, D. and Poterba, J. (2002), 'Do after-tax returns affect mutual fund inflows?', *Journal of financial economics* **63**(3), 381–414.
- Bollerslev, T., Tauchen, G. and Zhou, H. (2009), 'Expected stock returns and variance risk premia', *The Review of Financial Studies* **22**(11), 4463–4492.

- Bollerslev, T., Xu, L. and Zhou, H. (2015), ‘Stock return and cash flow predictability: The role of volatility risk’, *Journal of Econometrics* **187**(2), 458–471.
- Box, T., Davis, R., Evans, R. and Lynch, A. (2021), ‘Intraday arbitrage between etfs and their underlying portfolios’, *Journal of Financial Economics* **141**(3), 1078–1095.
- Brown, D. C., Davies, S. W. and Ringgenberg, M. C. (2021), ‘Etf arbitrage, non-fundamental demand, and return predictability’, *Review of Finance* **25**(4), 937–972.
- Caballero, R. J. and Simsek, A. (2020), ‘A risk-centric model of demand recessions and speculation’, *The Quarterly Journal of Economics* **135**(3), 1493–1566.
- Campbell, J. Y. and Cochrane, J. H. (1999), ‘By force of habit: A consumption-based explanation of aggregate stock market behavior’, *Journal of political Economy* **107**(2), 205–251.
- Campbell, J. Y. and Hentschel, L. (1992), ‘No news is good news: An asymmetric model of changing volatility in stock returns’, *Journal of financial Economics* **31**(3), 281–318.
- Campbell, J. Y. and Shiller, R. J. (1988), ‘The dividend-price ratio and expectations of future dividends and discount factors’, *The Review of Financial Studies* **1**(3), 195–228.
- Case, K. E. and Shiller, R. J. (2003), ‘Is there a bubble in the housing market?’, *Brookings papers on economic activity* **2003**(2), 299–362.
- Chabi-Yo, F., Gonçalves, A. and Loudis, J. (2021), ‘An intertemporal risk factor model’, *Kenan Institute of Private Enterprise Research Paper Forthcoming* .
- Chabi-Yo, F. and Loudis, J. (2020), ‘The conditional expected market return’, *Journal of Financial Economics* **137**(3), 752–786.
- Chaudhuri, S. E., Burnham, T. C. and Lo, A. W. (2020), ‘An empirical evaluation of tax-loss-harvesting alpha’, *Financial Analysts Journal* **76**(3), 99–108.
- Chen, G., Liu, X., Liu, X. and Zhao, Z. (2024), ‘Etf ownership and stock pricing efficiency: The role of etf arbitrage’, *Finance Research Letters* p. 105108.
- Chen, G.-m., Firth, M. and Rui, O. M. (2001), ‘The dynamic relation between stock returns, trading volume, and volatility’, *Financial Review* **36**(3), 153–174.

- Chen, G. and Strother, T. S. (2008), ‘On the contribution of index exchange traded funds to price discovery in the presence of price limits without short selling’, *Available at SSRN 1094485*.
- Chen, Z. (2020), ‘Inferring stock duration around fomc surprises: Estimates and implications’, *Journal of Financial and Quantitative Analysis* pp. 1–35.
- Cheng, I.-H. (2019), ‘The vix premium’, *The Review of Financial Studies* **32**(1), 180–227.
- Collin-Dufresne, P., Johannes, M. and Lochstoer, L. A. (2017), ‘Asset pricing when ‘this time is different’’, *The Review of Financial Studies* **30**(2), 505–535.
- Colon, J. M. (2017), ‘The great etf tax swindle: The taxation of in-kind redemptions’, *Penn St. L. Rev.* **122**, 1.
- Colon, J. M. (2023), ‘Unplugging heartbeat trades and reforming the taxation of etfs’.
- Dammon, R. M., Spatt, C. S. and Zhang, H. H. (2001), ‘Optimal consumption and investment with capital gains taxes’, *The Review of Financial Studies* **14**(3), 583–616.
- Dammon, R. M., Spatt, C. S. and Zhang, H. H. (2004), ‘Optimal asset location and allocation with taxable and tax-deferred investing’, *The Journal of Finance* **59**(3), 999–1037.
- Daniel, K. and Titman, S. (1997), ‘Evidence on the characteristics of cross sectional variation in stock returns’, *the Journal of Finance* **52**(1), 1–33.
- David, M. (1964), ‘Economic effects of the capital gains tax’, *The American Economic Review* **54**(3), 288–299.
- De La O, R. and Myers, S. (2021), ‘Subjective cash flow and discount rate expectations’, *The Journal of Finance* **76**(3), 1339–1387.
- Dechow, P. M., Erhard, R. D., Sloan, R. G. and SOLIMAN, M. T. (2021), ‘Implied equity duration: A measure of pandemic shutdown risk’, *Journal of Accounting Research* **59**(1), 243–281.
- Del Guercio, D. and Tkac, P. A. (2008), ‘Star power: The effect of monrningstar ratings on mutual fund flow’, *Journal of Financial and Quantitative Analysis* **43**(4), 907–936.

- DeMiguel, V. and Uppal, R. (2005), ‘Portfolio investment with the exact tax basis via nonlinear programming’, *Management Science* **51**(2), 277–290.
- Dew-Becker, I., Giglio, S., Le, A. and Rodriguez, M. (2017), ‘The price of variance risk’, *Journal of Financial Economics* **123**(2), 225–250.
- Dimson, E., Marsh, P. and Staunton, M. (2006), ‘The worldwide equity premium: A smaller puzzle’, *Chapter 11*, 467–514.
- Drechsler, I. and Yaron, A. (2011), ‘What’s vol got to do with it’, *The Review of Financial Studies* **24**(1), 1–45.
- Duarte, V., Fonseca, J., Goodman, A. S. and Parker, J. A. (2021), Simple allocation rules and optimal portfolio choice over the lifecycle, Technical report, National Bureau of Economic Research.
- D’acunto, F., Hoang, D., Paloviita, M. and Weber, M. (2023), ‘IQ, expectations, and choice’, *The Review of Economic Studies* **90**(5), 2292–2325.
- Easley, D. and O’Hara, M. (1992), ‘Adverse selection and large trade volume: The implications for market efficiency’, *Journal of Financial and Quantitative Analysis* **27**(2), 185–208.
- Elton, E. J. and Gruber, M. J. (2013), Mutual funds, *in* ‘Handbook of the Economics of Finance’, Vol. 2, Elsevier, pp. 1011–1061.
- Fama, E. F. and French, K. R. (1993), ‘Common risk factors in the returns on stocks and bonds’, *Journal of financial economics* **33**(1), 3–56.
- Fama, E. F. and French, K. R. (2002), ‘The equity premium’, *The Journal of Finance* **57**(2), 637–659.
- Fama, E. F. and French, K. R. (2015), ‘A five-factor asset pricing model’, *Journal of financial economics* **116**(1), 1–22.
- Fang, Y. and Sanger, G. C. (2011), Index price discovery in the cash market, *in* ‘Midwest Finance Association 2012 Annual Meetings Paper’.

- Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2012), ‘The flow-performance relationship around the world’, *Journal of Banking & Finance* **36**(6), 1759–1780.
- Franzoni, F. and Schmalz, M. C. (2017), ‘Fund flows and market states’, *The Review of Financial Studies* **30**(8), 2621–2673.
- Frei, C. and Welsh, L. (2022), ‘How the closure of a us tax loophole may affect investor portfolios’, *Journal of Risk and Financial Management* **15**(5), 209.
- French, K. R., Schwert, G. W. and Stambaugh, R. F. (1987), ‘Expected stock returns and volatility’, *Journal of financial Economics* **19**(1), 3–29.
- Gabaix, X. (2012), ‘Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance’, *The Quarterly journal of economics* **127**(2), 645–700.
- Gardner, R. and Welch, J. (2005), ‘Increasing after-tax return with exchange-traded funds’, *Journal of Financial Planning* **18**(6), 30.
- Garlappi, L., Naik, V. and Slive, J. (2001), ‘Portfolio selection with multiple assets and capital gains taxes’, *Available at SSRN 274939* .
- Gastineau, G. L. (2004), ‘The benchmark index etf performance problem’, *The Journal of Portfolio Management* **30**(2), 96–103.
- Giglio, S., Kelly, B. T. and Kozak, S. (2021), ‘Equity term structures without dividend strips data’, *Available at SSRN 3533486* .
- Giglio, S., Maggiori, M., Stroebel, J. and Utkus, S. (2021), ‘Five facts about beliefs and portfolios’, *American Economic Review* **111**(5), 1481–1522.
- Goetzmann, W. N., Massa, M., Rouwenhorst, K. G. et al. (2000), *Behavioral factors in mutual fund flows*, INSEAD.
- Golez, B. and Jackwerth, J. C. (2020), ‘Quarter-of-a-century of dividend pricing’, *Available at SSRN 3575450* .
- Gonçalves, A. S. (2021a), ‘Reinvestment risk and the equity term structure’, *The Journal of Finance* .
- Gonçalves, A. S. (2021b), ‘The short duration premium’, *Journal of Financial Economics* .

- Gorbatikov, E. and Sikorskaya, T. (2022), ‘Two aps are better than one: Etf mispricing and primary market participation’, *Available at SSRN 3923503*.
- Gormsen, N. J. (2021), ‘Time variation of the equity term structure’, *The Journal of Finance* **76**(4), 1959–1999.
- Gormsen, N. J. and Lazarus, E. (2021), ‘Duration-driven returns’, *Available at SSRN 3359027*.
- Greenwood, R. and Shleifer, A. (2014), ‘Expectations of returns and expected returns’, *The Review of Financial Studies* **27**(3), 714–746.
- Grinblatt, M. and Moskowitz, T. J. (2004), ‘Predicting stock price movements from past returns: the role of consistency and tax-loss selling’, *Journal of Financial Economics* **71**, 541–579.
- Gruber, M. J. (1996), ‘Another puzzle: The growth in actively managed mutual funds’, *The journal of finance* **51**(3), 783–810.
- Hanlon, M. and Heitzman, S. (2010), ‘A review of tax research’, *Journal of accounting and Economics* **50**(2-3), 127–178.
- Hasbrouck, J. (2003), ‘Intraday price formation in us equity index markets’, *The Journal of Finance* **58**(6), 2375–2400.
- Herskovic, B., Kelly, B., Lustig, H. and Van Nieuwerburgh, S. (2016), ‘The common factor in idiosyncratic volatility: Quantitative asset pricing implications’, *Journal of Financial Economics* **119**(2), 249–283.
- Hodaszy, S. Z. (2016), ‘Tax-efficient structure or tax shelter: Curbing etfs’ use of section 852 (b)(6) for tax avoidance’, *Tax Law*. **70**, 537.
- Hodaszy, S. Z. (2022), ‘Exchange-traded funds use section 852 (b)(6) for tax avoidance, not just tax deferral: So why is this loophole still open?’, *Tax Lawyer* **75**(3), 489.
- Holt, C. C. and Shelton, J. P. (1961), ‘The implications of the capital gains tax for investment decisions’, *The Journal of Finance* **16**(4), 559–580.

- Holt, C. C. and Shelton, J. P. (1962), ‘The lock-in effect of the capital gains tax’, *National Tax Journal* **15**(4), 337–352.
- Holthausen, R. W. and Verrecchia, R. E. (1990), ‘The effect of informedness and consensus on price and volume behavior’, *Accounting Review* pp. 191–208.
- Hou, K., Xue, C. and Zhang, L. (2015), ‘Digesting anomalies: An investment approach’, *The Review of Financial Studies* **28**(3), 650–705.
- Huang, J., Wei, K. D. and Yan, H. (2007), ‘Participation costs and the sensitivity of fund flows to past performance’, *The journal of finance* **62**(3), 1273–1311.
- Ivković, Z., Poterba, J. and Weisbenner, S. (2005), ‘Tax-motivated trading by individual investors’, *American Economic Review* **95**(5), 1605–1630.
- Ivković, Z. and Weisbenner, S. (2009), ‘Individual investor mutual fund flows’, *Journal of Financial Economics* **92**(2), 223–237.
- Jagannathan, R., McGrattan, E. and Scherbina, A. (2001), ‘The declining us equity premium’.
- Karpoff, J. M. (1986), ‘A theory of trading volume’, *The journal of finance* **41**(5), 1069–1087.
- Karpoff, J. M. (1987), ‘The relation between price changes and trading volume: A survey’, *Journal of Financial and Quantitative Analysis* **22**(1), 109–126.
- Kashner, E. (2017), ‘The heartbeat of etf tax efficiency’.  
**URL:** <https://insight.factset.com/the-heartbeat-of-etf-tax-efficiency>
- Kim, O. and Verrecchia, R. E. (1991), ‘Trading volume and price reactions to public announcements’, *Journal of accounting research* **29**(2), 302–321.
- Kostovetsky, L. (2003), ‘Index mutual funds and exchange-traded funds’, *Journal of Portfolio Management* **29**(4), 80.
- Lettau, M. and Madhavan, A. (2018), ‘Exchange-traded funds 101 for economists’, *Journal of Economic Perspectives* **32**(1), 135–154.

- Llorente, G., Michaely, R., Saar, G. and Wang, J. (2002), ‘Dynamic volume-return relation of individual stocks’, *The Review of financial studies* **15**(4), 1005–1047.
- Manela, A. and Moreira, A. (2017), ‘News implied volatility and disaster concerns’, *Journal of Financial Economics* **123**(1), 137–162.
- Manski, C. F. (2018), ‘Survey measurement of probabilistic macroeconomic expectations: progress and promise’, *NBER Macroeconomics Annual* **32**(1), 411–471.
- Marshall, B. R., Nguyen, N. H. and Visaltanachoti, N. (2013), ‘Etf arbitrage: Intraday evidence’, *Journal of Banking & Finance* **37**(9), 3486–3498.
- Martin, I. (2017), ‘What is the expected return on the market?’, *The Quarterly Journal of Economics* **132**(1), 367–433.
- Martin, I. W. and Wagner, C. (2019), ‘What is the expected return on a stock?’, *The Journal of Finance* **74**(4), 1887–1929.
- Mehra, R. (2003), ‘The equity premium: Why is it a puzzle?’.
- Mehra, R. and Prescott, E. C. (1985), ‘The equity premium: A puzzle’, *Journal of monetary Economics* **15**(2), 145–161.
- Moench, E. and Siavash, S. S. (2021), ‘What moves treasury yields?’, *CEPR Discussion Paper Series* .
- Moussawi, R., Shen, K. and Velthuis, R. (2022), ‘The role of taxes in the rise of etfs’, *Available at SSRN 3744519* .
- Nagel, S. and Xu, Z. (2022), ‘Asset pricing with fading memory’, *The Review of Financial Studies* **35**(5), 2190–2245.
- Odean, T. (1998), ‘Volume, volatility, price, and profit when all traders are above average’, *The journal of finance* **53**(6), 1887–1934.
- Piazzesi, M. and Schneider, M. (2009), ‘Momentum traders in the housing market: Survey evidence and a search model’, *American Economic Review* **99**(2), 406–411.
- Pindyck, R. S. (1984), ‘Risk, inflation, and the stock market’, *The American Economic Review* **74**(3), 335–351.

- Poterba, J. M. (1987), ‘Tax evasion and capital gains taxation’.
- Poterba, J. M. and Shoven, J. B. (2002), ‘Exchange-traded funds: A new investment option for taxable investors’, *American Economic Review* **92**(2), 422–427.
- Poterba, J. M. and Weisbenner, S. J. (2001), ‘Capital gains tax rules, tax-loss trading, and turn-of-the-year returns’, *The Journal of Finance* **56**(1), 353–368.
- Scheinkman, J. A. and Xiong, W. (2003), ‘Overconfidence and speculative bubbles’, *Journal of Political Economy* **111**(6), 1183–1220.
- Shiller, R. J. et al. (1981), ‘Do stock prices move too much to be justified by subsequent changes in dividends?’.
- Sialm, C. (2009), ‘Tax changes and asset pricing’, *American Economic Review* **99**(4), 1356–1383.
- Sirri, E. R. and Tufano, P. (1998), ‘Costly search and mutual fund flows’, *The journal of finance* **53**(5), 1589–1622.
- Van Binsbergen, J., Brandt, M. and Koijen, R. (2012), ‘On the timing and pricing of dividends’, *American Economic Review* **102**(4), 1596–1618.
- Van Binsbergen, J. H. and Koijen, R. S. (2017), ‘The term structure of returns: Facts and theory’, *Journal of Financial Economics* **124**(1), 1–21.
- Van Binsbergen, J., Hueskes, W., Koijen, R. and Vrugt, E. (2013), ‘Equity yields’, *Journal of Financial Economics* **110**(3), 503–519.
- Vasquez, A. (2017), ‘Equity volatility term structures and the cross section of option returns’, *Journal of Financial and Quantitative Analysis* **52**(6), 2727–2754.
- Wachter, J. A. and Kahana, M. J. (2024), ‘A retrieved-context theory of financial decisions’, *The Quarterly Journal of Economics* **139**(2), 1095–1147.
- Wang, J. (1994), ‘A model of competitive stock trading volume’, *Journal of Political Economy* **102**(1), 127–168.
- Weber, M. (2018), ‘Cash flow duration and the term structure of equity returns’, *Journal of Financial Economics* **128**(3), 486–503.

- Weil, P. (1989), ‘The equity premium puzzle and the risk-free rate puzzle’, *Journal of monetary economics* **24**(3), 401–421.
- Wu, D. (2020), ‘A disaster explanation of equity term structures’, *Available at SSRN 3571921* .
- Yu, L. (2005), ‘Basket securities, price formation, and informational efficiency’, *Price Formation, and Informational Efficiency (March 25, 2005)* .

## Part II

### Original Research Essays

# 3

## Term Structure of Equity Return Volatility

### Abstract

The term structure of equity return volatility fluctuates across time. It affects the term structure of equity returns through the volatility feedback effect and explains the cyclical nature of equity return term structure. By analysing the dividend strip futures, this paper finds that volatility feedback effects of dividend strips exist and decrease with the horizon. Using realised and implied volatilities as business cycle indicators, this paper confirms that the term structure of equity returns is pro-cyclical. The decomposition of cyclical nature shows that the pro-cyclical term structure of equity returns comes from the high relative sensitivity of short-duration volatility. The predictability of cyclical nature by the term structure of volatility is a novel feature that can be used to test macro-finance models. The rare disaster model proposed by [Gabaix \(2012\)](#) is rejected by the test.

**Keywords:** Term Structure of Equity Returns, Volatility, Return Prediction

### 3.1 Introduction

In recent years, stock index dividend strips have attracted much attention from the financial industry and academia. The trading volume of one-year dividend strip futures in December 2020 is almost four times as high as that in December 2015.<sup>1</sup> [Van Binsbergen et al. \(2012\)](#) first document dividend strip returns across horizons, showing that the unconditional term structure of equity returns is downward sloping.

A long strand of literature links stock return volatility with stock returns. [Campbell and Hentschel \(1992\)](#) find that volatility news raises the required return, especially during volatile periods. They call it the volatility feedback effect. [Bansal et al. \(2014\)](#) and [Campbell et al. \(2018\)](#) provide news-based theoretical models to include volatility as a risk factor. However, whether the volatility feedback effect is present for dividend strips at all horizons is unknown. This paper fills the research gap by linking the term structure of equity returns and the term structure of equity return volatility.

[Giglio and Kelly \(2018\)](#) discover an unconditional upward-sloping term structure of equity return volatility with variance swap data. As shown in Figure 3.1, this paper also documents an upward-sloping and concave term structure of equity return volatility. The shape is similar even if the most turbulent periods are excluded.

However, the variation of the term structure behaves in a less monotonic way. As shown in Figure 3.2, in February 2020, before the first lockdown in the US, the return volatility is calm and increases with the horizon. In March, volatilities of all horizons increase, and the term structure has a hump shape. Volatility drops dramatically in the following two months, and the shape turns back to be upward sloping. Equity return volatility varies in both cross-section and time series.

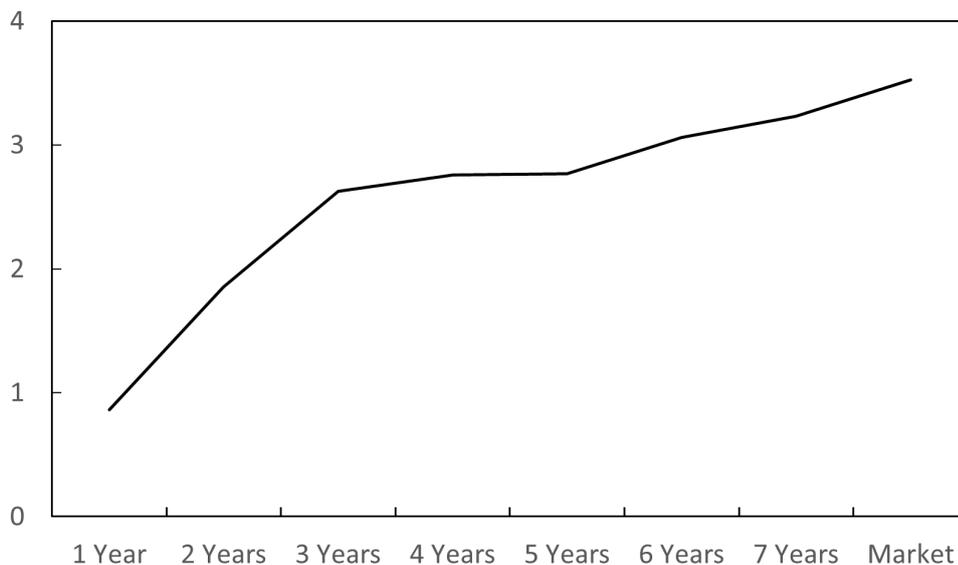
This paper asks the question: Does the term structure of equity return volatility explain the variation of equity return term structure? The term structure of equity returns can be investigated in two aspects—the variations of its level and its slope. In the literature, the variation of the slope is usually referred to as cyclical. To the best of my knowledge, this paper is the first one that connects the term structure of equity return volatility and the term structure of equity returns.

This research question is essential in the finance literature. Empirically, it extends the

---

<sup>1</sup>The number comes from publicly traded S&P 500 dividend futures in CME. Source: Bloomberg.

Figure 3.1: Unconditional Term Structure of Volatility



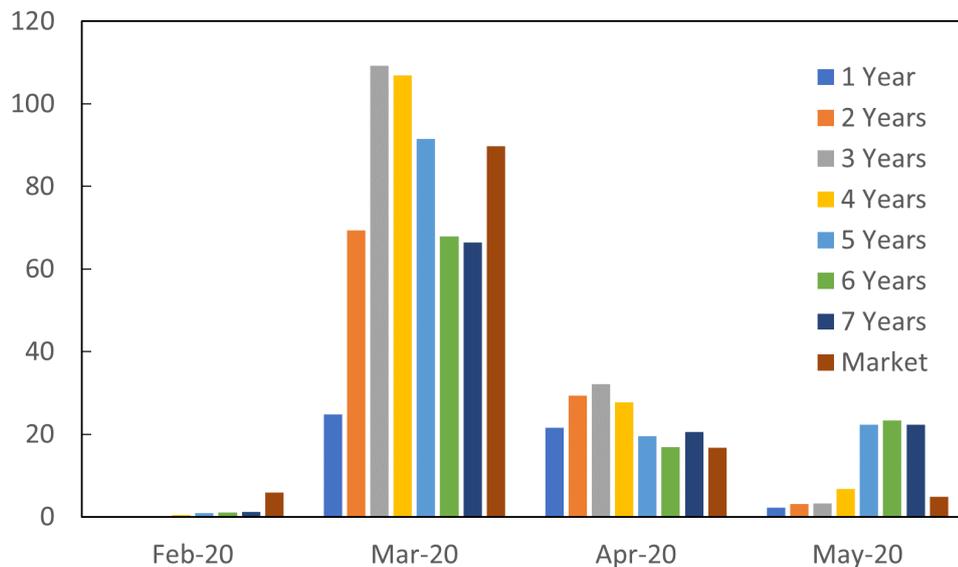
Volatility is measured by the realised variance of daily dividend strip returns within a month. Volatilities of 1- to 5-year dividend strips and the market use S&P 500 dividend futures data from December 2015 to July 2021. Volatilities of 6- and 7-year dividend strips use S&P 500 dividend futures data from January 2017 to July 2021.

research on the volatility feedback effect to dividend strips. Besides, none of the selected mainstream macro-finance models—the preferred habit model ([Campbell and Cochrane 1999](#)), the long-run risk model ([Bansal and Yaron 2004](#)), and the rare disaster model ([Gabaix 2012](#))—predict that the term structure of volatility explains cyclicity. This paper can potentially reject these models by running empirical tests.

To answer this question, this paper dissects return volatility into different horizons and explores how they affect dividend strip returns through the volatility feedback effect. Then, it examines the cyclicity of the equity return term structure with realised and implied volatilities. This paper also decomposes the cyclicity based on the term structure of equity return volatility. Finally, this study simulates three popular macro-finance models to see whether they are compatible with the empirical finding of cyclicity.

The main findings of this paper can be summarised as follows. First, although the realised volatility of the S&P 500 index does not explain the future market returns, the realised variance of dividend strips positively explains future dividend strip returns. The volatility feedback effect decreases with the horizon. These results are consistent with [Berger et al. \(2020\)](#) and [Dew-Becker et al. \(2021\)](#). Both expected and unexpected volatilities contribute to the monotonic pattern of the volatility feedback effects, but the expected

Figure 3.2: Term Structure of Volatility in Early 2020



Volatility is measured by the realised variance of daily dividend strip returns within a month. This figure uses S&P 500 dividend futures data.

volatility has opposite effects in short and medium horizons. Besides, the empirical results show that the equity term premium is pro-cyclical, supporting the findings of [Bansal et al. \(2021\)](#). The decomposition of cyclicity indicates that the relatively high sensitivity of short-duration volatility leads to the pro-cyclical term structure. Finally, simulations show that the rare disaster model should be rejected based on the cyclicity test. The preferred habit model and the long-run risk model provide no evidence of fitting the data, although the test does not reject them.

The paper proceeds as follows. First, Section 3.2 reviews two related literature strands: the volatility feedback effect and the term structure of equity returns. Section 3.3 presents how the term structure of volatility affects returns. Next, Section 3.4 delves into the cyclicity of the term structure of equity returns. Section 3.5 simulates three macro-finance models and contrasts them with empirical findings. Robustness analysis will be discussed in Section 3.6. Finally, Section 3.7 concludes.

## 3.2 Literature Review

This paper contributes mainly to two strands of literature—the volatility feedback effect and the term structure of equity returns.

## A. Volatility Feedback Effect

Some evidence of the volatility feedback effect can be traced back to [Pindyck \(1984\)](#). Notably, [French et al. \(1987\)](#) first document the effect systemically, finding that the predicted volatility explains the required return positively. Using the QGARCH model, [Campbell and Hentschel \(1992\)](#) find that the volatility feedback effect is asymmetric—high volatility periods have high feedback effects. [Martin \(2017\)](#) derives the lower bound of equity premium with implied volatility calculated from option prices, showing that the lower bound is tight and predicts the realised equity premium. [Chabi-Yo and Loudis \(2020\)](#) further derive the upper bound with high-order moments.

Volatility matters not only in time series but also in cross-sections. [Herskovic et al. \(2016\)](#) find that the common idiosyncratic volatility (CIV) predicts dividend growth and explains cross-sectional equity returns. It can also explain the value factor. [Chabi-Yo et al. \(2021\)](#) derive an extended ICAPM model with equity yield and realised volatility as risk factors. They argue that the volatility factor conveys a negative price and cannot be explained by other prevalent factors. [Martin and Wagner \(2019\)](#) derive risk neutral variance from stock options, showing that high variance is associated with small and value stocks. Moreover, their model explains over 30% of equity returns. [Vasquez \(2017\)](#) shows implied volatility is positively associated with straddle returns in cross-section, which popular risk factors cannot explain.

Recent research recognises that the difference between risk-neutral and physically expected volatility, called variance risk premium (VRP), is vital in explaining asset returns. [Bollerslev et al. \(2009\)](#) build a model based on the long-run risk model with time-varying volatility of volatility (VoV). With options data from 1990 to 2007, they find that high VRP predicts high returns, especially at the quarterly horizon. [Drechsler and Yaron \(2011\)](#) construct VRP as the difference between the squared VIX index and the expected realised volatility, making a similar conclusion. [Dew-Becker et al. \(2017\)](#) use variance swap to estimate the prices of variance risk in different horizons, finding that only very-short-term variance (up to 2 months) is priced. With option prices, [Andries et al. \(2015\)](#) find that the prices of variance risk are negative and decrease with the horizon in absolute terms. They also show that high volatility is associated with a steep price structure. VRP predicts both return ([Bekaert and Hoerova 2014](#)) and dividend growth ([Bollerslev et al. 2015](#)). [Bekaert](#)

[et al. \(2019\)](#) argue that time-varying risk aversion can explain changes in VRP. Instead of using VRP, [Cheng \(2019\)](#) considers the VIX premium, which is defined as the difference between the VIX futures index and the expected VIX forecasted by ARMA. The author shows that the VIX premium predicts future returns.

## B. Term Structure of Equity Returns

The term structure of equity returns literature is relatively new due to data availability. [Van Binsbergen et al. \(2012\)](#) first document the different returns of dividend strips, showing that the unconditional term structure of equity returns is downward sloping. They use S&P 500 options data (Long-Term Equity Anticipation Securities, LEAPS) from 1996 to 2009. Proprietary datasets help researchers explore dividend strip returns with longer horizons, hence the term structure. Using private data from investment banks, [Van Binsbergen and Koijen \(2017\)](#) show that short dividend strips have smaller betas but higher returns and Sharpe ratios.

Recent literature pays much attention to the cyclicity of equity return term structure. [Van Binsbergen et al. \(2013\)](#) decompose forward dividend yields, which are closely related to expected hold-to-maturity dividend strip returns. They show that the term structure of equity returns is pro-cyclical, while the term structure of expected dividend growth is counter-cyclical. [Bansal et al. \(2021\)](#) develop a regime-switching model with Bayesian methods, showing that the term structure of equity returns is pro-cyclical. [Golez and Jackwerth \(2020\)](#) extend the observations of Van Binsbergen et al. (2012) to a more extended period (1995–2019). They find that the 2-year dividend strip outperforms the index in recessions and underperforms in normal periods. [Chabi-Yo and Loudis \(2020\)](#) focus on shorter horizons (less than one year), showing that the short-end term structure is pro-cyclical. On the contrary, [Gormsen \(2021\)](#) argues that the term structure is counter-cyclical, with the business cycle measured by equity yield. [Gonçalves \(2021a\)](#) develops an ICAPM model to support this finding.

Many papers try to build theoretical models to explain the term structure of equity returns. Based on the long-run risk model proposed by [Bansal and Yaron \(2004\)](#), [Ai et al. \(2018\)](#) introduce a general equilibrium framework to model the dynamics of discount factor and cash flow. They show that the model matches some stylised facts and helps to extend the term structure to a longer period. [Andries et al. \(2019\)](#) introduce risk aversions that are

decreasing with the horizon into the long-run risk model. The model implies that liquidity is critical for the term structure slope. [Wu \(2020\)](#) incorporates the dividend recovery feature into the variable disaster model of [Gabaix \(2012\)](#). The model generates a lot of stylised features.

Equity duration is closely related to the term structure of equity returns, but it focuses on stocks rather than dividend strips. Equity duration literature cares about dividends further into the future because equity duration is usually very long. [Weber \(2018\)](#) discovers that short-duration equities have higher returns than long-duration equities, which risk factors cannot explain. [Gonçalves \(2021b\)](#) finds a similar result, showing that short-duration stocks are exposed to expected return variation risk. [Gormsen and Lazarus \(2021\)](#) provide a duration-based explanation for the premia on major equity factors. Some papers apply the concept of duration in different contexts. [Chen \(2020\)](#) proposes that effective equity duration accounts for both cash flow and discount rate changes. [Dechow et al. \(2021\)](#) suggest using the implied equity duration to measure the pandemic shutdown risk.

### 3.3 Volatility Feedback Effect

#### A. Dividend Strip

A stock or stock index can be viewed as a collection of dividend strips with different horizons. The decomposition can be written as an equation,

$$S_t = \sum_{n=1}^{\infty} P_{n,t} = \sum_{n=1}^{\infty} \mathbb{E}_t[M_{t+n}D_{t+n}] \quad (3.1)$$

where  $n$  is the time to maturity,  $t$  is the time indicator,  $P$  is the present value of dividend strips,  $M$  refers to discount factors, and  $D$  represents dividends. By the no-arbitrage condition, the present value should also be the price of the dividend strip. Holding the dividend strips for one period, investors can get returns from the dividend strips, which need not be equal across horizons. Mathematically, the  $n$ -period dividend strip return can be expressed as follows,

$$R_{n,t+1} = \frac{P_{n-1,t+1}}{P_{n,t}} - 1 \quad (3.2)$$

This paper focuses on monthly returns, so one period is defined as a month. The relationship between the horizon and the dividend strip return is called the term structure of equity returns.

The prices of dividend strips cannot be directly observed because dividend strips are contingent assets. Therefore, they must be inferred from derivatives, usually futures or options. This paper focuses on the prices derived from dividend strip futures, which have been traded publicly since 2015. The current price of the dividend strip can be derived with the no-arbitrage condition.

$$P_{n,t} = F_{n,t}(1 + y_{n,t})^{-n} \quad (3.3)$$

where  $F_{n,t}$  is the future price at time  $t$  and  $y_{n,t}$  is the risk-free zero-coupon rate.

The most liquid future contracts are delivered every December, so I interpolate future prices to get finer grids (monthly) futures prices. For example, dividend futures with 2-year maturity in July 2016 will be the combination of futures delivered in December 2017 and December 2018,  $F_{24,t} = \frac{5}{12}F_{17,t} + \frac{7}{12}F_{29,t}$ .

## B. Volatility

The term structure of volatility refers to the relationship between horizon and dividend strip return volatility. The most relevant volatility should be the expected realised variance  $\mathbb{E}_{t+1|t}[RV_{t+1}]$ , which is not directly observable. However, the realised volatility is persistent, and the correlogram indicates an AR(1) structure. Therefore, using the current volatility  $RV_t$  as a proxy of  $\mathbb{E}_{t+1|t}[RV_{t+1}]$  is suitable.<sup>2</sup>

I measure the realised variance as the sum of squared daily log returns within the month with daily observations of future prices. This definition is consistent with the literature (Bollerslev et al. 2009, Dew-Becker et al. 2017), which can be written as

$$RV_{n,t} = \sum_{\tau=1}^{m_t} r_{n,t,\tau}^2 \quad (3.4)$$

where  $m_t$  is the number of trading days in month  $t$ , and  $r_{n,t,\tau}$  is the daily log return on day  $\tau$  in month  $t$ .

---

<sup>2</sup>To generate the expected realised variance, VAR is sometimes used in the literature. This paper shows a strong relationship between  $RV_t$  and  $R_{n,t+1}$ , even without other predictors.

Another popular measurement of volatility is the implied volatility derived from options. As discussed in the literature, the implied volatility behaves differently from the expected realised volatility, and the difference generates the volatility risk premium. Therefore, including the implied volatility of dividend strips in the analysis helps further investigate the role of volatility. However, there is no currently available data for dividend strip options in the US market. [Gormsen et al. \(2021\)](#) collect the illiquid European dividend strip options data for the COVID-19 pandemic period. Their findings are similar to those in this paper with realised volatility.

Nonetheless, the study uses the market's implied volatility in the next section when analysing the cyclicalities. The implied volatility is measured by the VIX index.

## C. Data

Unlike many other papers, this research does not rely on proprietary datasets. Instead, all the data in this paper comes from Bloomberg. For 1- to 5-year S&P 500 dividend strip future prices, the data covers a period ranging from December 2015 to July 2021. For 6- and 7-year dividend strips, the data is available from January 2017 to July 2021. Although the sample period is relatively short, it covers normal times and a disaster, namely the COVID-19 pandemic. I also collect the zero yield curve, S&P 500 index (SPX), S&P 500 dividend index (SPXDIV), and VIX index (VIX) from the same source. The S&P 500 index is available from December 1989, and the VIX index starts in June 1990.

S&P 500 dividend strip futures are traded in CME with a maximum maturity of 10 years. However, dividend futures with very long maturities are rarely traded. Lack of liquidity makes the prices unreliable. As a result, this paper only considers dividend strips with maturities of up to 7 years, consistent with [Van Binsbergen and Koijen \(2017\)](#).

The key variables used in this paper are summarised in Table 3.1. Panel A summarises the market (S&P 500) and dividend strip returns up to 7 years; panel B summarises the realised variance and the VIX index. All numbers are shown in percentages.

## D. Empirical Results

By decomposition of the S&P 500 index, this paper documents the volatility feedback effects of dividend strips. I run the following regressions for dividend strips with horizons

Table 3.1: Descriptive Statistics of Returns and Volatility

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
A. Return					
$R_{market}$	379	0.75	4.21	-16.94	12.68
$R_1$	68	0.65	5.19	-30.27	17.24
$R_2$	68	1.93	5.84	-38.18	20.53
$R_3$	68	2.15	5.67	-33.62	22.57
$R_4$	68	1.14	5.66	-31.07	23.29
$R_5$	68	1.07	4.79	-24.62	16.06
$R_6$	55	1.15	4.57	-18.39	12.84
$R_7$	55	1.03	4.10	-9.87	8.36
B. Realised Variance					
$RV_{market}$	379	0.28	0.58	0.01	7.47
$RV_1$	68	0.07	0.33	0.00	2.06
$RV_2$	68	0.15	0.76	0.00	5.78
$RV_3$	68	0.22	1.14	0.00	9.10
$RV_4$	68	0.23	1.11	0.00	8.91
$RV_5$	68	0.23	0.96	0.01	7.62
$RV_6$	55	0.26	0.81	0.01	5.66
$RV_7$	55	0.27	0.80	0.01	5.54
VIX	373	19.55	7.72	9.51	59.89

All numbers are shown in percentages. 1- to 5-year S&P 500 dividend strips cover December 2015 to July 2021. 6- and 7-year dividend strips cover January 2017 to July 2021. S&P 500 (market) return and realised volatility range from December 1989 to July 2021, and the VIX index ranges from June 1990 to July 2021.

from 1 to 7 years and for the market index.

$$R_{n,t+1} = \beta_{n,0} + \beta_n RV_{n,t} + \epsilon_{t+1} \quad (3.5)$$

where  $n$  refers to 1, 2,  $\dots$ , 7, and *market*.

Table 3.2 shows regression results for 1- to 7-year dividend strips and the market index. As shown in columns (1) to (7), the coefficients of  $RV_n$  for dividend strips are significant for all horizons, even though the sample covers fewer than six years. The  $R^2$  of the regressions are all above 10% except for the horizon of 7 years. Furthermore, all the coefficients are positive, meaning that the past realised variance (and hence the current expected realised variance) demands a risk premium, which is consistent with the literature. (Berger et al. 2020)

Column (8) reports regression (3.5) for the market index. Unlike those for dividend strips, the coefficient of  $RV_n$  is small and insignificant. The  $R^2$  of the regression is virtually zero. The volatility feedback effect is not observed for the S&P 500 index, meaning that the

Table 3.2: Volatility Feedback Effect of S&amp;P 500 Dividend Strips

Dependent Variable: $r_n$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$n =$	1 Year	2 Years	3 Years	4 Years	5 Years	6 Years	7 Years	Market
$RV_n$	6.690*** (0.472)	2.533*** (0.318)	1.822*** (0.209)	2.133*** (0.196)	1.786*** (0.159)	1.771*** (0.226)	0.703*** (0.215)	0.022 (1.021)
$R^2$	0.180	0.108	0.135	0.175	0.128	0.100	0.019	0.000
$N$	67	67	67	67	67	54	54	378

Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

realised variance of the market does not demand a risk premium. Note that this regression covers a long period ranging from 1990 to 2021, much longer than the other regressions. Section 3.6 analyses the impact of changing regression periods.

The betas of dividend strip regressions generally decrease with the horizon, meaning that the volatility feedback effect is stronger for short-duration dividends than for long-duration ones. As a result, short-end dividends demand higher volatility risk premiums. Equities usually have very long durations, meaning short-end dividends constitute only an arguably small fraction of equity value. The decreasing volatility feedback effect can explain the small beta for the regression of the market index.

The magnitudes of coefficients are economically significant. Table 14 in the appendix shows the same regression with standardised variables. For the dividend strip with a 1-year horizon, a one standard deviation increase in the realised volatility predicts a 0.424 standard deviation higher return in the next period. For dividend strips with 2- to 7-year maturities, a one standard deviation increase is associated with 0.328, 0.367, 0.417, 0.357, 0.314, and 0.137 standard deviation increase in returns of the subsequent period, respectively. The standardised volatility feedback effect for the market index is 0.003, virtually zero. The decreasing pattern of the volatility feedback effect is preserved even after standardising variables, although it is less clear than using original variables.

This paper shows that the volatility feedback effect for the S&P 500 cannot be picked up by past realised volatility. On the contrary, the volatility feedback effect is significantly present in short-to-medium-term dividend strips, which only constitute a small fraction of equity value.

## E. Expected and Unexpected Volatility

Decomposing realised volatility into expected and unexpected parts helps identify the source of monotonic predictive power. Realised volatility is persistent, which means past realised volatilities can predict the current. Thus, the simplest way to decompose the realised volatility is to run autoregressions. Correlograms suggest that the optimal lag is 1, so this research uses the AR(1) model to decompose expected and unexpected volatilities as follows.

$$RV_{n,t+1} = \beta_0 + \beta_1 RV_{n,t} + \eta_{t+1} \quad (3.6)$$

The expected part in month  $t+1$  is  $RV_{t+1}^e = \hat{\beta}_1 RV_{n,t}$ , and the unexpected part is  $RV_{t+1}^{ue} = \hat{\eta}_{t+1}$ . The hat sign ( $\hat{\cdot}$ ) means estimates derived from regression (3.6). Then, the study runs the following regression with expected and unexpected realised volatilities as explanatory variables.

$$R_{n,t+1} = \beta_{n,0} + \beta_{n,1} RV_{n,t}^e + \beta_{n,2} RV_{n,t}^{ue} + \epsilon_{n,t+1} \quad (3.7)$$

Results for dividend strips are reported in columns (1) to (7) of Table 3.3. The first row provides slope coefficients of the expected volatility  $RV_n^e$ . Column (1) is the regression with a 1-year horizon. The coefficient of expected volatility is significantly positive. However, the coefficient drops to below 0 and becomes insignificant for the horizon of 2 years. It keeps decreasing until the 5-year horizon. The coefficients for 6- and 7-year horizons are also negative but are slightly higher than that of the 5-year horizon. All the coefficients of  $RV_n^e$  are significant except for the 2-year-horizon one.

Table 3.3: Expectation of Volatility and Equity Return Term Structure

Dependent Variable: $r_n$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
n=	1 Year	2 Years	3 Years	4 Years	5 Years	6 Years	7 Years	Market
$RV_n^e$	7.789*** (1.044)	-0.148 (0.930)	-3.199*** (0.597)	-4.311*** (0.576)	-5.080*** (0.537)	-4.751*** (0.518)	-3.859*** (0.635)	0.189 (1.164)
$RV_n^{ue}$	6.247*** (0.824)	2.905*** (0.246)	2.166*** (0.097)	2.504*** (0.072)	2.142*** (0.100)	2.317*** (0.126)	1.321*** (0.152)	-0.030 (1.077)
$R^2$	0.268	0.123	0.205	0.268	0.227	0.216	0.130	0.000
$N$	66	66	66	66	66	53	53	377

Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

There is a monotonically decreasing pattern of the  $RV_n^e$  coefficients. In the short end, a higher realised volatility predicts higher future returns, but with the horizon increasing, a higher realised volatility predicts lower future returns. The finding suggests that in the medium horizon, the expected volatility is "good volatility", but in the short end (1 year), the expected volatility is "bad volatility". This result complements the literature on the dichotomy behaviour of volatilities—signed volatilities based on price movement directions (Bollerslev et al. 2020) and implied and realised volatilities based on measurement methods (Dew-Becker et al. 2021).

The coefficients of  $RV_n^e$  are also economically significant. Table 15 in the appendix reports the same regressions with standardised variables. A one standard deviation increase in 1-year dividend realised volatility predicts 0.493 standard deviations higher returns in the next month. The impacts of one standard deviation increase for 2- to 7-year horizons are -0.019, -0.644, -0.843, -1.015, -0.842, and -0.754, respectively. The downside impact of medium-term volatility is higher than the upside impact of 1-year volatility. Notably, the monotonic pattern preserves after standardisation, meaning that the different levels of dividend strip volatilities do not drive the result.

The second row in Table 3.3 reports coefficients of the unexpected volatility. Unexpected volatility is generally considered a negative shock, so it is unsurprising that coefficients in all seven columns are significantly positive. The higher the unexpected volatility, the higher the required return in the next period.

Unexpected volatility also contributes to the decreasing pattern of coefficients. In the short end, higher unexpected volatility drives up the required returns more than in the long end. The result is not surprising because the cumulative unexpected volatility tends to 0 with time going by. A shock of long-duration volatility is likely to reverse in the following years. Therefore, long-duration returns should be less sensitive to shocks of unexpected volatility.

The impact of unexpected volatility is also economically significant. A one standard deviation increase in 1-year dividend realised volatility predicts 0.396 standard deviations higher returns in the next month. The impacts of one standard deviation increase for 2- to 7-year horizons are 0.376, 0.436, 0.490, 0.428, 0.411, and 0.258, respectively.

Again, the market index behaves differently from dividend strips. Column (8) of Table 3.3 reports the regression result. For the market index, neither the coefficient of the

expected volatility nor that of unexpected volatility is significant. Both coefficients are very small, and the  $R^2$  is virtually zero. Decomposing the realised volatility does not help discover the volatility feedback effect for the market. Column (8) of Table 15 reports standardised coefficients of the regression, which are very close to 0.

## F. Liquidity

There is a concern that the liquidity difference across horizons can explain the results in Table 3.2. Typically, dividend futures with long horizons are traded less frequently than those with short horizons. To address this issue, this research follows Amihud (2002) to construct the mean-adjusted illiquidity index and include it in the regressions. Dividend futures are not traded every day, especially for the long-horizon ones. As a result, this study constructs the index based on daily trading volumes instead of monthly trading volumes. As discussed in Amihud (2002), the construction of the illiquidity index is very robust to different horizons.

This paper uses the following method to create the mean-adjusted liquidity index (*ILLIQMA*). First, the study constructs the illiquidity index for every horizon and month based on the absolute price impact.

$$ILLIQ_{n,t} = \sum_{\tau=1}^{m_t} |R_{n,t,\tau}| / VOLD_{n,t} \quad (3.8)$$

where  $n$  refers to the horizon,  $|R_{n,t,\tau}|$  is the absolute value of dividend future return on day  $\tau$  in month  $t$ ,  $m_t$  is the number of trading days in month  $t$ , and  $VOLD_{n,t}$  is the dollar trading volume of dividend futures with horizon  $n$  in month  $t$ . Since dividend futures are delivered every December, the trading volumes are interpolated like the dividend future prices. The illiquidity index is set to the highest in history for the rare situation where no trade occurs in the month. The intuition is that if the market is illiquid, a given trading volume will have a big price impact and vice versa.

The illiquidity index varies a lot over time, so the index should be adjusted to account for common liquidity shocks. As in Amihud (2002), the average illiquidity index is defined as

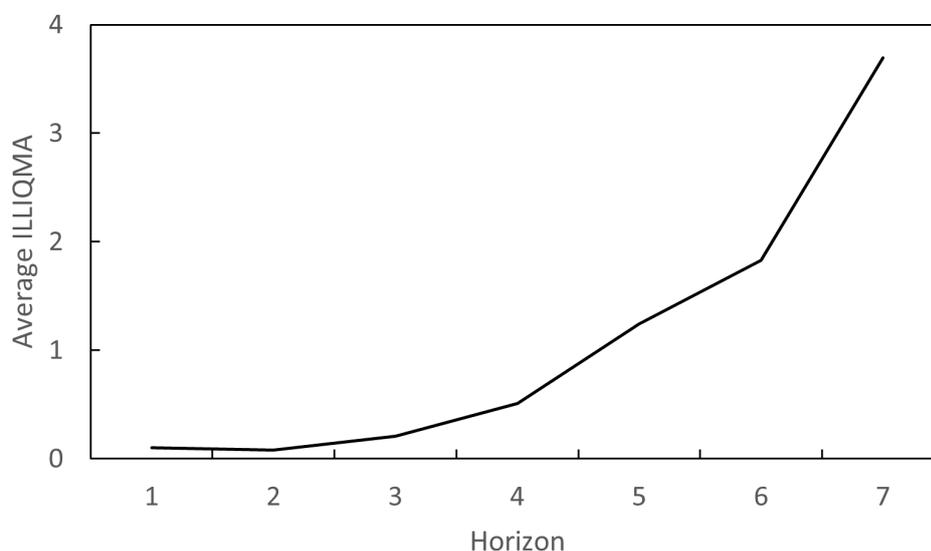
$$AIIIIQ_t = \frac{1}{7} \sum_{n=1}^7 ILLIQ_{n,t} \quad (3.9)$$

and the mean-adjusted illiquidity index is

$$ILLIQMA_{n,t} = ILLIQ_{n,t}/AILLIQ_t \quad (3.10)$$

By definition,  $ILLIQMA$  hovers around 1. As shown in Figure 3.3, illiquidity increases with the horizon, indicating the potential impact of liquidity on the volatility feedback effect. Now, I include  $ILLIQMA_{n,t}$  in regression 3.5 to see whether the decreasing pattern of  $RV$  coefficients is weakened.

Figure 3.3: Average Mean-adjusted Illiquidity Index of Dividend Futures



The figure shows the mean-adjusted illiquidity index (ILLIQMA) of dividend futures average across time. The horizontal axis is the maturity of dividend futures, ranging from 1 to 7 years. This figure uses S&P 500 dividend futures data from December 2015 to July 2021 for 1- to 5-year horizons and from January 2017 to July 2021 for 6- and 7-year horizons.

Regression results are reported in Table 3.4. The table has seven panels showing dividend horizons from 1 to 7 years. The first column of every panel is the same as Table 3.2, and the second column shows the regression of dividend strip return on the lagged mean-adjusted illiquidity index. Although 6 out of 7 regressions have positive slope coefficients, none is statistically significant. In other words, the illiquidity of dividend futures does not predict returns on its own. The third columns of every panel show regressions with both realised volatility and mean-adjusted illiquidity index as independent variables. The coefficients of  $RV_n$  are visually identical to the ones in the first columns, implying that the volatility feedback effect cannot be explained by liquidity. Six  $ILLIQMA_n$  coefficients

out of seven are positive, and those of 1- and 4-year horizons are significantly positive.

Combining all three columns in the same panel, it is clear that the illiquidity index on its own does not predict dividend strip returns. Still, including realised volatility helps discover that illiquidity positively predicts future returns, at least for some horizons. On the contrary, realised volatility is a strong predictor of returns. Controlling for liquidity does not alter the predictive power of realised volatility. In an unreported table, I show that using *ILLIQ* instead of *ILLIQMA* results in almost the same outcome.

### 3.4 Cyclicalities of Equity Return Structure

This section addresses the variation of equity return term structure slope, which is also called cyclicalities. If the term structure is upward-sloping in expansions but downward-sloping in recessions, the term structure of equity returns is pro-cyclical. However, if the relationship is the reverse, the term structure is counter-cyclical.

#### A. Measurement of Equity Return Structure

Cyclicalities are usually measured by the equity term premium (ETP), which is the difference between long-horizon returns and the 1-year dividend strip return.

$$ETP_n = R_n - R_1 \quad (3.11)$$

The market portfolio has a very long duration. This paper also investigates the market term premium to reflect the very long-duration term structure, which is defined by

$$MTP = R_{market} - R_1 \quad (3.12)$$

If the equity and market term premia are positive in expansions but negative in recessions, the term structure of equity returns is pro-cyclical and vice versa.

It is possible to find a measure that can reflect the slope of the term structure. In the term structure of interest rates literature, it is common to use the first two or three principal components to capture most of the term structure's variation. The three principal components are level, slope, and shape. The term structure of equity returns can also be decomposed similarly.

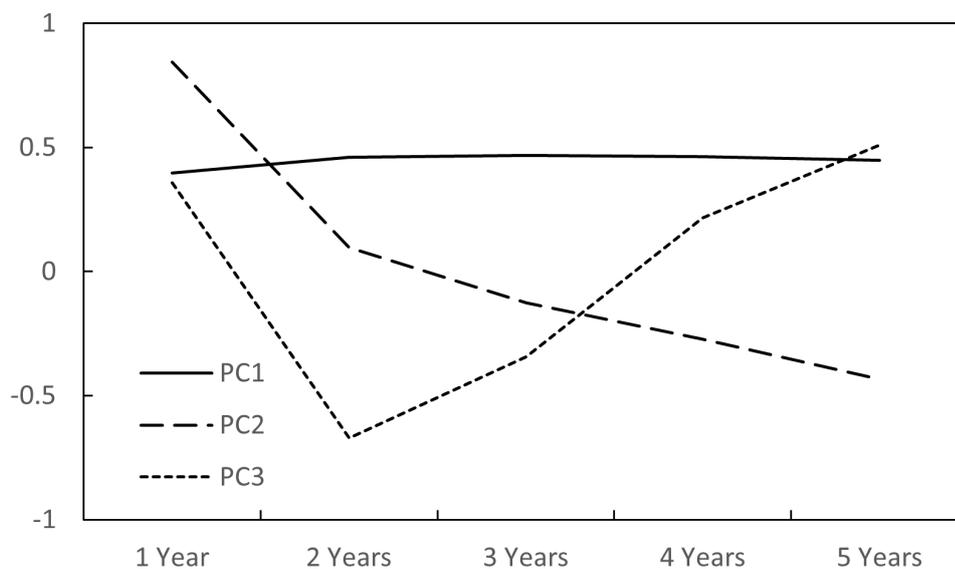
Table 3.4: Volatility Feedback Effect and Liquidity

Dependent Variable: $r_n$		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		1 Year			2 Years		3 Years		4 Years				
$n=$													
$RV_n$		6.690*** (0.472)	6.793*** (0.490)	6.793*** (0.490)	2.533*** (0.318)	2.564*** (0.341)	1.822*** (0.209)	1.848*** (0.224)	2.133*** (0.196)	2.133*** (0.196)	2.200*** (0.207)		
$ILLIQMA_n$			0.010 (0.023)	0.016*** (0.006)	0.000 (0.052)	0.016 (0.021)	0.003 (0.019)	0.009 (0.011)	0.006 (0.009)	0.006 (0.009)	0.009* (0.005)		
$R^2$		0.180	0.003	0.187	0.108	0.109	0.135	0.138	0.175	0.175	0.190		
$N$		67	67	67	67	67	67	67	67	67	67	67	67
Dependent Variable: $r_n$		(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)			
$n=$		5 Years			6 Years		7 Years						
$RV_n$		1.786*** (0.159)	1.809*** (0.177)	1.771*** (0.226)	1.791*** (0.260)	0.703*** (0.215)	0.701*** (0.222)	0.701*** (0.222)					
$ILLIQMA_n$			-0.000 (0.005)	0.001 (0.003)	0.000 (0.005)	-0.001 (0.003)	0.000 (0.002)	0.000 (0.003)	0.000 (0.002)				
$R^2$		0.128	0.000	0.129	0.100	0.100	0.019	0.019	0.000	0.019	0.019		
$N$		67	67	67	54	54	54	54	54	54	54		

Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

This study only decomposes 1- to 5-year dividend strip returns to maximise available observations. The first two principal components explain 97.8% of the total variation, and the first three can explain 99.7%. As shown in Figure 3.4, the loadings of the first principle component are almost flat, which reflects the volatility level. The second principle component has a downward-sloping loading curve, capturing the term structure slope. As expected, the third principal component is associated with the shape.

Figure 3.4: Loadings for Principle Components of Equity Return Term Structure



The figure shows the first three principal components of 1- to 5-year dividend strip returns. This figure uses S&P 500 dividend futures data from December 2015 to July 2021.

The research adds the negative of the second principal component (Average Term Premium,  $ATP = -PC_2$ ) to the left-hand side of regressions to measure the term structure slope.  $ATP$  can be viewed as a combination of equity term premia up to 5 years. It serves as an indicator of overall significance. The term structure of equity returns is pro-cyclical if the slope is positive in good times but negative in bad times. If the relationship is the reverse, the term structure is counter-cyclical.

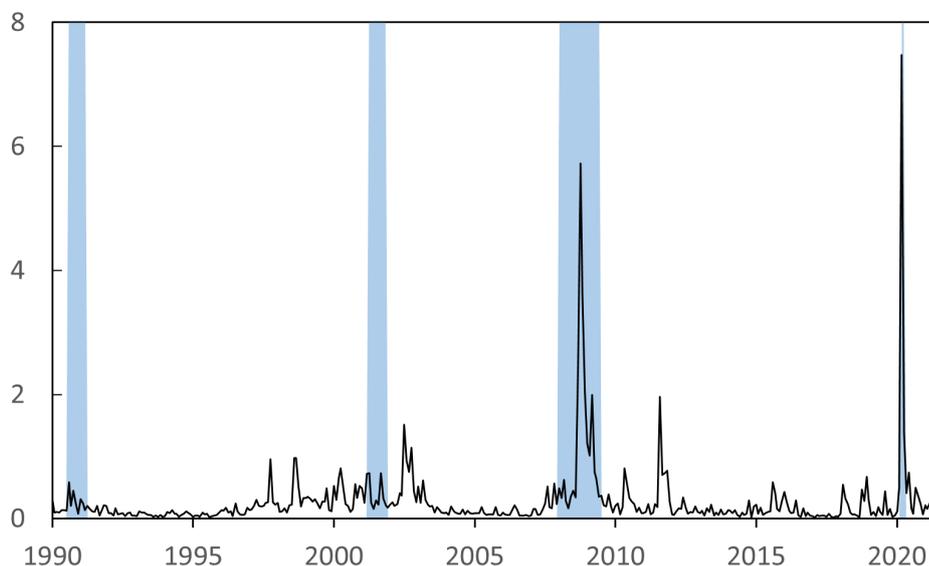
## B. Cyclicalities

There are many ways to define good and bad times in the literature. To accommodate their regime-switching model, [Bansal et al. \(2021\)](#) use NBER recession dates as the indicator of bad times. [Gormsen \(2021\)](#) proposes that the equity yield of the market index can serve as an indicator of good and bad times. [Bloom \(2014\)](#) finds that the VIX index is

"clearly counter-cyclical" and advocates using the stock market volatility as an indicator of cyclicity. Choudhry et al. (2016) also find a close relationship between the stock market volatility and the business cycle across countries.

This paper follows Bloom's suggestion and uses stock market volatilities as indicators of cyclicity. Figure 3.5 depicts the realised volatility of the S&P 500 index and recession periods defined by NBER. There are four recessions between 1990 and 2021, all of which are associated with spikes of realised volatility. The financial crisis (great recession) around 2008 and the COVID-19 pandemic in 2020 witnessed exceptionally high volatility in the stock market. On the other hand, some small spikes of realised volatility are not classified as recessions, including the burst of the dot-com bubble around 2000 and the European sovereign debt crisis between 2010 and 2012. The stock market volatility picks up not only big recessions but also small adverse shocks, making it a good measurement of the business cycle. This section analyses the cyclicity of the equity return term structure, with the realised volatility as the business cycle measurement. The implied volatility will also be discussed.

Figure 3.5: Realised Volatility and NBER Recessions



The shaded areas correspond to recessions defined by NBER. The solid line shows the realised volatility of the S&P 500 index in percentage points.

**Realised Volatility** This paper first uses the realised variance of the S&P 500 index to measure the realised volatility of the overall market. I run the following regressions,

$$Y_{t+1} = \beta_0 + \beta_1 RV_{market,t} + \epsilon_{t+1} \quad (3.13)$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}, MTP_{t+1}$ , and  $ATP_{t+1}$ . Without confusion,  $RV_{market}$  is abbreviated as  $RV$  in the following analysis.

Regression results are shown in Table 3.5. The realised variance predicts equity term premia but not the market return itself. The slope coefficients for 2-, 3-, and 4-year equity term premia are insignificant, but those for 5- to 7-year ETP are significantly negative. For the significant coefficients, the absolute values of  $\beta_1$  increase with the horizon. In addition,  $R^2$  increases with the horizon. The findings show that spreads between mid-term dividend strip returns and the 1-year dividend strip return are negatively associated with the realised variance of the index. In conclusion, the term structure of equity returns is pro-cyclical.

Table 3.5: Realised Volatility and Equity Return Term Structure

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$ETP_2$	$ETP_3$	$ETP_4$	$ETP_5$	$ETP_6$	$ETP_7$	$MTP$	$ATP$
$RV$	-0.031 (0.110)	0.098 (0.169)	0.322 (0.196)	-0.564*** (0.180)	-0.971*** (0.189)	-1.709*** (0.193)	-0.534*** (0.147)	-0.373 (0.436)
$R^2$	0.000	0.001	0.006	0.016	0.045	0.097	0.010	0.011
$N$	68	68	68	68	55	55	68	68

$ETP_n = R_n - R_1$  refers to the term premia of  $n$ -year dividend strips.  $ETP = R_{market} - R_1$  is the term premium of the market.  $ATP$  is the negative of the second principal component of 1- to 5-year dividend strip returns. The table reports regressions of equity term premia on the realised variance of the S&P 500 index. Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

The slope coefficient for the regression of  $MTP$  is also significantly negative, but the magnitude is smaller than  $ETP_7$ . It suggests that the term spread generated by realised variance may reverse in the longer horizon. The coefficient for the regression of  $ATP$  is not significant. It results from incorporating the information of short-term equity term premia, which are not individually significant.

**Implied Volatility** With data for 19 markets, Dew-Becker et al. (2021) find that implied volatility and realised volatility convey very different risk prices. Therefore, the implied

volatility may affect the term structure of equity returns differently.

This paper measures the implied volatility mainly with the VIX index published by CBOE, which is widely used in the literature. Unlike the realised volatility, the implied volatility does not usually appear in theoretical models. There is a concern that some other variables may affect the regression results. Section 3.2 introduces a long literature about VIX premium, which can arguably explain future index returns. I construct the VIX premium (*VIXP*) according to [Cheng \(2019\)](#) and include it as a control variable in the regressions. Another essential variable related to ETP is the current equity yield (*ey*) investigated by [Gormsen \(2021\)](#). I include both variables as control variables in the following regressions,

$$Y_{t+1} = \beta_0 + \beta_1 VIX_t + \mathbf{Z}_t \delta_t \epsilon_{t+1} \quad (3.14)$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}, MTP_{t+1},$  and  $ATP_{t+1}$ , and  $\mathbf{Z}$  refers to the vector of control variables.

Table 3.6 reports the regression results. Unlike the regressions with realised volatility, regressions with implied volatility do not present a monotonic pattern. The point estimates of slope coefficients for equity term premia are negative for all horizons, but only those for 2-, 3-, and 7-year ETP are significant at the 5% level. The coefficients for *MTP* and *ATP* are not significant. The implied volatility picks up both short- and mid-term ETP but not the MTP. In an unreported table, I show that removing the control variables does not change the results qualitatively, but coefficients become slightly larger. The coefficients again indicate that the term structure of equity returns is pro-cyclical.

Table 3.6: Implied Volatility and Equity Return Term Structure

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>ETP</i> <sub>2</sub>	<i>ETP</i> <sub>3</sub>	<i>ETP</i> <sub>4</sub>	<i>ETP</i> <sub>5</sub>	<i>ETP</i> <sub>6</sub>	<i>ETP</i> <sub>7</sub>	<i>MTP</i>	<i>ATP</i>
<i>VIX</i>	-0.115*** (0.030)	-0.127*** (0.046)	-0.068 (0.052)	-0.104 (0.066)	-0.120* (0.064)	-0.158** (0.076)	0.020 (0.055)	-0.081 (0.051)
<i>R</i> <sup>2</sup>	0.069	0.072	0.033	0.051	0.132	0.140	0.033	0.056
<i>N</i>	68	68	68	68	55	55	68	68

$ETP_n = R_n - R_1$  refers to the term premia of  $n$ -year dividend strips.  $ETP = R_{market} - R_1$  is the term premium of the market. *ATP* is the negative of the second principal component of 1- to 5-year dividend strip returns. [Newey and West \(1987\)](#) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

### C. Decomposition of Cyclical

The previous analysis shows that the term structure of equity returns is pro-cyclical, but what drives this cyclical remains unknown. The term structure of equity return volatility provides a new angle on discovering the driving source.

The realised variance of the S&P 500 index reflects the volatility of the whole market, which has a very long duration. It is natural to use  $RV$  to measure the long-duration volatility. The volatility of short-duration assets may complement the result. I use the first principal component of 1- to 5-year dividend strip realised variance ( $PC_1$ ) to measure the short-duration volatility, which captures 96% of the total variation. The advantage of  $PC_1$  is that it captures the short-duration volatility without overweighting idiosyncratic volatilities.

The long-duration volatility  $RV$  and the short-duration one  $PC_1$  summarise the information in the term structure of equity return volatility. The two volatility measurements are highly correlated, with a correlation coefficient of 0.98. It means that they largely co-move. If I regress  $PC_1$  on  $RV$ , the result is

$$PC_{1,t} = -0.194 + 2.047 RV_t \quad (3.15)$$

(0.044) (0.038)

The slope coefficient is significantly positive. One unit increase in the long-duration volatility  $RV$  is associated with about two unit increases in the short-duration volatility  $PC_1$ . In other words, the short-duration volatility is relatively more sensitive than the measurement of the business cycle.

Then I re-run the regressions,

$$Y_{t+1} = \beta_0 + \beta_1 PC_{1,t} + \epsilon_{t+1} \quad (3.16)$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}, MTP_{t+1},$  and  $ATP_{t+1}$ . As  $RV$  and  $PC_1$  are highly correlated, the regression results should not be very different when using either.

The outputs are reported in Panel A of Table 3.7. Not surprisingly, the coefficients for 2- to 4-year  $ETP$  are insignificant, while those for 5- to 7-year  $ETP$  are significantly

negative. The magnitudes of betas are increasing with the horizon. The coefficient for  $MTP$  is significant but is smaller than that of  $ETP_7$  in absolute value. The coefficient for  $ATP$  is significant now, showing that the overall short-term term structure slope is pro-cyclical.

Next, the research puts both long-duration and short-duration information on the right-hand side of regressions,

$$Y_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 PC_{1,t} + \epsilon_{t+1} \quad (3.17)$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}, MTP_{t+1}$ , and  $ATP_{t+1}$ .

Interestingly, when both the realised variance of the S&P 500 and the first principal component are included in regressions, they become significant for all horizons. The regression results are in Panel B of Table 3.7. The coefficients of the long-duration volatility  $RV$  are positive and increase with the horizon. In contrast, the coefficients of short-horizon volatility  $PC_1$  are negative and increase with the horizon in absolute values.  $R^2$  increases with the horizon as well. For  $ETP_1$ , the two variables explain about 6% of the total variation, while for  $ETP_7$ , they explain over 28% of the total variation.

The coefficients for the regression of  $MTP$  are again significant but have smaller magnitudes than  $ETP_7$ , suggesting a reversal effect in the long horizon. The coefficients for the regression of  $ATP$  are also significant.

Compared with Table 3.5, including both variables strongly enhances the explanatory power with significant coefficients. It means that  $RV$  and  $PC_1$  contain different sets of information, even though both volatilities are highly correlated. Collinearity does not result in biases but leads to much bigger standard errors in Panel B than in Panel A. The enhancement of predictive powers (increase in the absolute value of slope coefficients) is more than enough to offset the large standard errors so that the coefficients are significant.

The signs of the coefficients are as expected. Term premia are constructed as long-duration minus short-duration returns. The long-duration volatility picks up the long-duration information of term premia, so term premia should move up in response to a positive volatility shock due to the positive volatility feedback effect reported in Table 3.2. On the contrary, the short-duration volatility picks up the short-duration information of term premia, so the signs of coefficients should be exactly the opposite.

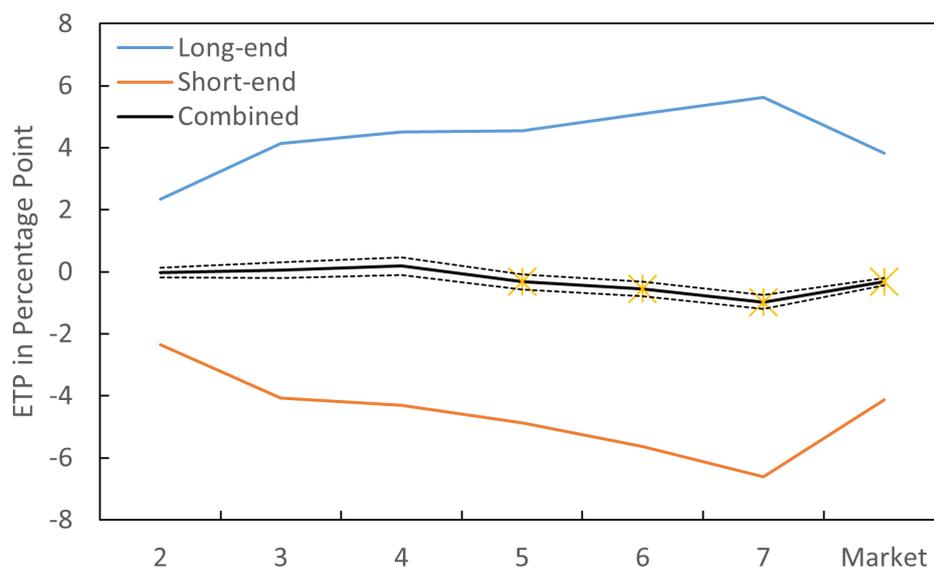
Table 3.7: Realised Volatility and Equity Return Term Structure

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$ETP_2$	$ETP_3$	$ETP_4$	$ETP_5$	$ETP_6$	$ETP_7$	$MTP$	$ATP$
A. Short-Duration Volatility								
$PC_1$	-0.104 (0.091)	-0.106 (0.132)	-0.012 (0.143)	-0.446*** (0.123)	-0.649*** (0.116)	-1.028*** (0.106)	-0.406*** (0.032)	-0.329*** (0.101)
$R^2$	0.004	0.003	0.000	0.044	0.089	0.153	0.011	0.037
$N$	67	67	67	67	55	55	67	67
B. Short-Duration and Index Volatility								
$RV$	4.036** (1.558)	7.147** (2.778)	7.777** (3.494)	7.868** (3.101)	8.796*** (3.217)	9.726*** (3.193)	6.605*** (1.983)	6.763** (2.727)
$PC_1$	-1.989*** (0.743)	-3.445*** (1.291)	-3.645** (1.631)	-4.121*** (1.441)	-4.766*** (1.497)	-5.580*** (1.526)	-3.491*** (0.942)	-3.488*** (1.268)
$R^2$	0.059	0.144	0.142	0.179	0.240	0.282	0.095	0.196
$N$	67	67	67	67	55	55	67	67

$ETP_n = R_n - R_1$  refers to the term premia of  $n$ -year dividend strips.  $ETP = R_{market} - R_1$  is the term premium of the market.  $ATP$  is the negative of the second principal component of 1- to 5-year dividend strip returns. Panel A presents the regressions of equity term premia on the first principle component of 1- to 5-year dividend strip realised variance; panel B presents regressions with both explanatory variables. [Newey and West \(1987\)](#) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

The opposite effects of long- and short-duration volatilities can be visualised in Figure 3.6. The blue line depicts the change of term premia in response to a one standard deviation  $RV$  shock due to the long-duration volatility information. The orange line shows the change of term premia credited to the short-duration volatility information. The dark solid line in the middle shows the combined effect of both types of information, and the two dotted lines mark 95% confidence intervals. Asterisks indicate (negative) significance at the 5% level.

Figure 3.6: Decomposition of Cyclicity



The figure shows the decomposition of one standard deviation shock of  $RV$  to 2- to 7-year ETP and MTP. The orange line depicts the change in the short-end part of ETP and MTP ( $\Delta - R_1$ ), and the blue line depicts the change in the long-end part of ETP and MTP ( $\Delta R_n$ ). Dotted lines represent 95% confidence intervals, and asterisks indicate significance at 5% level.

Both changes of ETP due to long- and short-duration information increase monotonically in absolute values with the horizon, resulting in a monotonic pattern of the combined effect. This result can be explained by the decreasing pattern of volatility feedback effects in Table 3.2. Returns of short-duration dividend strips are more sensitive than those of long-duration ones. When the volatility increases, the short end of the equity return term structure moves up more than the long end. Since ETP measures the difference between long-duration dividend return and the 1-year dividend return, the longer the horizon, the greater the difference. As a result, there is a monotonic pattern in Figure 3.6.

The decreasing pattern of volatility feedback effects also explains the increasing  $R^2$  in Table 3.7. Table 3.1 shows dividend strip returns have roughly the same sizes and standard

deviations across horizons. When ETP moves more (long duration) in response to volatility, the explained variation compared to total variation is higher, so  $R^2$  is higher for the long horizon.

The short-duration volatility has a greater impact than the long-duration one because the short-duration volatility is more sensitive to the business cycle. As a result, the combined effects are negative, indicating a pro-cyclical term structure. The combined effects are almost identical to the results of regression (3.13) in which  $RV$  is the only explanatory variable. As in Table 3.5, the combined effect of the  $MTP$  regression is smaller than that of  $ETP_7$ , showing a reversal effect in the long horizon.

To conclude, long- and short-duration volatilities contain different sets of information, both of which have strong explanatory power on term premia. A decomposition shows that long-duration volatility is associated with the counter-cyclical behaviour of term premia, while the short-duration one is associated with the pro-cyclical one. Since short-duration volatility is more sensitive to the business cycle, the combined effect exhibits a pro-cyclical pattern.

## D. Good Volatility, Bad Volatility

To further delve into the source of term structure cyclicity, I decompose the realised volatilities into "good volatility and bad volatility" based on the direction of price movement. [Patton and Sheppard \(2015\)](#) find that volatility associated with upward price movement ("good volatility") predicts lower future required returns and volatility, while that associated with downward movement ("bad volatility") predicts higher future returns and volatility. This paper follows [Bollerslev et al. \(2020\)](#) to analyse the effect of good and bad volatilities of the market index, which can be represented by the relative signed jump variation ( $RSJ$ ).

"Good volatility" is defined by positive realise semi-variance, and "bad volatility" is defined by negative realise semi-variance. The formulas are very similar to equation (3.4).

$$RV_t^+ = \sum_{\tau=1}^{m_t} r_{t,\tau}^2 \cdot \mathbf{1}[r_{t,\tau} > 0] \quad (3.18)$$

$$RV_t^- = \sum_{\tau=1}^{m_t} r_{t,\tau}^2 \cdot \mathbf{1}[r_{t,\tau} < 0] \quad (3.19)$$

where  $r_{t,\tau}$  is log daily market (S&P 500 index) return in month  $t$ ,  $m_t$  is the number of trading days in month  $t$ , and  $\mathbf{1}[\cdot]$  is an indicative function. Note that  $RV^+$  and  $RV^-$  sum up to  $RV$ .

Then, the relative signed jump variation is defined as

$$RSJ_t = \frac{RV_t^+ - RV_t^-}{RV_t} \quad (3.20)$$

This measure standardises the difference between "good and bad volatilities" by total volatility, removing the overall volatility level.  $RSJ$  is restricted between -1 and 1.

Now, the study runs the following regressions with  $RSJ$  as one of the independent variables,

$$Y_{t+1} = \beta_0 + \beta_3 RSJ_t + \epsilon_{t+1} \quad (3.21)$$

$$Y_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 PC_{1,t} + \beta_3 RSJ_t + \epsilon_{t+1}$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}, MTP_{t+1}$ , and  $ATP_{t+1}$ . The results are listed in Table 3.8. Columns (1) to (6) show regressions with dividend strip term premia as dependent variables. Column (7) uses the market term premium as the dependent variable, and column (8) uses the average term structure slope.

Panel A in Table 3.8 shows the first regression, where the relative signed jump is the only explanatory variable. Coefficients of  $RSJ$  are insignificant in the first six columns, indicating that the short- to medium-duration term structure is not predicted by the relative signed jump on its own. The coefficient for the average slope regression is also insignificant. Interestingly, the market term premium regression coefficient is significantly negative, indicating that high "good volatility" relative to "bad volatility" predicts a downward-sloping term structure in the long end. The magnitude is reasonably large. The coefficient (-0.028) means that compared with an all-bad-news month, an all-good-news month has a 5.6% lower  $MTP$ .

Table 3.8: Relative Signed Jump Variation and Equity Return Term Structure

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$ETP_2$	$ETP_3$	$ETP_4$	$ETP_5$	$ETP_6$	$ETP_7$	$MTP$	$ATP$
A. Relative Signed Jump Variation								
$RSJ$	0.007 (0.008)	0.004 (0.006)	0.003 (0.008)	0.003 (0.011)	-0.002 (0.014)	-0.006 (0.020)	-0.028*** (0.010)	-0.001 (0.008)
$R^2$	0.007	0.002	0.001	0.001	0.000	0.002	0.052	0.000
$N$	68	68	68	68	55	55	68	68
B. $RSJ$ and Volatility								
$RV$	6.530*** (2.000)	10.32*** (3.233)	11.19*** (3.817)	11.06*** (2.947)	11.56*** (2.861)	11.91*** (2.610)	4.655** (1.980)	9.020*** (2.661)
$PC_1$	-3.111*** (0.905)	-4.871*** (1.474)	-5.182*** (1.765)	-5.555*** (1.369)	-6.006*** (1.342)	-6.558*** (1.251)	-2.614*** (0.936)	-4.503*** (1.236)
$RSJ$	0.023 (0.014)	0.029** (0.014)	0.031** (0.013)	0.029** (0.011)	0.026* (0.013)	0.021 (0.018)	-0.018** (0.009)	0.021** (0.008)
$R^2$	0.115	0.217	0.215	0.238	0.277	0.297	0.111	0.242
$N$	67	67	67	67	55	55	67	67

$ETP_n = R_n - R_1$  refers to the term premia of  $n$ -year dividend strips.  $ETP = R_{market} - R_1$  is the term premium of the market.  $ATP$  is the negative of the second principal component of 1- to 5-year dividend strip returns. Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

Panel B in Table 3.8 shows the second regression, in which  $RSJ$ , short-duration volatility, and long-duration volatility are all included as independent variables. The coefficients of  $RV$  and  $PC_1$  follow the same increasing pattern as in Table 3.5, but their magnitudes are bigger than without  $RSJ$ . In other words, the inclusion of relative signed jump variation enhances the previous result. Besides, the coefficients of  $RSJ$  are bigger than in Panel A, except for the regression of market term premium.  $RSJ$  coefficients for 3- to 6-year equity term premia become significantly positive, showing that the short- to medium-duration term structure is positively associated with the difference between "good volatility" relative to "bad volatility". On the contrary, the coefficient is significantly negative for the  $MTP$  regression, again implying that the long-duration term structure is negatively associated with the difference between good and bad news. In the last column (regression of the average term structure), the coefficient of  $RSJ$  is significantly positive.

The most important conclusion drawn from the table is that the increasing pattern of predictability is preserved after incorporating the information on volatility directions.  $RSJ$  cannot explain the monotonic pattern of term structure predictability. The predictive powers of long- and short-duration volatilities become more substantial when controlling the relative signed jump. Notably, the short and long ends of the term structure move in opposite directions in response to the change of  $RSJ$ .  $RSJ$  is not a good indicator of the business cycle because it moves too fast. This finding does not jeopardise the cyclicity of the equity term structure. However, it reminds us that the short and long ends of the term structure do not always move together, so it is crucial to think about the curvature of the equity term structure.

### 3.5 Simulation

So far, this paper has demonstrated that the term structure of equity return volatility explains—(a) the variation of equity return term structure level through volatility feedback effects and (b) the cyclicity of equity return term structure with both long- and short-duration information. This section will test macro-finance models with empirical findings in Section 3.4.

This paper simulates three popular macro-finance models—the preferred habit model (Campbell and Cochrane 1999), the long-run risk model (Bansal and Yaron 2004), and the

rare disaster model (Gabaix 2012). Simulations help identify whether these models fit the empirical cyclical of equity return term structure. Key coefficients to simulate are slope coefficients of the following regressions,

$$Y_{t+1} = \beta_0 + \beta_1 RV_t + \epsilon_{t+1} \quad (3.13)$$

$$Y_{t+1} = \beta_0 + \beta_1 PC_{1,t} + \epsilon_{t+1} \quad (3.16)$$

$$Y_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 PC_{1,t} + \epsilon_{t+1} \quad (3.17)$$

where  $Y_{t+1} = ETP_{2,t+1}, ETP_{2,t+1}, \dots, ETP_{7,t+1}$ .

## A. Preferred habit Model

The preferred habit model is initially developed by Campbell and Cochrane (1999), which argues that making the agent's preference dependent on surplus consumption instead of consumption explains the equity premium puzzle. The preference can be written as

$$U = \mathbb{E} \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma} \quad (3.22)$$

where  $\delta$  is the subjective discount factor,  $C_t$  is the consumption level,  $X_t$  is the habit level (reference point), and  $\gamma$  is the coefficient of relative risk aversion.

The consumption and dividend processes are external, with a constant component and a stochastic normal independent and identically distributed (i.i.d.) component of growth rates. Price-dividend ratios of market and dividend strips depend on the only state variable—surplus consumption ratio ( $S = (C - X)/C$ ). Market and dividend strip returns can be calculated with price-dividend ratios and simulated dividend levels.

This paper follows Wachter (2005) and Van Binsbergen and Koijen (2017) to simulate the model. Particularly, the model is simulated with daily rather than monthly frequency to construct realised volatility  $RV_n$  from the simulation. This paper uses parameters given in the original Campbell and Cochrane (1999) paper but adjusts them to daily frequency.

## B. Long-run Risk Model

Bansal and Yaron (2004) propose the long-run risk model, which uses Epstein and Zin (1991) preference and a stochastic process of volatility. A system of equations characterises

the model setup,

$$\begin{aligned} x_{t+1} &= \rho x_t + \phi_e \sigma_t e_{t+1} \\ g_{t+1} &= \mu + x_t + \sigma_t \eta_{t+1} \\ g_{d,t+1} &= \mu_d + \phi x_t + \phi_d \sigma_t u_{t+1} \\ \sigma_{t+1}^2 &= \sigma^2 + \nu_1 (\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1} \end{aligned} \tag{3.23}$$

where  $\rho$  is the persistence coefficient of the latent process  $\{x_t\}$ ,  $\phi_e$  and  $\phi_d$  are constants that scale the overall volatility  $\sigma_t$ .  $g$  is the consumption growth rate, and  $\mu$  is its constant part.  $g_d$  is the dividend growth rate, and  $\mu_d$  is its constant part.  $\sigma^2$  is the equilibrium level of overall volatility.  $e_t$ ,  $\eta_t$ ,  $u_t$ , and  $w_t$  are i.i.d. standard normal distributed shocks.

The latent process  $\{x_t\}$  and the volatility process  $\{\sigma_t\}$  determine price-dividend ratios of the market and dividend strips. Market and dividend strip returns can be calculated with price-dividend ratios and simulated dividend levels.

I follow [Beeler and Campbell \(2009\)](#) and [Van Binsbergen and Koijen \(2017\)](#) to simulate the model with daily frequency. The parameters are identical to the original Bansal and Yaron paper but adjusted to daily frequency accordingly. The coefficient of relative risk aversion  $\gamma$  is chosen to be 10.

### C. Rare Disaster Model

There is a long literature on the rare disaster. A recent improvement of the rare disaster model is made by [Gabaix \(2012\)](#), which introduces cross-sectional settings and explains ten famous puzzles in financial economics. The model considers a small probability of big disasters and a time-varying resilience. These features result in a decent level of equity premium and price variation with a small coefficient of relative risk aversion.

Resilience is a state variable closely associated with equity prices, which measures how well equity performs when a disaster arrives. Resilience follows the process below,

$$H_t = H^* + \hat{H}_t \tag{3.24}$$

$$\hat{H}_{t+1} = \frac{1 + H^*}{1 + H_t} e^{-\phi_H \hat{H}_t} \hat{H}_t + \epsilon_{t+1}^H \tag{3.25}$$

where  $H^*$  is the long-run level of resilience,  $\hat{H}$  is the variable part of resilience, and  $\phi_H$  is

the persistence of the process. Resilience follows a twisted AR(1) process.

The price of the market index depends on the resilience,

$$P_t = \frac{D_t}{\delta_i} \left( 1 + \frac{\hat{H}_t}{\delta_i + \phi_H} \right) \quad (3.26)$$

where  $\delta_i = \delta - g_d - \log(H^*)$  is the discount factor adjusted by dividend growth rate and equilibrium resilience.

The following formula gives the dividend strip price,

$$P_{n,t} = D_t e^{-n\delta_i} \left( 1 + \frac{1 - e^{-n\phi_H}}{\delta_i} \hat{H}_t \right) \quad (3.27)$$

where  $n = 1, 2, 3, 4, 5, 6,$  and  $7$  years.

I follow the original [Gabaix \(2012\)](#) to simulate the model, using daily rather than monthly frequency. Parameters are adjusted accordingly.

## D. Simulation Results

I simulate all three models 1000 times with a length of 68 months and report the median and 95% confidence interval (CI) for all regressions mentioned above. Simulation and empirical results are reported in Table 3.9.

The data column is the same as in Tables 3.5 and 3.7. Panel A presents the regressions of equity term premia on the realised variance of the market index. Regarding the median simulated coefficients, both the rare disaster model and the long-run risk model produce coefficients close to 0. In contrast, the preferred habit model produces a monotonically increasing pattern of coefficients contrary to the empirical finding. All three models have 95% confidence intervals covering 0, meaning they fail to reproduce the empirical results significantly. However, since the confidence interval is vast, it is also impossible to reject the null hypothesis that the models are wrong, except for the rare disaster model with a 7-year horizon.

Panel B presents the regressions of equity term premia on the first principle component of 1- to 5-year dividend strip realised variance. Again, the rare disaster model produces median coefficients close to 0, but the confidence intervals are wide enough to cover point estimates from the data. Judging from the median simulation, both the preferred habit

Table 3.9: Realised Volatility and Equity Return Term Structure

Panel A. Long-Duration Volatility										
$n$	Data	Rare Disaster Model			Preferred Habit Model			Long-run Risk Model		
		Median	95% CI		Median	95% CI		Median	95% CI	
2	-0.031	0.002	-0.221	0.231	0.243	-0.565	0.974	0.022	-1.531	1.567
3	0.098	0.004	-0.469	0.490	0.493	-1.164	1.961	0.041	-3.817	3.833
4	0.322	0.006	-0.734	0.783	0.755	-1.797	2.980	0.086	-6.422	6.427
5	-0.564	0.009	-1.019	1.111	1.029	-2.467	4.031	0.100	-9.254	9.265
6	-0.971	0.011	-1.331	1.487	1.313	-3.178	5.093	0.074	-12.05	11.85
7	-1.709	0.015	-1.645	1.926	1.599	-3.934	6.193	0.009	-14.76	14.53
Panel B. Short-Duration Volatility										
$n$	Data	Rare Disaster Model			Preferred Habit Model			Long-run Risk Model		
		Median	95% CI		Median	95% CI		Median	95% CI	
2	-0.104	0.000	-0.201	0.167	0.899	-4.553	6.189	0.605	-7.250	8.431
3	-0.106	0.000	-0.428	0.356	1.828	-9.148	12.59	1.166	-17.95	20.48
4	-0.012	0.001	-0.681	0.568	2.784	-13.86	19.25	1.734	-30.83	33.70
5	-0.446	0.001	-0.965	0.808	3.794	-19.01	26.29	2.394	-44.01	47.75
6	-0.649	0.001	-1.281	1.076	4.828	-24.18	33.61	2.907	-57.19	62.01
7	-1.028	0.002	-1.599	1.378	5.908	-29.71	40.90	3.410	-70.11	75.46
Panel C. Both Volatilities										
$n$	Data	Rare Disaster Model			Preferred Habit Model			Long-run Risk Model		
		Median	95% CI		Median	95% CI		Median	95% CI	
Short-Duration Volatility										
2	-1.989	-0.007	-0.440	0.482	-7.029	-52.90	34.48	50.81	-26.07	202.7
3	-3.445	-0.015	-0.937	1.008	-14.15	-107.9	70.90	104.1	-75.21	456.1
4	-3.645	-0.024	-1.497	1.564	-21.20	-165.0	109.5	160.8	-131.1	742.8
5	-4.121	-0.034	-2.127	2.153	-28.37	-224.4	150.4	215.2	-194.6	1038
6	-4.766	-0.046	-2.835	2.834	-36.08	-286.4	193.9	265.7	-257.8	1328
7	-5.580	-0.059	-3.628	3.613	-43.69	-351.3	240.3	310.7	-326.1	1605
Long-Duration Volatility										
2	4.036	0.009	-0.832	0.702	1.152	-5.067	7.441	-10.18	-42.20	6.073
3	7.147	0.018	-1.724	1.493	2.326	-10.42	15.17	-21.50	-97.41	16.87
4	7.777	0.029	-2.695	2.385	3.548	-16.07	23.19	-32.48	-158.9	29.32
5	7.868	0.042	-3.701	3.388	4.782	-22.05	31.54	-43.37	-222.1	44.74
6	8.796	0.055	-4.755	4.515	6.046	-28.41	40.25	-53.54	-284.3	60.56
7	9.726	0.071	-5.850	5.780	7.368	-35.16	49.37	-63.32	-344.1	75.19

This table reports slope coefficients of regressions (3.13), (3.16), and (3.17) with empirical and simulation data.  $n$  refers to the horizon of the dividend strip. The empirical part is identical to Table 3.5. All three models are simulated with a length of 68 months for 1000 times. Median parameters and 95% confidence intervals from simulations are reported in the table. Panel A presents the regressions of equity term premia on the realised variance of the market index; panel B presents the regressions of equity term premia on the first principle component of 1- to 5-year dividend strip realised variance; panel C presents regressions with both explanatory variables.

model and the long-run risk model generate an increasing pattern of coefficients, which is opposite to the finding in the data. Again, the confidence intervals are vast and cover 0. It is also impossible to reject the null hypothesis that the models are wrong, but they do not fit the data well.

Panel C presents regressions with both explanatory variables. Again, the rare disaster model produces median coefficients close to 0, but the confidence intervals can no longer cover the point estimates. Therefore, the null hypothesis that the rare disaster model is correct is rejected at the 5% significance level for all horizons. The preferred habit model generates median coefficients with the same pattern as the empirical data, but the confidence intervals are much higher than in Panels A and B and cover 0. There is not enough evidence to say the model fits the data well. However, the model cannot be rejected due to the wide confidence interval. In terms of median coefficients, the long-run risk model produces an opposite pattern to the data and wide confidence intervals. Again, there is no evidence to either prove that the model fits the data or reject the model.

In conclusion, none of the three popular models fit the empirical finding well. The rare disaster model predicts that the cyclicity of equity return term structure does not depend on volatility. The preferred habit model and the long-run risk model produce such wide confidence intervals that no evidence of fitting the monotonic pattern in data is presented. Judged from the regression with both long- and short-duration volatilities, the rare disaster model is rejected at the 5% level for all horizons. However, the other two models are not rejected because of their wide confidence intervals.

### 3.6 Robustness Check

This section includes several robustness checks to ensure that the sample length and the volatility specification do not drive the results. This section considers several alternatives for realised and implied volatilities.

## A. Sample Length

The first exercise investigates whether the realised and implied volatilities are linked with S&P 500 returns. I run the following regressions,

$$R_{market,t+1} = \beta_0 + \beta_{market}RV_t + \epsilon_{t+1} \quad (3.28)$$

$$R_{market,t+1} = \delta_0 + \delta_{market}VIX_t + \eta_{t+1} \quad (3.29)$$

The full sample results are reported in columns (1) and (4) of Table 3.10. The  $R^2$  in both regressions are close to 0, while none of the slope coefficients is significant. Since the dividend strip data is only available from December 2015, I present the sub-sample analysis with post-2016 observations in columns (2) and (5). Both slope coefficients are significant in the sub-sample. However, as shown in Figure 3.2, the volatilities are incredibly high at the beginning of the COVID-19 pandemic. Therefore, observations in April and May 2021 are not representative and should be considered outliers. I add two dummy variables to eliminate their impact. The regression outcomes are reported in columns (3) and (6), in which  $R^2$  is calculated without the outliers. The  $R^2$  drops to close to 0, and slope coefficients become insignificant. It shows that the explanatory power of index volatility comes from the COVID-19 period only.

Table 3.10: Volatility Feedback Effect of S&P 500 Index

Dependent Variable: $R_{market}$						
	(1)	(2)	(3)	(4)	(5)	(6)
$RV$	0.022 (1.021)	1.625*** (0.093)	1.441 (2.904)			
$VIX$				0.028 (0.048)	0.118** (0.053)	0.014 (0.123)
$R^2$	0.000	0.125	0.004	0.003	0.051	0.001
$N$	378	67	67	373	67	67

Columns (1) and (4) show regressions with full samples; columns (2) and (5) use the sub-sample from January 2016 to July 2021; columns (3) and (6) also use the sub-sample from January 2016 to July 2021, but include two dummies to eliminate periods at the beginning of Covid-19 pandemic crisis (April and May 2020). [Newey and West \(1987\)](#) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

The two outliers discussed above are not a concern of the results for dividend strips in Table 3.2. Unlike those for the stock index, the results for dividend strips are robust to the COVID-19 period. The results without the turbulent periods are qualitatively similar

in an unreported table, although the coefficients are less significant.

## B. Forward Equity Yield

Van Binsbergen et al. (2013) suggest that the log forward equity yield can be decomposed as the difference between hold-to-maturity risk premium and expected dividend growth.

$$fy_{n,t} = n\theta_{n,t} - ng_{n,t} \quad (3.30)$$

where  $fy_{n,t} = \ln(D_t/F_{n,t})$  refers to the (log) forward equity yield,  $\theta_{n,t}$  is the average hold-to-maturity risk premium, and  $g_{n,t}$  is the average expected log dividend growth rate.  $D_t$  is measured as the total S&P 500 dividend in the past calendar year.

Annual dividend only fluctuates mildly in a given month, so the variation of  $fy$  can serve as a proxy of the realised volatility of dividend strip returns. The volatility is defined by

$$V_{n,t} = \text{Var}[fy_{n,t}] \quad (3.31)$$

where  $n$  ranges from 1 to 7 years.

There is no hold-to-maturity return for the S&P 500 index. Following the logic of equation (3.31), the volatility of (log) equity yield  $ey$  can be viewed as a proxy for the expected realised volatility, which is defined by

$$V_{market,t} = \text{Var}[ey_t] \quad (3.32)$$

The correlation coefficients of realised volatilities are reported in Table 3.11, where  $V$  represents the respective  $V_n$  defined in equation (3.31) and  $V_{market}$  defined by equation (3.32). All correlation correlations are above 90% in the last column, indicating that the proxies are highly correlated with realised variance.

Like Section 3.4.2, this study estimates the first principle component of  $V_1$  to  $V_5$  ( $PC_1$ ) as the measure of short-duration volatility and uses  $V_{market}$  as a measure of long-duration volatility. The results are reported in Table 3.12.

Panel A reports the regressions of ETP on the volatility of equity yield. The slope coefficients are significant only for 6- and 7-year ETP. Both coefficients are negative, consistent with the findings in Table 3.5. Panel B shows the regressions of ETP on the short-duration

Table 3.11: Correlation Coefficients of Realised Volatilities

	$RV_1$	$RV_2$	$RV_3$	$RV_4$	$RV_5$	$RV_6$	$RV_7$	$V$
$RV_1$								0.915
$RV_2$	0.947							0.978
$RV_3$	0.904	0.993						0.998
$RV_4$	0.888	0.987	0.998					0.998
$RV_5$	0.853	0.960	0.974	0.983				0.982
$RV_6$	0.850	0.944	0.952	0.963	0.993			0.962
$RV_7$	0.876	0.954	0.955	0.963	0.989	0.997		0.946
$RV$	0.840	0.964	0.984	0.986	0.971	0.950	0.948	0.900

$V$  refers to proxies of the respective realised volatility defined by equations (3.31) and (3.32).

volatility of forward yields. The coefficients for 5- to 7-year equity term premia and the market term premium are significantly negative. The results indicate that the term structure of equity returns is pro-cyclical, as in Section 3.4.

For the significant coefficients in both panels, the absolute values of the coefficients increase with the horizon. The coefficient for  $MTP$  is smaller than the one for  $ETP_7$  in absolute terms. The results are the same as in Sections 3.4.2 and 3.4.3.

Panel C reports the regressions with both long- and short-duration volatilities. The results are similar but less significant than the results in Table 3.7. The coefficients are significant for columns (4) to (8). As expected, among the significant coefficients, coefficients for  $V_{market}$  are positive, and those for  $PC_1$  are negative. The magnitudes of slope coefficients and  $R^2$  follow the same monotonic pattern as in the previous analysis.

To conclude, changing the realised volatility measurement does not alter the result qualitatively, although the coefficients are generally less significant. This outcome is primarily because forward yield measures the hold-to-maturity return, an imperfect proxy for the one-period return.

## C. VIX Square

The VIX index measures the standard deviation of risk-neutral expected volatility rather than the variance, which differs from the definition of  $RV$ . The implied risk-neutral variance can be easily obtained by the following formula,

$$VIX^2 = (VIX)^2/100 \quad (3.33)$$

Table 3.12: (Forward) Equity Yield Volatility and Equity Return Term Structure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable:	$ETP_2$	$ETP_3$	$ETP_4$	$ETP_5$	$ETP_6$	$ETP_7$	$MTP$	$ATP$
A. Long-Duration Volatility								
$V_{market}$	-1.039 (1.227)	0.151 (1.637)	1.745 (1.813)	-3.225 (2.173)	-5.128** (2.451)	-9.146*** (3.016)	-1.374 (2.553)	-2.086 (1.614)
$R^2$	0.002	0.000	0.003	0.011	0.026	0.057	0.001	0.007
$N$	68	68	68	68	55	55	68	68
B. Short-Duration Volatility								
$PC_1$	-0.115 (0.170)	-0.035 (0.242)	0.193 (0.269)	-0.844*** (0.235)	-1.316*** (0.221)	-2.213*** (0.208)	-0.872*** (0.071)	-0.600 (0.494)
$R^2$	0.001	0.001	0.002	0.028	0.065	0.127	0.021	0.022
$N$	67	67	67	67	55	55	67	67
C. Short- and Long-Duration Volatilities								
$V_{market}$	-5.029 (18.21)	4.290 (13.15)	6.996 (8.904)	22.08** (9.202)	34.88** (16.06)	51.90* (27.66)	47.26** (20.67)	18.29** (8.818)
$PC_1$	0.668 (2.909)	-0.703 (2.118)	-0.896 (1.393)	-4.282*** (1.380)	-6.738*** (2.444)	-10.28** (4.271)	-8.230** (3.227)	-3.449** (1.333)
$R^2$	0.004	0.002	0.006	0.067	0.157	0.268	0.151	0.065
$N$	67	67	67	67	55	55	67	67

$ETP_n = R_n - R_1$  refers to the term premia of  $n$ -year dividend strips.  $MTP = R_{market} - R_1$  is the term premium of the market.  $ATP$  is the negative of the second principle component of 1- to 5-year dividend strip returns. Panel A presents the regressions of equity term premia on the realised variance of equity yield; panel B presents the regressions of equity term premia on the first principle component of 1- to 5-year forward yield realised variance; panel C presents regressions with both explanatory variables. Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

The denominator 100 in the equation exists because both  $VIX$  and  $VIX^2$  are in percentage. I replace  $VIX$  with  $VIX^2$  in equation (3.14) and re-run the regressions.

Table 3.13 shows the regression results. The slope coefficients for 2-, 3-, 6-, and 7-year ETP are significantly negative, similar to the findings in Table 3.6. The only difference is that the coefficient for  $ETP_6$  regression moves from marginally significant to significant at the conventional level. The  $R^2$  are also similar. The coefficient for  $MTP$  is insignificant, indicating that  $VIX^2$  has no impact on the long-duration premium. The coefficient for  $ATP$  is insignificant, as in Table 3.6.

In conclusion, the alternative measure of implied volatility does not change the previous results. The term structure of equity returns is still pro-cyclical with the alternative measure. The outcome should not be a surprise since  $VIX$  and  $VIX^2$  are highly correlated, with a correlation coefficient of 97%.

Table 3.13: VIX Square and Equity Return Term Structure

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$ETP_2$	$ETP_3$	$ETP_4$	$ETP_5$	$ETP_6$	$ETP_7$	$MTP$	$ATP$
$VIX^2$	-0.216*** (0.038)	-0.234*** (0.067)	-0.108 (0.095)	-0.221* (0.123)	-0.268** (0.125)	-0.383** (0.158)	-0.014 (0.123)	-0.170 (0.106)
$R^2$	0.062	0.063	0.027	0.056	0.144	0.168	0.032	0.060
$N$	68	68	68	68	55	55	68	68

Newey and West (1987) standard errors with 12 lags are reported in parentheses. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

### 3.7 Conclusion

This paper documents a series of findings related to the term structures of equity returns and the term structure of equity return volatility. First of all, realised volatility can positively predict the dividend strip return. The volatility feedback effect of the dividend strip decreases with the horizon, while that of the stock index is insignificant. Second, both expected and unexpected realised volatilities contribute to the monotonic pattern of the volatility feedback effect. Expected volatility positively predicts 1-year dividend strip returns and negatively predicts medium-term dividend strip returns. Third, realised and implied volatilities are negatively associated with equity term premiums, showing that the equity term structure is pro-cyclical. With different sets of information, long-duration and short-duration volatilities explain the term structure slope of equity returns. The predictive power is more substantial for the long than the short horizon, but it reverses in the very long horizon. Furthermore, the decomposition of cyclicalities shows that the pro-cyclical term structure is due to the high relative sensitivity of short-duration volatility. Finally, simulations show that the rare disaster model should be rejected by the data. Although not rejected by the test, the preferred habit and long-run risk models do not have sufficient evidence to support the empirical discovery of cyclicalities.

This paper contributes mainly to two strands of literature. The first one is the volatility feedback effect. To the best of my knowledge, this is the first paper investigating the volatility feedback effect of dividend strips and mapping out the term structure. The paper also contributes to the equity return term structure literature by showing that the term structure of equity returns is pro-cyclical. The empirical findings support Bansal et al. (2021) and Ai et al. (2018). Notably, the empirical finding of cyclicalities provides a brand

new tool to test macro-finance models. The rare disaster model fails to match the impacts of long- and short-duration volatilities.

Follow-up research can make potential improvements in the following areas. First, this paper only considers a short sample from 2015 to 2021. It would be interesting to include extended data from the past with proprietary datasets. Second, later research can use more sophisticated methods to construct the expected future realised volatility, potentially with VAR. Finally, if the data for long-horizon variance swap is available, one can construct risk-neutral implied variance and variance premium to investigate the term structure of implied volatility.

## References

- Ai, H., Croce, M. M., Diercks, A. M. and Li, K. (2018), ‘News shocks and the production-based term structure of equity returns’, *The Review of Financial Studies* **31**(7), 2423–2467.
- Amihud, Y. (2002), ‘Illiquidity and stock returns: cross-section and time-series effects’, *Journal of financial markets* **5**(1), 31–56.
- Andries, M., Eisenbach, T. M. and Schmalz, M. C. (2019), ‘Horizon-dependent risk aversion and the timing and pricing of uncertainty’, *FRB of New York Staff Report* .
- Andries, M., Eisenbach, T. M., Schmalz, M. C. and Wang, Y. (2015), ‘The term structure of the price of variance risk’, *FRB of New York Staff Report* .
- Bansal, R., Kiku, D., Shaliastovich, I. and Yaron, A. (2014), ‘Volatility, the macroeconomy, and asset prices’, *The Journal of Finance* **69**(6), 2471–2511.
- Bansal, R., Miller, S., Song, D. and Yaron, A. (2021), ‘The term structure of equity risk premia’, *Journal of Financial Economics* .
- Bansal, R. and Yaron, A. (2004), ‘Risks for the long run: A potential resolution of asset pricing puzzles’, *The journal of Finance* **59**(4), 1481–1509.
- Beeler, J. and Campbell, J. Y. (2009), The long-run risks model and aggregate asset prices: An empirical assessment, Technical report, National Bureau of Economic Research.

- Bekaert, G., Engstrom, E. C. and Xu, N. R. (2019), ‘The time variation in risk appetite and uncertainty’, *National Bureau of Economic Research Working Paper Series* .
- Bekaert, G. and Hoerova, M. (2014), ‘The vix, the variance premium and stock market volatility’, *Journal of econometrics* **183**(2), 181–192.
- Berger, D., Dew-Becker, I. and Giglio, S. (2020), ‘Uncertainty shocks as second-moment news shocks’, *The Review of Economic Studies* **87**(1), 40–76.
- Bloom, N. (2014), ‘Fluctuations in uncertainty’, *Journal of Economic Perspectives* **28**(2), 153–76.
- Bollerslev, T., Li, S. Z. and Zhao, B. (2020), ‘Good volatility, bad volatility, and the cross section of stock returns’, *Journal of Financial and Quantitative Analysis* **55**(3), 751–781.
- Bollerslev, T., Tauchen, G. and Zhou, H. (2009), ‘Expected stock returns and variance risk premia’, *The Review of Financial Studies* **22**(11), 4463–4492.
- Bollerslev, T., Xu, L. and Zhou, H. (2015), ‘Stock return and cash flow predictability: The role of volatility risk’, *Journal of Econometrics* **187**(2), 458–471.
- Campbell, J. Y. and Cochrane, J. H. (1999), ‘By force of habit: A consumption-based explanation of aggregate stock market behavior’, *Journal of political Economy* **107**(2), 205–251.
- Campbell, J. Y., Giglio, S., Polk, C. and Turley, R. (2018), ‘An intertemporal capm with stochastic volatility’, *Journal of Financial Economics* **128**(2), 207–233.
- Campbell, J. Y. and Hentschel, L. (1992), ‘No news is good news: An asymmetric model of changing volatility in stock returns’, *Journal of financial Economics* **31**(3), 281–318.
- Chabi-Yo, F., Gonçalves, A. and Loudis, J. (2021), ‘An intertemporal risk factor model’, *Kenan Institute of Private Enterprise Research Paper Forthcoming* .
- Chabi-Yo, F. and Loudis, J. (2020), ‘The conditional expected market return’, *Journal of Financial Economics* **137**(3), 752–786.
- Chen, Z. (2020), ‘Inferring stock duration around fomic surprises: Estimates and implications’, *Journal of Financial and Quantitative Analysis* pp. 1–35.

- Cheng, I.-H. (2019), ‘The vix premium’, *The Review of Financial Studies* **32**(1), 180–227.
- Choudhry, T., Papadimitriou, F. I. and Shabi, S. (2016), ‘Stock market volatility and business cycle: Evidence from linear and nonlinear causality tests’, *Journal of Banking & Finance* **66**, 89–101.
- Dechow, P. M., Erhard, R. D., Sloan, R. G. and SOLIMAN, M. T. (2021), ‘Implied equity duration: A measure of pandemic shutdown risk’, *Journal of Accounting Research* **59**(1), 243–281.
- Dew-Becker, I., Giglio, S. and Kelly, B. (2021), ‘Hedging macroeconomic and financial uncertainty and volatility’, *Journal of Financial Economics* .
- Dew-Becker, I., Giglio, S., Le, A. and Rodriguez, M. (2017), ‘The price of variance risk’, *Journal of Financial Economics* **123**(2), 225–250.
- Drechsler, I. and Yaron, A. (2011), ‘What’s vol got to do with it’, *The Review of Financial Studies* **24**(1), 1–45.
- Epstein, L. G. and Zin, S. E. (1991), ‘Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis’, *Journal of political Economy* **99**(2), 263–286.
- French, K. R., Schwert, G. W. and Stambaugh, R. F. (1987), ‘Expected stock returns and volatility’, *Journal of financial Economics* **19**(1), 3–29.
- Gabaix, X. (2012), ‘Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance’, *The Quarterly journal of economics* **127**(2), 645–700.
- Giglio, S. and Kelly, B. (2018), ‘Excess volatility: Beyond discount rates’, *The Quarterly Journal of Economics* **133**(1), 71–127.
- Golez, B. and Jackwerth, J. C. (2020), ‘Quarter-of-a-century of dividend pricing’, *Available at SSRN 3575450* .
- Gonçalves, A. S. (2021a), ‘Reinvestment risk and the equity term structure’, *The Journal of Finance* .
- Gonçalves, A. S. (2021b), ‘The short duration premium’, *Journal of Financial Economics* .

- Gormsen, N. J. (2021), ‘Time variation of the equity term structure’, *The Journal of Finance* **76**(4), 1959–1999.
- Gormsen, N. J., Kojien, R. S. and Martin, I. W. (2021), Implied dividend volatility and expected growth, in ‘AEA Papers and Proceedings’, Vol. 111, pp. 361–65.
- Gormsen, N. J. and Lazarus, E. (2021), ‘Duration-driven returns’, *Available at SSRN 3359027*.
- Herskovic, B., Kelly, B., Lustig, H. and Van Nieuwerburgh, S. (2016), ‘The common factor in idiosyncratic volatility: Quantitative asset pricing implications’, *Journal of Financial Economics* **119**(2), 249–283.
- Martin, I. (2017), ‘What is the expected return on the market?’, *The Quarterly Journal of Economics* **132**(1), 367–433.
- Martin, I. W. and Wagner, C. (2019), ‘What is the expected return on a stock?’, *The Journal of Finance* **74**(4), 1887–1929.
- Newey, W. K. and West, K. D. (1987), ‘Hypothesis testing with efficient method of moments estimation’, *International Economic Review* pp. 777–787.
- Patton, A. J. and Sheppard, K. (2015), ‘Good volatility, bad volatility: Signed jumps and the persistence of volatility’, *Review of Economics and Statistics* **97**(3), 683–697.
- Pindyck, R. S. (1984), ‘Risk, inflation, and the stock market’, *The American Economic Review* **74**(3), 335–351.
- Van Binsbergen, J., Brandt, M. and Kojien, R. (2012), ‘On the timing and pricing of dividends’, *American Economic Review* **102**(4), 1596–1618.
- Van Binsbergen, J. H. and Kojien, R. S. (2017), ‘The term structure of returns: Facts and theory’, *Journal of Financial Economics* **124**(1), 1–21.
- Van Binsbergen, J., Hueskes, W., Kojien, R. and Vrugt, E. (2013), ‘Equity yields’, *Journal of Financial Economics* **110**(3), 503–519.
- Vasquez, A. (2017), ‘Equity volatility term structures and the cross section of option returns’, *Journal of Financial and Quantitative Analysis* **52**(6), 2727–2754.

Wachter, J. A. (2005), ‘Solving models with external habit’, *Finance Research Letters* **2**(4), 210–226.

Weber, M. (2018), ‘Cash flow duration and the term structure of equity returns’, *Journal of Financial Economics* **128**(3), 486–503.

Wu, D. (2020), ‘A disaster explanation of equity term structures’, *Available at SSRN 3571921* .

## Appendix

### A. Standardised Regressions

Table 14: Standardised Volatility Feedback Effect

Dependent Variable: $r_n$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
n=	1 Year	2 Years	3 Years	4 Years	5 Years	6 Years	7 Years	Market
$RV_n$	0.424*** (0.030)	0.328*** (0.041)	0.367*** (0.042)	0.417*** (0.038)	0.357*** (0.032)	0.314*** (0.040)	0.137*** (0.042)	0.003 (0.140)
$R^2$	0.180	0.108	0.135	0.175	0.128	0.100	0.019	0.000
$N$	67	67	67	67	67	54	54	378

This table presents results of regression (3.5) with standardised variables. [Newey and West \(1987\)](#) standard errors with 12 lags are reported in parentheses. All variables are standardised to have mean 0 and variance 1. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

Table 15: Standardised Volatility Feedback Effect and Expectation

Dependent Variable: $r_n$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
n=	1 Year	2 Years	3 Years	4 Years	5 Years	6 Years	7 Years	Market
$RV_n^e$	0.493*** (0.066)	-0.019 (0.120)	-0.644*** (0.120)	-0.843*** (0.113)	-1.015*** (0.107)	-0.842*** (0.092)	-0.754*** (0.124)	0.026 (0.160)
$RV_n^{ue}$	0.396*** (0.052)	0.376*** (0.032)	0.436*** (0.020)	0.490*** (0.014)	0.428*** (0.020)	0.411*** (0.022)	0.258*** (0.030)	-0.004 (0.148)
$R^2$	0.268	0.123	0.205	0.268	0.227	0.216	0.130	0.000
$N$	66	66	66	66	66	53	53	377

This table presents results of regression (3.7) with standardised variables. [Newey and West \(1987\)](#) standard errors with 12 lags are reported in parentheses. All variables are standardised to have mean 0 and variance 1. \*\*\* represents significance at 1% level; \*\* represents significance at 5% level; and \* represents significance at 10% level.

## B. Details of Simulations

This section provides additional simulation details of the three macro-finance models in Section 3.6—preferred habit model, long-run risk model, and rare disaster model.

**Preferred Habit Model** Surplus consumption ( $S_t = (C_t - X_t)/C_t$ ) is defined as the consumption ( $C_t$ ) relative to the habit level ( $X_t$ ). Its logarithm ( $s_t = \log(S_t)$ ) is the only state variable in the preferred habit model. It follows the process

$$s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t) (\Delta c_{t+1} - \mathbb{E}(\Delta c_{t+1})) \quad (34)$$

where  $\bar{s}$  is the unconditional mean of  $s_t$ ,  $\lambda(s_t)$  is the sensitivity function, and  $\phi$  is the coefficient of persistence. Consumption growth  $\Delta c_{t+1}$  is a random walk process given by

$$\Delta c_{t+1} = g + \nu_{t+1} \quad (35)$$

where  $g$  is the average growth rate, and  $\nu_{t+1}$  follows a normal distribution with mean 0 and variance  $\sigma_\nu^2$ .

The sensitivity function is set to be 0 if  $s_t \leq \bar{s} + (1 - \bar{S}^2)/2$  and otherwise given by the following formula

$$\begin{aligned} \lambda(s_t) &= (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1 \\ \bar{S} &= \exp(\bar{s}) = \sigma_\nu \sqrt{\frac{\gamma}{1 - \phi - b/\gamma}} \end{aligned} \quad (36)$$

where  $\gamma$  is the coefficient of relative risk aversion, and  $b$  is a preference parameter that determines the behaviour of the risk-free rate.

Following [Wachter \(2005\)](#), I define the price-dividend ratio of the  $n$ -period dividend strip as

$$\begin{aligned} F_n(s_t) &= \mathbb{E}_t \left[ \left( \prod_{j=1}^n M_{t+j} \right) \frac{C_{t+n}}{C_t} \right] \\ &= \mathbb{E}_t \left[ M_{t+1} \frac{C_{t+n}}{C_t} F_{n-1}(s_{t+1}) \right] \\ &= \delta e^{(1-\gamma)g - \gamma(1-\phi)(\bar{s} - s_t)} \int_{-\infty}^{\infty} p(\nu) e^{(1-\gamma)\nu - \gamma\lambda(s_t)\nu} F_{n-1}((1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)\nu) d\nu \end{aligned} \quad (37)$$

where  $\delta$  is the subjective discount factor. The price-dividend ratio is only a function of  $s_t$ . I use the fine grid of  $s_t$  (Grid 3) defined in Wachter (2005) to search for values of  $F_n$  at those grid points. Linear interpolation is used to determine the values of  $F_n$  for the rest of the points. The calculation requires doing numerical integration for  $\nu$ . I use 40 points determined by the Gauss-Legendre quadrature to conduct the integration.

The price-dividend ratio of the market is the sum of all dividend strip price-dividend ratios. Technically, I sum price-dividend ratios for dividend strips up to 100 years (25200 daily observations) and set all negative values to 0.

The dividend process is externally given by equation (35), with which prices of dividend strips and the market index can be calculated.

**Long-run Risk Model** There are two state variables in the long-run risk model—the varying component of the expected consumption growth  $x_t$  and the variance of consumption  $\sigma_t^2$ . Their values are given by system (3.23).

The log price-consumption ratio for a consumption claim linearly depends on the two state variables,

$$z_t = A_0 + A_1 x_t + A_2 \sigma_t^2 \quad (38)$$

The dividend process is analysed separately from the consumption process. The log price-dividend ratio for the aggregate stock market is given by

$$z_{m,t} = A_{0,m} + A_{1,m} x_{m,t} + A_{2,m} \sigma_{m,t}^2 \quad (39)$$

The parameters  $A_0$ ,  $A_1$ ,  $A_2$ ,  $A_{0,m}$ ,  $A_{1,m}$ , and  $A_{2,m}$  are functions of  $z_t$  and  $z_{m,t}$  given in Bansal and Yaron (2004). To get specific numbers for these parameters, I solve the following equations for fixed points  $\bar{z}$  and  $\bar{z}_m$ , and insert them back to functions for parameter values.

$$\bar{z} = A_0(\bar{z}) + A_2(\bar{z})\bar{\sigma}^2 \quad (40)$$

$$\bar{z}_m = A_{0,m}(\bar{z}_m) + A_{2,m}(\bar{z}_m)\bar{\sigma}^2 \quad (41)$$

As in Van Binsbergen and Koijen (2017), price dividend ratios of  $n$ -period dividend strips are also linear to the two state variables. The formula is given by

$$F_{n,t} = C_{0,n} + C_{1,n} x_t + C_{2,n} \sigma_t^2 \quad (42)$$

Multipliers in equation (42) are calculated iteratively by

$$C_{0,n} = C_{0,1} + C_{0,n-1} + C_{2,n-1}\bar{\sigma}^2(1-\nu) + \frac{1}{2}C_{2,n-1}^2\sigma_w^2 - C_{2,n-1}\sigma_w^2\lambda_w \quad (43)$$

$$C_{1,n} = C_{1,1} + C_{1,n-1}\rho$$

$$C_{2,n} = C_{2,1} + C_{2,n-1}\nu + \frac{1}{2}C_{1,n-1}^2\phi_e^2 - C_{1,n-1}\lambda_e\phi_e$$

$$C_{0,1} = \theta\log(\delta) - \theta\psi^{-1}\mu_c + \lambda_w^2\sigma_w^2/2 + (\theta-1)[\kappa_0 + (\kappa_1-1)A_0 + \kappa_1A_2\bar{\sigma}^2(1-\nu) + \mu_c] + \mu_d$$

$$C_{1,1} = -\theta\psi^{-1} + (\theta-1)[1 + A_1(\kappa_1\rho - 1)] + \phi$$

$$C_{2,1} = (\theta-1)(\kappa_1\nu - 1)A_2 + (\gamma^2 + \lambda_e^2 + \phi_d^2 + \pi^2)/2 - \gamma\pi$$

where  $\lambda_e = (1-\theta)\kappa_1A_1\phi_e$  and  $\lambda_w = (1-\theta)A_2\kappa_1$ . Other parameters are defined in [Bansal and Yaron \(2004\)](#).

Dividends are determined by system (3.23), with which prices of dividend strips and the market index can be obtained.

**Rare Disaster Model** The rare disaster model proposed by [Gabaix \(2012\)](#) has many asset classes. In the simulation of this paper, only the equity market part is considered. Besides, the simulation does not feature the cross-sectional behaviour of the equity market. Only one equity asset, the market index, is considered in the simulation.

Resilience, defined in equations (3.24) and (3.26), is the key variable to simulate. The error term  $\epsilon_{t+1}^H$  can take many forms and even be heteroskedastic. However, the process  $\hat{H}_t$  must be stable, and prices should be greater than 0. To ensure these properties, the value of  $\epsilon_{t+1}^H$  should be chosen so that  $\hat{H}_t$  is greater than  $\hat{H}_{min} = \max((e^{-\phi_H} - 1)(1 + H^*), -p - H^*)$  at all times. Here,  $p$  is the probability of disaster happening. In this paper,  $\hat{H}_t$  follows a heteroskedastic binary distribution.

$$\epsilon_{t+1}^H = \begin{cases} \frac{1}{2} \left( \hat{H}_{min} - \frac{1+H^*}{1+\hat{H}_t} e^{-\phi_H} \hat{H}_t \right), & \text{with probability } 1-h \\ \frac{h-1}{2h} \left( \hat{H}_{min} - \frac{1+H^*}{1+\hat{H}_t} e^{-\phi_H} \hat{H}_t \right), & \text{with probability } h \end{cases} \quad (44)$$

The value of  $h$  is chosen to be 0.0841 to match the volatility of  $H_t$ .

The dividend process has a constant growth rate with random shocks, with which prices of dividend strips and the market index can be obtained.

# 4

## Short-Duration Equity Return Puzzle

### Abstract

Short-duration dividend strips have high conditional Sharpe ratios during crises, far beyond the theoretical upper bound. The finding is called the short-duration equity return puzzle. Using dividend prices and dividend forecasts, this paper constructs the required rate of return and the conditional Sharpe ratio for dividend strips between 2002 and 2021. During the crisis period in the sample, the required rate of return of 1-year dividend peaks at 55% and its conditional Sharpe ratio rises above 14, 2 to 17 times higher than the theoretical upper bounds predicted by mainstream macrofinance models. Returns and Sharpe ratios for other horizons are also too high to be explained. The finding is robust to measurement errors and transaction costs.

**Keywords:** Dividend Strip; Conditional Sharpe Ratio; Stochastic Discount Factor

## 4.1 Introduction

Equity refers to the residual interest of a firm held by shareholders. In the financial market, it is usually traded in the form of stock or stock index. The value of equity can be viewed as a combination of dividends with different durations, provided that there is no bubble in the long run. Dividends are relatively low compared with equity. The average dividend yield of the S&P 500 index between 1991 and 2021 is only 1.95%, so most of the equity value is contained in dividends with long durations.

Traditionally, financial economists focus on topics related to equity itself because stocks and stock indexes are widely traded in the market. Dividends are not well investigated due to data availability. Using S&P 500 options data, [Van Binsbergen et al. \(2012\)](#) first document the different returns of dividend strips. They find that returns and Sharpe ratios of holding one- and two-year dividend strips are much higher than those of the S&P 500 index. Some other papers also find that the dividend return is decreasing with duration. ([Cassella et al. 2021](#), [Van Binsbergen and Koijen 2017](#))

The high return and Sharpe ratio of short-duration dividends are somewhat puzzling because, intuitively, short-duration dividends should have lower risks than the stock index. Many companies provide forward guidance on the time and level of future dividends. They tend to avoid deviating from the guidance due to the signalling effect ([Woolridge 1983](#)), which makes the short-duration dividend similar to a debt. Even if companies do not provide forward guidance, dividend announcements can still be predicted with high precision. ([Kalay and Loewenstein 1986](#), [DeAngelo et al. 2000](#))

However, there are oppositions to the "puzzle" because of some drawbacks of the current finding. Most papers in the literature measure dividend returns with realised returns, which are inherently backwards-looking. The high return of dividend strips could merely be a result of luck. More importantly, the term structure of dividend strip returns is cyclical ([Van Binsbergen et al. 2013](#), [Gormsen 2021](#)), which makes the average return vary with sample periods ([Bansal et al. 2021](#)).

This paper intends to sharpen the "short-duration equity return puzzle" by asking the following question: *Can mainstream macrofinance models explain conditional Sharpe ratios of dividend strips over time?* The research question avoids the drawbacks because it does not rely on the average value, and the conditional Sharpe ratio is forward-looking.

This paper considers the conditional Sharpe ratio instead of the expected return because dividends and the stock index have different risks.

The research question is interesting because it provides a novel direct test of macro-finance models (conditional Sharpe ratio), in addition to equity premium (Mehra and Prescott 1985), excess volatility (Shiller 1979, LeRoy and Porter 1981), return predictability (Campbell and Shiller 1988), and other puzzles.

This paper focuses on dividend strips of the S&P 500 index. The construction of conditional Sharpe ratios depends on the dividend strip price and the forecasted dividend level. Since 2015, dividend futures are publicly traded in CME. The current price of dividends can be calculated directly from the price of futures. I also use the dividend price data from Van Binsbergen et al. (2012) for the period before 2009. Dividend forecasts come from IBES.

The result shows that the maximum annualised conditional Sharpe ratio of the 1-year dividend strip between 2002 and 2021 is 14, far beyond the theoretical maxima given by the preferred habit model, the long-run risk model, the rare disaster model, and the term structure model. The maximum annualised expected Sharpe ratios of 2- to 4-year dividend strips are also too high to be explained, although smaller than that of the 1-year dividend strip. This finding is robust to transaction costs and the measurement of conditional variance of forecasts.

The next section introduces some related literature. Section 4.3 provides a detailed discussion of the methodology. Next, empirical results are shown in Section 4.4, and the robustness is discussed in Section 4.5. Then, Section 4.6 simulates several mainstream macrofinance models and shows that the short-duration equity return puzzle exists. Finally, Section 4.7 concludes.

## 4.2 Literature Review

The term structure of equity returns measures the relationship between dividend strip return and duration. The concept is relatively new due to data availability. Van Binsbergen et al. (2012) first show that the unconditional term structure of equity returns is downward sloping, using S&P 500 options data (Long-Term Equity Anticipation Securities, LEAPS) ranging from 1996 to 2009. Proprietary datasets help researchers explore dividend strip

returns with longer horizons, hence the term structure. Using private data from investment banks, [Van Binsbergen and Kojien \(2017\)](#) show that short dividend strips have smaller betas but higher returns and Sharpe ratios.

Recent literature pays much attention to the cyclical nature of equity return term structure. [Van Binsbergen et al. \(2013\)](#) decompose forward dividend yields, which are closely related to expected hold-to-maturity dividend strip returns. They show that the term structure of equity returns is upward-sloping in the expansion (pro-cyclical), while the term structure of expected dividend growth is upward-sloping in recession (counter-cyclical). [Bansal et al. \(2021\)](#) develop a regime-switching model with Bayesian methods, showing that the term structure of equity returns is pro-cyclical. [Golez and Jackwerth \(2020\)](#) extend the observations of [Van Binsbergen et al. \(2012\)](#) to a more extended period (1995–2019). They find that the 2-year dividend strip outperforms the index in recessions and underperforms in normal periods. [Chabi-Yo and Loudis \(2020\)](#) focus on shorter horizons (less than one year), showing that the short-end term structure is pro-cyclical. On the contrary, [Gormsen \(2021\)](#) argue that the term structure is counter-cyclical with business cycle measured by equity yield. [Gonçalves \(2021a\)](#) develops an ICAPM model to support this finding.

However, most papers in the literature analyse realised returns instead of required returns, so they are naturally backwards-looking. As a result, the term structure of equity returns is not comparable to the term structure of interest rates.<sup>1</sup> To solve this problem, [Bansal et al. \(2021\)](#) construct the expected equity return by estimating a conditional dividend growth rate with historical data. [Gormsen \(2021\)](#) uses predictive regressions to estimate the expected return with the realised return. [Giglio et al. \(2021\)](#) employ a factor model to estimate the expected return without dividend strip data. These papers do not directly observe the term structure of expected returns.

Many papers try to build theoretical models to explain the term structure of equity returns. Based on the long-run risk model proposed by [Bansal and Yaron \(2004\)](#), [Ai et al. \(2018\)](#) introduce a general equilibrium framework to model the dynamics of discount factor and cash flow. They show that the model matches some stylised facts and helps to extend the term structure to a longer period. [Andries et al. \(2019\)](#) introduce risk aversions that decrease with the horizon into the long-run risk model. The model implies that liquidity is critical for the term structure slope. [Wu \(2020\)](#) incorporates the dividend recovery feature

---

<sup>1</sup>See [Moench and Siavash \(2021\)](#) for a review.

into the variable disaster model of [Gabaix \(2012\)](#). The model generates a lot of stylised features.

Equity duration is closely related to the term structure of equity returns, but it focuses on stocks rather than dividend strips. Equity duration literature cares about dividends further into the future because equity duration is usually very long. [Weber \(2018\)](#) discovers that short-duration equities have higher returns than long-duration equities, which risk factors cannot explain. [Gonçalves \(2021b\)](#) finds a similar result and shows that short-duration stocks are exposed to expected return variation risk. [Gormsen and Lazarus \(2021\)](#) provide a duration-based explanation for the premia on major equity factors. Some papers apply the concept of duration in different contexts. [Chen \(2020\)](#) proposes the effective equity duration to account for both cash flow and discount rate changes. [Dechow et al. \(2021\)](#) suggest using the implied equity duration to measure the pandemic shutdown risk.

### 4.3 Methodology

This section presents the methodology to construct the forward-looking required rate of return for S&P 500 dividend strips and the conditional Sharpe ratio.

#### A. Required Rate of Return

Dividend strips are similar to zero-coupon bonds. A dividend strip pays all its value when it matures, but the final payoff is uncertain. The required rate of return of a dividend strip can be written as

$$R_{n,t} = \left( \mathbb{E}_t \left[ \frac{D_{t+n}}{P_{n,t}} \right] \right)^{\frac{12}{n}} - 1 = \left( \frac{\mathbb{E}_t [D_{t+n}]}{P_{n,t}} \right)^{\frac{12}{n}} - 1 \quad (4.1)$$

where  $n$  represents the duration of the dividend strip,  $t$  indicates the current time,  $D_{t+n}$  is the dividend paid in period  $t + n$ ,  $P_{n,t}$  is the current price of the dividend strip with duration  $n$ , and  $\mathbb{E}_t[\cdot]$  is the expectation operator conditional on the information in period  $t$ .

Equation (4.1) is the equity counterpart of the zero-coupon rate. This paper uses monthly observations in the empirical analysis, so the interval of  $t$  and  $n$  is a month. The returns are annualised so that those with different durations are comparable.

In order to obtain the required rate of return, one must observe the expected dividend and the dividend price.

**Dividend Price** The prices of dividend strips cannot be directly observed because dividend strips are contingent assets. Therefore, they must be inferred from derivatives, usually futures or options. This paper focuses on the prices derived from dividend strip futures, which have been traded publicly in CME since 2015. The current price of the dividend strip can be derived with the no-arbitrage condition.

$$P_{n,t} = F_{n,t}(1 + y_{n,t})^{-n} \quad (4.2)$$

where  $F_{n,t}$  is the future price at time  $t$  and  $y_{n,t}$  is the risk-free zero-coupon rate.

Dividend futures contracts are delivered every December, so I interpolate future prices to get finer grids (monthly) futures prices. For example, dividend futures with 2-year maturity in July 2016 will be the combination of futures delivered in December 2017 and December 2018,  $F_{24,t} = \frac{5}{12}F_{17,t} + \frac{7}{12}F_{29,t}$ .

Before 2015, there was no publicly traded dividend futures. This paper uses dividend prices from [Van Binsbergen et al. \(2012\)](#), which are calculated from the put-call parity of the S&P 500 index options.

$$P_{n,t} = p_{n,t} - c_{n,t} + S_t - Xe^{-ny_{n,t}} \quad (4.3)$$

where  $p_{n,t}$  refers to the price of the S&P 500 put option with maturity  $n$  in period  $t$ ,  $c_{n,t}$  is the price of the S&P 500 call option,  $S_t$  is the S&P 500 index level, and  $X$  is the strike price of given options.

The dividend strip obtained from options is essentially a synthetic asset of a put, a call, a stock index, and a specific amount of cash. Although [Van Binsbergen et al. \(2012\)](#) match put and call options within the same minute and use the median value to estimate the dividend price, measurement errors are unavoidable. [Boguth et al. \(2012\)](#) argue that small measurement errors in base assets can lead to large measurement errors in the synthetic asset. The reason is that  $P$  is much smaller than  $S$ , so there is a leverage effect. [Golez and Jackwerth \(2022\)](#) find that a holding period longer than one year can reduce the measurement error of the Sharpe ratio, but it does not solve the problem entirely.

Dividend futures are not synthetic assets, so the dividend price calculated from equation (4.2) is not subject to the measurement error problem.

**Expectation** The construction of dividend expectations relies on analyst forecasts. To the best of my knowledge, there is no available forecast for the dividend of the S&P 500 index, so this paper uses the bottom-up method to estimate the forecast. IBES provides analyst forecasts of dividends from 2002. The maximum forecast horizon is 5 years, and most companies have dividend forecasts for up to 3 years. Only constituents of the S&P 500 index are included in the sample. This study uses the median (consensus) forecast of a month as the expected dividend.

Dividend forecasts are reported in terms of the fiscal year, not the calendar year. There is a gap between the end of the fiscal year and the ex-date, when the dividend officially becomes the firm's debt. The average gap is one and a half months, so I move the forecast horizon forward for 2 months. This adjustment only has a mild effect on the result.

Not all firms are observed every month. If the forecast is unavailable, this observation is treated as missing instead of replaced by the last observation. It guarantees that all information is up-to-date. To avoid fluctuation due to the changing number of observations, this paper uses the following formula to calculate the expected dividend,

$$\mathbb{E}_t[D_{t+n}] = \frac{\sum_i w_{i,t} \mathbb{E}_t[D_{i,t+n}]}{\sum_i w_{i,t} D_{i,t}} D_t \quad (4.4)$$

where  $D_i$  refers to the 12-month cumulative dividend of a firm, and  $w_i$  is the weight. Since the S&P 500 index is a market-capitalisation-weighted index, this research aggregates the dividends with the number of shares outstanding as the weight.

## B. Conditional Sharpe Ratio

The required rate of return can be calculated with the dividend price and the dividend expectation. To maintain consistency, this paper only uses the excess return in the empirical analysis. The term "return" refers to the excess return thereafter. The conditional variance of the excess return can be written as

$$\mathbb{V}_t[R_{n,t} - R_{f,t}] = \mathbb{V}_t[R_{n,t}] \quad (4.5)$$

where  $R_{f,t}$  is the risk-free rate, a constant in period  $t$ , and  $\mathbb{V}_t[\cdot]$  is the variance operator conditional on the information in period  $t$ .

The variance of dividend expectation cannot be aggregated from the forecasts of individual firms because the correlation structure among forecasts is unknown. Therefore, the conditional variance must be extracted from the aggregate dividend expectation.

This paper uses a multivariate generalised autoregressive conditional heteroskedasticity (MGARCH) model to extract the conditional variance of the aggregate forecasts. There are four time series in the model— $\mathbb{E}_t[D_{t+n}]$  for  $n = 12, 24, 36, \text{ and } 48$ . To keep the model parsimonious, this paper uses the dynamic conditional correlation (DCC) structure developed by Engle (2002).<sup>2</sup> I choose AR(1)-DCC(1,1) model for the empirical analysis. Numbers of lags are suggested by correlograms of the expected dividend series.

Expected dividend increases with time, so a trend term must be included in the estimation. I take the logarithms of the time series first.

$$d_{n,t} = \log(\mathbb{E}_t[D_{t+n}]) \quad (4.6)$$

The level equations are AR(1) with a trend.

$$\begin{aligned} d_{n,t+1} &= \beta_{n,0} + \beta_{n,1}d_{n,t} + \beta_{n,2}t + \epsilon_{n,t+1} \\ \epsilon_t &= H_t^{1/2}z_t \end{aligned} \quad (4.7)$$

where  $\epsilon_t$  is the vector of error terms,  $H_t$  is the conditional variance-covariance matrix, and  $z_t$  is a vector with independent standard normal terms.

The dynamic of  $H_t$  is defined by the following equations,

$$\begin{aligned} H_t &= D_t R_t D_t \\ D_t &= \text{diag}(h_{11,t}^{1/2}, \dots, h_{NN,t}^{1/2}) \\ R_t &= \text{diag}(q_{11,t}^{-1/2}, \dots, q_{NN,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2}, \dots, q_{NN,t}^{-1/2}) \\ Q_{t+1} &= (1 - \alpha - \beta)Q + \alpha\Psi_t + \beta Q_t \\ \Psi_t &= u_t u_t' \end{aligned} \quad (4.8)$$

where each element in  $D_t$  follows a GARCH process,  $Q_t = (q_{ij,t})$  is symmetric positive def-

<sup>2</sup>See Bauwens et al. (2006) for a detailed survey of different MGARCH models.

inite,  $Q$  is the unconditional variance-covariance matrix,  $H_t = (h_{ij,t})$ , and  $u_{n,t} = \epsilon_{n,t}/\sqrt{h_{ii,t}}$  is the scaled residual. When the number of equations  $N$  is 1, the model converges to GARCH. When the adjustment coefficients satisfy  $\alpha + \beta = 0$ , the model reduces to the constant conditional correlation (CCC) structure.

By assuming log-normality of the dividend expectation, one can calculate the variance of excess return with the following formula,

$$\mathbb{V}_t[R_{n,t} - R_{f,t}] = (e^{h_{ii,t}} - 1)e^{2\mu_{n,t} + h_{ii,t}} \quad (4.9)$$

where  $\mu_{n,t}$  is the conditional expectation of  $d_{n,t}$ .

The conditional Sharpe ratio is given by

$$SR_{n,t} = \frac{R_{n,t} - R_{f,t}}{\sqrt{\mathbb{V}_t[R_{n,t} - R_{f,t}]}} = \frac{R_{n,t} - R_{f,t}}{\sigma_t[R_{n,t} - R_{f,t}]} \quad (4.10)$$

where  $\sigma_t[\cdot]$  is the standard deviation operator conditional on the information in period  $t$ . In the following analysis, all Sharpe ratios are annualised so that they are comparable across horizons.

**Maximum conditional Sharpe ratio** This paper will compare the empirical conditional Sharpe ratios of dividend strips with the theoretical maximum conditional Sharpe ratio. Exploiting the duality of the mean-standard deviation frontier for the intertemporal marginal rate of substitution (IMRS) and the mean-standard deviation frontier for the asset payoff, [Hansen and Jagannathan \(1991\)](#) show that the conditional Sharpe ratio for all risk assets must obey the following relationship,

$$SR_{n,t} = \frac{R_{n,t} - R_{f,t}}{\sigma_t[R_{n,t} - R_{f,t}]} = -\rho_t(M_{t+1}, R_{n,t+1} - R_{f,t}) \frac{\sigma_t[M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} \leq \frac{\sigma_t[M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} \quad (4.11)$$

where  $M$  is the stochastic discount factor and  $\rho_t[\cdot]$  is the correlation coefficient operator conditional on the information in period  $t$ .

In other words, the theoretical maximum conditional Sharpe ratio for all risk assets is,

$$SR_{max,t} = \frac{\sigma_t[M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} \quad (4.12)$$

This number can be calculated for theoretical asset pricing models analytically or numerically.

## 4.4 Empirical Analysis

This section discusses the empirical analysis of the required rate of return and the conditional Sharpe ratio for dividend strips.

### A. Data

I collect dividend futures data from Bloomberg for the period between November 2015 and December 2021. The longest horizon of dividend futures is 10 years, but the liquidity is usually insufficient for horizons over 7 years. This paper only uses dividend futures with horizons of up to 5 years. I also collect bid and ask prices of dividend strip futures from FactSet. Dividend prices between January 1996 and October 2009 are obtained from [Van Binsbergen et al. \(2012\)](#)<sup>3</sup>, which are calculated with S&P 500 options. Since the maximum maturity of S&P 500 options is 3 years, only 1- and 2-year horizon dividend prices are available. Note that the data is subject to the measurement error problem.

I also collect the S&P 500 index (*SPX*), the S&P 500 dividend index (*SPXDIV*), S&P 500 constituents lists, and risk-free rates from Bloomberg.

Dividend forecast data comes from IBES, covering the period between May 2002 and December 2021. This paper uses forecasts with forecast periods of 1 to 5 years. Only US companies are included in the sample. The median dividend forecast of a firm in a month is taken as the consensus forecast of the month. Actual dividend data is also extracted from IBES. Outliers are manually verified and picked out. This paper only focuses on regular dividends because special dividends are not included in the *SPXDIV* index.

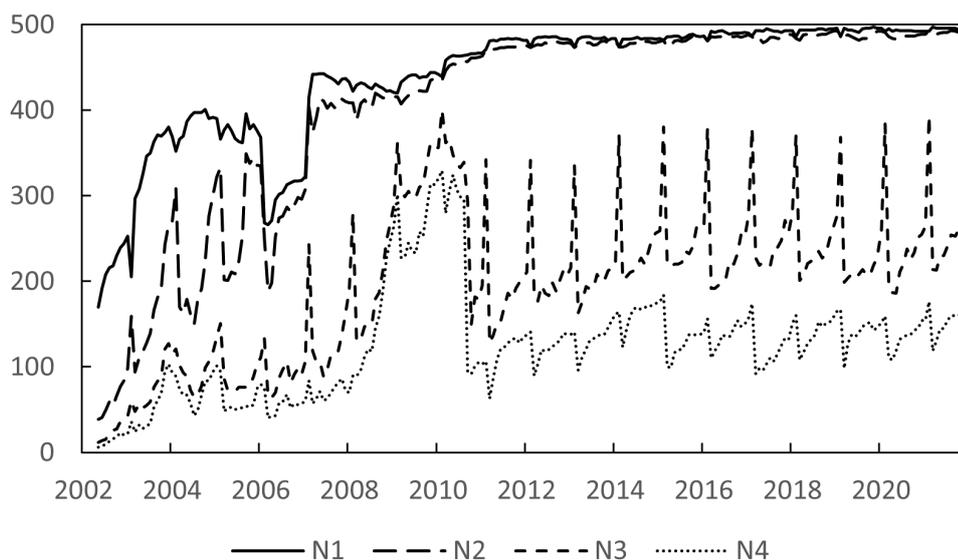
### B. Validity of Expectation

**Sample Coverage** The bottom-up method can produce a reliable aggregate dividend forecast. Figure 4.1 shows the number of S&P 500 constituents covered by the sample. The four time series, N1 to N4, represent forecast periods from 1 to 4 years. At the beginning of the sample, fewer than half of the constituent firms are covered by the dataset, but the

<sup>3</sup>AER website: <https://www.aeaweb.org/articles?id=10.1257/aer.20131416>.

coverage goes up rapidly. About 80% of firms have 1-year dividend forecasts in 2004. Since 2010, over 90% of firms have dividend forecasts for 1- and 2-year horizons. Short horizons always have better coverage than long horizons. Spikes in the time series are due to the change of years.

Figure 4.1: Number of S&P 500 Firms with Observed Dividend Forecast

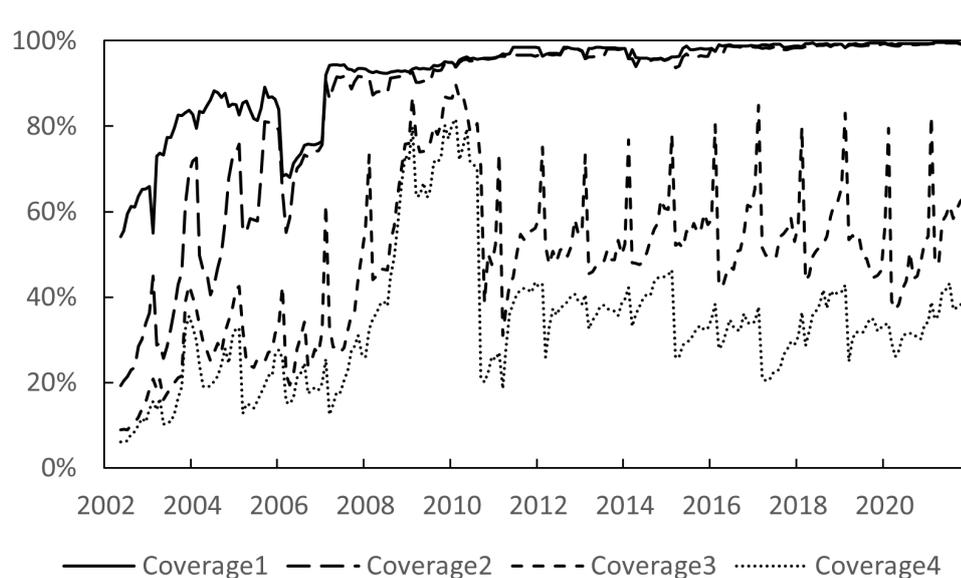


N1 to N4 represent the number of firms with observed dividend forecasts for 1- to 4-year horizons. This figure uses the IBES dividend forecast and actual dividend data between May 2002 and December 2021.

Since this paper uses value weights to aggregate the dividend forecast, the market share covered by the sample is more important than the number of firms. Figure 4.2 depicts the market share of S&P 500 firms with observed dividend expectations. The four time series, Coverage1 to Coverage4, represent forecast periods from 1 to 4 years. This figure is very similar to Figure 4.1, but the market share covered by the sample is higher than the number of firms. It is because analysts tend to make dividend forecasts for big firms. From 2011, the covered market shares for 1- and 2-year forecast periods are close to 100%. The potential measurement error is minimal.

**Forecast Error** The high coverage ratio does not necessarily guarantee that the aggregated IBES dividend forecast is precise because the analysts may not be marginal investors in the market. Following [Bilinski and Bradshaw \(2022\)](#), this paper measures the forecast

Figure 4.2: Market Share of S&amp;P 500 Firms with Observed Dividend Expectation



Coverage1 to Coverage4 represent the market share of S&P 500 firms with observed dividend forecast for 1- to 4-year horizons. This figure uses the IBES dividend forecast and actual dividend data between May 2002 and December 2021.

error with the following formula,

$$FE_{n,t} = \frac{|D_{t+n} - \mathbb{E}_t[D_{t+n}]|}{S_t} \quad (4.13)$$

The forecast error is expressed in absolute terms and scaled by the S&P 500 index. Table 4.1 shows the forecast errors of dividends in the sample. The average forecast error of the 1-year horizon forecast is 0.12%. Those of 2- to 4-year horizons are 0.23%, 0.29%, and 0.37%, respectively. [Bilinski and Bradshaw \(2022\)](#) find that the average forecast error of US stocks with an average duration of less than one year is 0.45%. My finding is comparable to their result and suggests that the dividend forecast of the SP500 index is more precise than individual stocks.

Another way to confirm the validity of the dividend expectation is to show that the actual dividend growth can be predicted by the forecast. The dividend growth rate is defined by the following formula,

$$dg_{n,t+n} = \frac{D_{t+n} - D_t}{D_t} \quad (4.14)$$

where  $D_t$  refers to the 12-month cumulative dividend ending in period  $t$ .

Table 4.1: Forecast Errors of Dividends

Horizon	Obs	Average	SD	Max
1 Year	224	0.12%	0.10%	0.53%
2 Years	212	0.23%	0.16%	0.73%
3 Years	200	0.29%	0.21%	0.78%
4 Years	188	0.37%	0.28%	0.98%

This table shows forecast errors of dividends calculated with equation (4.13). Forecast horizons are in the first column. The second column presents the average forecast error across time. The third and fourth columns show the standard deviation and the maximum value of the forecast error across time. This figure uses aggregated IBES dividend forecast and S&P 500 dividend index data between May 2002 and December 2021.

I run the regressions of dividend growth rates on the expected growth rates for  $n = 12, 24, 36,$  and  $48$ . To correct for the serial correlation, this paper uses [Newey and West \(1987\)](#) standard errors with 12 lags.

$$dg_{n,t+n} = \beta_0 + \beta_1 \mathbb{E}_t[dg_{n,t+n}] + \epsilon_{t+n} \quad (4.15)$$

Regression results are reported in Table 4.2. Columns (1) to (4) correspond to 1- to 4-year forecast periods. If  $\beta_0 = 0$  and  $\beta_1 = 1$ , the expected dividend growth rate predicts the realised growth rate without bias. The slope coefficients for 1- to 3-year horizons are 0.939, 1.015, and 1.060, which are virtually and statistically indifferent from 1. The intercepts for 1- to 3-year horizons are 0.040, 0.048, and 0.030, which are also small and statistically indifferent from 0 at the 5% level. The slope coefficient and the intercept for the 4-year horizon are 0.586 and 0.172. However, the  $p$ -value for the  $F$  test with the null hypothesis  $H_0 : \beta_0 = 0$  and  $\beta_1 = 1$  is 0.193. The "no bias" hypothesis is not rejected at the conventional significance level. The regression results are consistent with [Bilinski and Bradshaw \(2022\)](#), which find that analyst dividend forecasts are more accurate than alternative proxies based on extrapolations of past dividends.

Figure 4.3 presents the time series of realised and expected dividends of the S&P 500 index for the period between 2002 and 2021. The dark line is the realised dividend, and the other four are dividend forecasts with 1- to 4-year horizons. All the dividend forecasts move together and respond faster than the realised dividend. For example, In mid-2020, dividend forecasts drop deeply in response to the COVID-19 pandemic, but the realised dividend only decreases mildly and recovers very soon.

In conclusion, the bottom-up method to derive the aggregate S&P 500 dividend forecast

Table 4.2: Dividend Growth Prediction

Dependent Variable: $dg_n$				
	(1)	(2)	(3)	(4)
$n =$	12	24	36	48
$\mathbb{E}_t[dg_n]$	0.939	1.015	1.060	0.586
	(0.270)	(0.311)	(0.351)	(0.228)
$\beta_0$	0.040	0.048	0.030	0.172
	(0.021)	(0.054)	(0.103)	(0.105)
$p$ -value	0.001	0.194	0.483	0.193
Adjusted $R^2$	0.375	0.198	0.181	0.085
$N$	224	212	200	188

This table reports the results of regression (4.15) for forecast horizons  $n = 12, 24, 36,$  and  $48$  months. The  $p$ -value term is for the  $F$  test with the null hypothesis  $H_0 : \beta_0 = 0$  and  $\beta_1 = 1$ . [Newey and West \(1987\)](#) standard errors are reported in parentheses.

covers most firms among the 500 constituents. These forecasts have small forecast errors and predict the realised dividend growth very well. They are good estimates of market expectations.

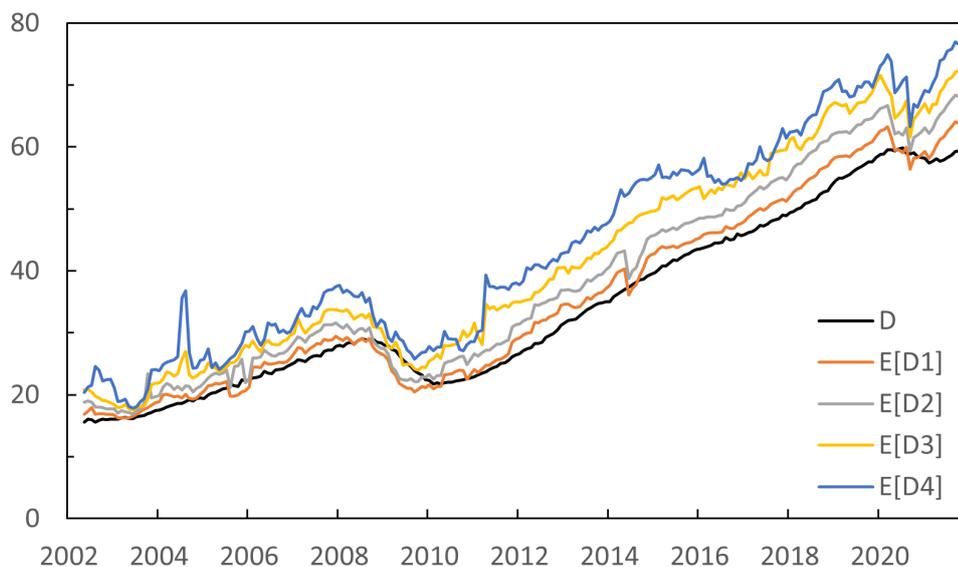
### C. Main Result

The required rate of return of dividend strips can be obtained with equation (4.1). Excess returns are calculated by subtracting risk-free rates of respective horizons. Figure 4.4 reports the excess returns. The left part (2002–2009) of returns uses dividend prices from [Van Binsbergen et al. \(2012\)](#), and the right part (2015–2021) uses dividend futures. Excess returns for 1- and 2-year horizons are available for both parts of the sample, while those for 3- and 4-year horizons are only available in the recent sub-sample.

The first thing to note is that excess returns in the left part are more volatile than in the right part. As shown in Figures 4.1 and 4.2, the coverage ratios are lower at the beginning of the sample than at the end, so there could be relatively big measurement errors for the dividend forecast. More importantly, dividend strips in [Van Binsbergen et al. \(2012\)](#) are synthetic assets, so there are measurement errors inherently.

In both sub-samples, the 2-year horizon dividend has a higher return than the 1-year horizon dividend most of the time. However, in 2002, 2008–2009, and 2020, the relationship is reversed. This finding is consistent with the pro-cyclical term structure of equity returns found by [Bansal et al. \(2021\)](#). More importantly, the required rates of returns for all dividend strips are high in these periods. The return of the 1-year dividend reaches 21%

Figure 4.3: Realised and Expected Dividends of S&amp;P 500 Index



The dark line is the 12-month cumulative realised dividend of the S&P 500 index. The other four time series depict the forecasts of S&P 500 dividends with forecast periods of 1 to 4 years. This figure uses the *SPXDIV* index, dividend forecast and actual dividend data between May 2002 and December 2021.

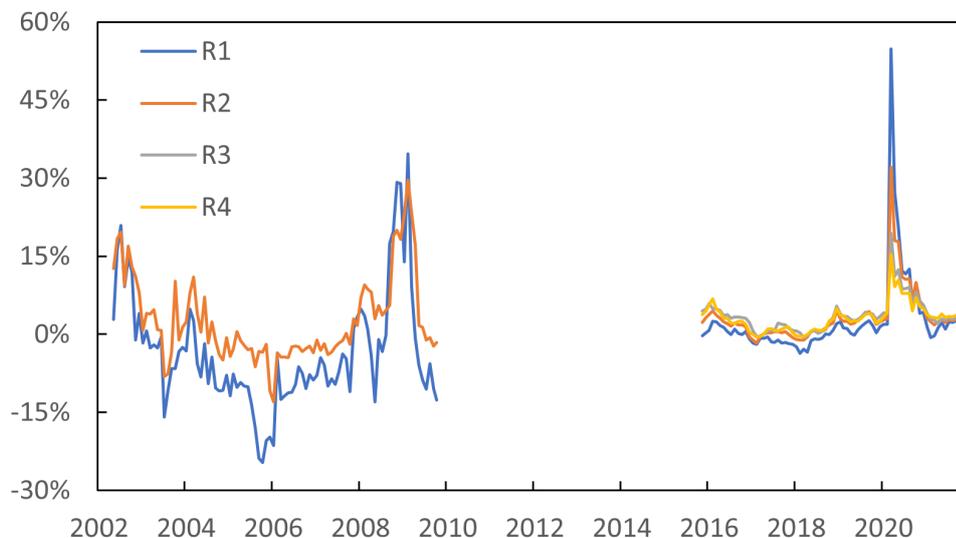
in July 2002, 35% in February 2009, and strikingly, 55% in March 2020. Returns of 2- to 4-year dividend strips in March 2020 are 32%, 19%, and 15%, respectively.

The high required rate of return during crises is puzzling because there is no apparent risk associated with dividend strips. As presented in Section 4.4.2, short-duration dividends can be predicted more precisely than long-duration ones. If a stock index is split into a series of dividend strips, the more predictable part should have less risk. Moreover, many companies provide future dividend guidance and are reluctant to deviate from it, making short-duration dividends similar to debts. (Woolridge 1983) It is hard to reconcile the high return with the low risk.

Some papers argue that there are risks associated with short-duration dividend strips. For example, Ai et al. (2018) argue that short-duration dividends are more volatile than long-duration ones. Wu (2020) provides a model with relatively stable long-duration dividends. Sharpe ratio, measured as the excess return scaled by its standard deviation, serves as a better measure of abnormal return.

First of all, I need to estimate the conditional variance of the forecasted dividend with the AR(1)-DCC(1,1) model described in Section 4.3.2. Regression results are presented in Table 4.3. The regression for the 3-year horizon is not estimated due to collinearity. The

Figure 4.4: Required Rate of Return for S&amp;P 500 Dividend Strips



This figure shows required rates of returns for S&P 500 dividend strips with 1- to 4-year horizons. All returns are excess returns. The left part of returns uses data from [Van Binsbergen et al. \(2012\)](#), and the right part uses dividend futures.

slope coefficients for the AR(1) process are 0.956, 0.954, 0.939, and 0.897 for 1- to 4-year horizons, respectively. These coefficients measure the persistence of dividend forecasts. The results indicate that short-duration dividend forecasts are more persistent than long-duration ones.

The GARCH(1,1) process is not persistent for 1- to 3-year dividend forecasts. The GARCH(1) coefficients for the three series are -0.027, 0.005, and -0.090, either negative or insignificant. The ARCH(1) terms for them are 1.223, 1.308, and 0.308, indicating that the conditional variance is sensitive to the most recent innovation. The 4-year dividend forecast is different. Its ARCH(1) term is small (0.068) but significant, and its GARCH(1) term is reasonably large (0.91) and significant. The conditional variance of the 4-year dividend forecast shows a persistent pattern.

The adjustment coefficients,  $\alpha$  and  $\beta$ , describe the process of the correlation matrix. Both of them are significantly different from 0. The  $p$ -value entry is for the  $\chi^2$  test with the null hypothesis  $H_0 : \alpha + \beta = 0$ . If the null hypothesis holds, the DCC structure degenerates to the CCC structure, in which the correlation matrix is constant. Since the  $p$ -value is 0.000,  $H_0$  should be rejected, and the DCC structure should be used.

Then, Sharpe ratios can be calculated with equations (4.9) and (4.10). Figure 4.5 shows Sharpe Ratios for S&P 500 dividend strips with 1- to 4-year horizons. The shape of this

Table 4.3: Regression Results for Conditional Variance of Dividend Forecast

Dependent Variable: $d_{n,t+1}$				
	(1)	(2)	(3)	(4)
$n=$	12	24	36	48
$d_{n,t}$	0.956 (0.015)	0.945 (0.015)	0.939	0.897 (0.019)
ARCH(1)	1.223 (0.253)	1.308 (0.229)	0.308	0.068 (0.031)
GARCH(1)	-0.027 (0.007)	0.005 (0.012)	-0.090	0.910 (0.030)
$\alpha$	0.010	(0.005)		
$\beta$	0.944	(0.008)		
$p$ -value	0.000			
$N$	235	235	235	235

This table reports the results of multivariate GARCH regression (4.7) and (4.8) for 1- to 4-year log dividend forecasts. The model specification is AR(1)-DCC(1).  $\alpha$  and  $\beta$  are adjustment coefficients. The  $p$ -value is for the  $\chi^2$  test with the null hypothesis  $H_0 : \alpha + \beta = 0$ . Standard errors are reported in parentheses.

figure is very similar to the required rate of return (Figure 4.4). The left part of this figure is more volatile than the right part. There are some extremely negative values in the left part, but none in the right part. This result should be explained by measurement errors.

In both sub-samples, long-duration dividends have higher Sharpe ratios than the short-duration ones most of the time, but the relationship is reversed in 2002, 2008, and 2020. It shows that the Sharpe ratio of the dividend strip is also pro-cyclical, similar to the required rate of return.

The most surprising finding is the extreme Sharpe ratios during crises. In March 2020, Sharpe ratios for the 1- to 4-year dividend strips are 14.08, 11.68, 8.26, and 7.92, respectively. These values are large compared with the average realised Sharpe ratio of the S&P 500 index between April 1990 and December 2021, which is only 0.79. These large Sharpe ratios of dividend strips appear more than once in history. In October 2022, the conditional Sharpe ratios of 1- and 2-year dividend strips are 4.22 and 5.62. In February 2009, these numbers are 9.46 and 10.78. Conditional Sharpe ratios of dividend strips tend to be high during a crisis. Section 4.6 will show that these high conditional Sharpe ratios are far beyond the theoretical upper bound predicted by mainstream models.

Figure 4.5: Conditional Sharpe Ratio for S&amp;P 500 Dividend Strips



This figure shows the conditional Sharpe Ratios for S&P 500 dividend strips with 1- to 4-year horizons. The left part of returns uses data from [Van Binsbergen et al. \(2012\)](#), and the right part uses dividend futures.

## 4.5 Robustness

This section discusses some robustness checks of the empirical analysis. Particularly, it is interesting to know whether the high required rate of return and the high Sharpe ratio are driven by measurement errors and transaction costs.

### A. Measurement Error

Two sources of measurement errors can affect the results—the dividend price and the dividend forecast. To reduce the measurement error from the dividend price, this part only uses the dividend prices derived from dividend futures, so the sample time is restricted to the period between November 2015 and December 2021.

As shown in Figure 4.2, the market capitalisation coverage ratio of dividend forecast varies across time. Since August 2008, the coverage ratios for 1- and 2-year dividend strips are always higher than 90%. To reduce the measurement error from the dividend forecast as much as possible, I use the period from August 2008 to estimate the conditional variance of 1- and 2-year dividend strips only. This paper also considers a more recent period starting in November 2015. The coverage ratios for 1- and 2-year dividend strips are higher than

98% and 96%, respectively. Reducing the sample period has a minimal impact on the result.

The paper also uses AR(1)-DCC(1,1) specification for the recent sample. Results are reported in Table 4.4. The slope coefficients for the AR(1) process are 0.908 and 0.891 for both dividend strips, which are lower than the numbers in Table 4.3. Recent dividend forecasts are not as persistent as in the full sample. Both time series have a more persistent conditional variance process in the short sample. GARCH(1) coefficients are 0.863 and 0.879 for 1- and 2-year dividend strips, significant and much higher than the ones estimated with the full sample. ARCH(1) terms of the two series are 0.155 and 0.121, which are significant but not as large as before. The adjustment coefficients are 0.112 and 0.771 in the short sample, and the  $p$ -value for the  $\chi^2$  test with the null hypothesis  $H_0 : \alpha + \beta = 0$  is virtually 0. As a result, DCC structure instead of CCC structure should be used.

Table 4.4: Regression Results for Conditional Variance of Dividend Forecast

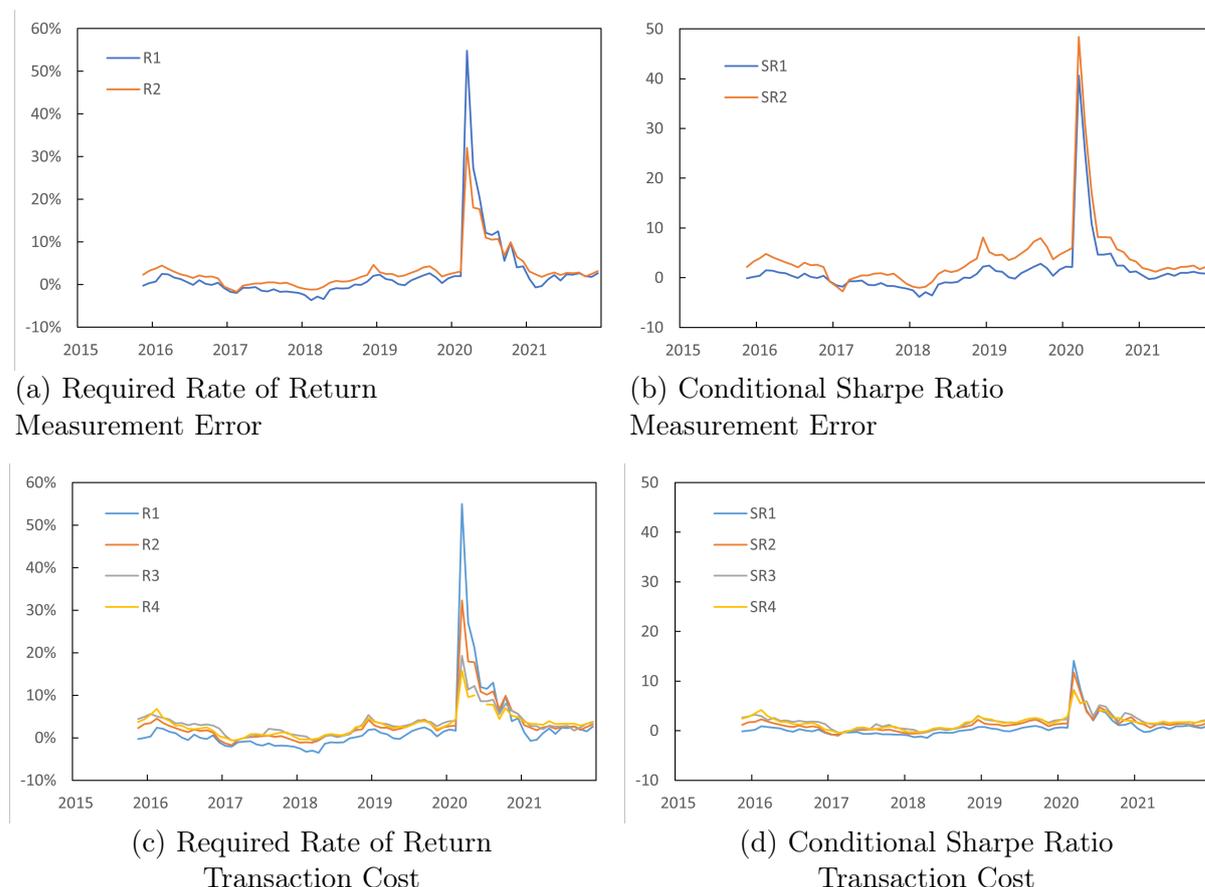
Dependent Variable: $d_{n,t+1}$		
	(1)	(2)
$n =$	12	24
$d_{n,t}$	0.908 (0.017)	0.891 (0.018)
ARCH(1)	0.155 (0.034)	0.121 (0.029)
GARCH(1)	0.863 (0.024)	0.879 (0.023)
$\alpha$	0.112	(0.055)
$\beta$	0.771	(0.058)
$p$ -value	0.000	
$N$	161	161

This table reports the results of multivariate GARCH regression (4.7) and (4.8) for 1- and 2-year log dividend forecasts. The sample period is August 2008 to December 2021. The model specification is AR(1)-DCC(1).  $\alpha$  and  $\beta$  are adjustment coefficients. The  $p$ -value is for the  $\chi^2$  test with the null hypothesis  $H_0 : \alpha + \beta = 0$ . Standard errors are reported in parentheses.

Panels (a) and (b) in Figure 4.6 show the required rate of return and the conditional Sharpe ratio with the recent sub-sample of dividend forecasts. Since the conditional variance of dividend forecasts only affects the Sharpe ratio but not the return, Panel (a) is identical to the right part of Figure 4.4. Panel (b) shows larger extreme Sharpe ratios during a crisis than Figure 4.5. In March 2020, conditional Sharpe ratios for 1- and 2-year dividends are 40.63 and 48.41, much larger than the ones estimated with the full sample. This is because the recent sample has smaller measurement errors and, hence, smaller con-

ditional variance of the estimates. Measurement errors do not result in an upward bias of estimates. On the contrary, if the more recent and precise dividend forecasts are used, extreme Sharpe ratios are larger and more puzzling.

Figure 4.6: Required Rate of Return and Conditional Sharpe Ratio for S&P 500 Dividend Strips



The figures show required rates of returns and conditional Sharpe ratios with different robust settings. Panel (a) shows the required rates of returns for 1- and 2-year dividend strips, identical to the right part of Figure 4.4. Panel (b) is the conditional Sharpe ratio for 1- and 2-year horizon dividend strips, estimated with only dividend forecasts from August 2008. Panel (c) depicts the required rates of returns for 1- to 4-year dividend strips with transaction costs. Panel (d) shows the conditional Sharpe ratio for 1- to 4-year horizon dividend strips with transaction costs.

## B. Transaction Cost

Transaction cost may explain high returns and Sharpe ratios. To determine the impact of the transaction cost, this research collects ask and bid prices of dividend futures from FactSet. The sample also ranges from November 2015 to December 2021. Since the trading strategy involves only a long position of dividend futures, the study replaces the dividend

price with the ask price and re-does the analysis.

The bid-ask spread of dividend futures is small. Using ask prices, the average required rate of return drops by 0.11%, 0.09%, 0.10%, and 0.07% for 1- to 4-year dividend strips, respectively. Transaction costs are negligible compared with the magnitudes of returns. Panel (c) in Figure 4.6 shows the required rates of returns for 1- to 4-year dividends. The shape and magnitude of this figure are almost identical to Figure 4.4. The maximum returns of 1- to 4-year horizon dividend strips are 55%, 32%, 19%, and 16%, respectively.

Panel (d) in Figure 4.6 depicts conditional Sharpe ratios, which are similar to the ones in Figure 4.5. The highest Sharpe ratios for 1- to 4-year dividends are 14.10, 11.72, 8.24, and 8.07, respectively. Transaction cost does not affect the conditional Sharpe ratio by much.

## 4.6 Simulation

So far, this paper has demonstrated that the required rate of return and conditional Sharpe ratio for dividend strips are high during crises. These numbers are too high to be rationalised intuitively. This section is going to make formal tests on whether the high conditional Sharpe ratios can be explained by selected macro-finance models.

This paper discusses four popular macro-finance models—the preferred habit model (Campbell and Cochrane 1999), the long-run risk model (Bansal and Yaron 2004), the rare disaster model (Gabaix 2012), and the term structure model (Lettau and Wachter 2011). Simulations help identify whether these models explain the empirical extreme Sharpe ratios.

The key moment of interest is the theoretical maximum conditional Sharpe ratio,  $SR_{max,t}$ , given by equation (4.12). The calculation requires the forward-looking stochastic discount factor,  $\mathbb{E}_t[M_{t+1}]$ , and its conditional standard deviation,  $\sigma_t[M_{t+1}]$ .

### A. Preferred habit Model

The preferred habit model is first developed by Campbell and Cochrane (1999), which argues that making the agent's preference dependent on surplus consumption instead of consumption explains the equity premium puzzle. The preference can be written as

$$U = \mathbb{E} \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma} \quad (4.16)$$

where  $\delta$  is the subjective discount factor,  $C_t$  is the consumption level,  $X_t$  is the habit level (reference point), and  $\gamma$  is the coefficient of relative risk aversion.

The consumption and dividend processes are external, with a constant component and a stochastic normal independent and identically distributed (i.i.d.) component of growth rates. Price-dividend ratios of market and dividend strips depend on the only state variable—surplus consumption ratio ( $S = (C - X)/C$ ). Market and dividend strip returns can be calculated with price-dividend ratios and simulated dividend levels.

The stochastic discount factor is also a function of the surplus consumption ratio,

$$M_{t+1} = \delta \frac{u_C(C_{t+1}, X_{t+1})}{u_C(C_t, X_t)} = \delta \left( \frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma} \quad (4.17)$$

where  $u_C(\cdot)$  is the derivative of  $U$  with respect to  $C$ .

The maximum conditional Sharpe ratio of the Campbell and Cochrane (CC) preferred habit model is given by the following formula,

$$SR_{max,t}^{CC} = \left\{ e^{\gamma^2 \sigma^2 [1 + \lambda(s_t)]^2} - 1 \right\}^{\frac{1}{2}} \quad (4.18)$$

where  $\sigma^2$  is the variance of the consumption growth innovation, and  $\lambda(s_t)$  is the sensitive function for the logarithm of surplus consumption ratio,  $s_t = \log(S_t)$ .

The paper follows [Wachter \(2005\)](#) and [Van Binsbergen and Koijen \(2017\)](#) to simulate the model. In particular, the study simulates the model with monthly frequency and uses the original parameters in [Campbell and Cochrane \(1999\)](#).

## B. Long-run Risk Model

The long-run risk model is proposed by [Bansal and Yaron \(2004\)](#), which uses [Epstein and Zin \(1991\)](#) preference and a stochastic process of volatility. A system of equations characterises the model setup,

$$\begin{aligned} g_{t+1} &= \mu + x_t + \sigma_t \eta_{t+1} \\ x_{t+1} &= \rho x_t + \phi_e \sigma_t \epsilon_{t+1} \\ g_{d,t+1} &= \mu_d + \phi x_t + \phi_d \sigma_t u_{t+1} \\ \sigma_{t+1}^2 &= \sigma^2 + \nu_1 (\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1} \end{aligned} \quad (4.19)$$

where  $\rho$  is the persistence coefficient of the latent process  $\{x_t\}$ ,  $\phi_e$  and  $\phi_d$  are constants that scale the overall volatility  $\sigma_t$ .  $g$  is the consumption growth rate, and  $mu$  is its constant part.  $g_d$  is the dividend growth rate, and  $mu_d$  is its constant part.  $\sigma^2$  is the equilibrium level of overall volatility.  $e_t$ ,  $\eta_t$ ,  $u_t$ , and  $w_t$  are i.i.d. standard normal distributed shocks.

The latent process  $\{x_t\}$  and the volatility process  $\{\sigma_t\}$  determine the price-dividend ratios of the market and dividend strips. Market and dividend strip returns can be calculated with price-dividend ratios and simulated dividend levels.

The logarithm of the stochastic discount factor ( $m_t = \log(M_t)$ ) depends on the consumption growth and consumption return.

$$m_t = \theta \log(\delta) - \frac{\theta}{\psi} g_t + (\theta - 1) r_{a,t} \quad (4.20)$$

where  $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$ ,  $\psi$  is the intertemporal elasticity of substitution (IES), and  $r_a$  refers to the consumption return, which depends on the price-consumption ratio.

It is not very easy to solve the conditional expectation and the conditional standard deviation of  $M_{t+1}$ , so this paper conducts one-step forward simulations for each period 1000 times. The conditional moments are replaced by the simulated sample moments.

I follow [Beeler and Campbell \(2009\)](#) and [Van Binsbergen and Kojen \(2017\)](#) to simulate the model with monthly frequency. The parameters are the same as the original Bansal and Yaron paper. The coefficient of relative risk aversion,  $\gamma$ , is chosen to be 10.

### C. Rare Disaster Model

There is a long literature on the rare disaster. A recent improvement of the rare disaster model is made by [Gabaix \(2012\)](#), which introduces cross-sectional settings and explains ten famous puzzles in financial economics. The model considers a small probability of big disasters and a time-varying resilience. These features result in a decent level of equity premium and price variation with a small coefficient of relative risk aversion.

Resilience is a state variable closely associated with equity prices, which measures how well equity performs when a disaster arrives. Resilience follows the process below,

$$\begin{aligned} H_{t+1} &= H^* + \hat{H}_{t+1} \\ \hat{H}_{t+1} &= \frac{1 + H^*}{1 + H_t} e^{-\phi_H} \hat{H}_t + \epsilon_{t+1}^H \end{aligned} \quad (4.21)$$

where  $H^*$  is the long-run level of resilience,  $\hat{H}$  is the variable part of resilience, and  $\phi_H$  is the persistence of the process. Resilience follows a twisted AR(1) process.

The dividend strip prices depend on the resilience,

$$P_{n,t} = D_t e^{-n\delta_i} \left(1 + \frac{1 - e^{-n\phi_H}}{\delta_i} \hat{H}_t\right) \quad (4.22)$$

where  $n = 12, 24, 36,$  and  $48$  months, and  $\delta_i = \delta - g_d - \log(H^*)$  is the discount factor adjusted by dividend growth rate and equilibrium resilience.

The stochastic discount factor is given by the following formula, which depends on whether the disaster occurs.

$$M_{t+1} = e^{-\delta - \gamma g_c} \times \begin{cases} 1 & \text{if there is no disaster at } t+1 \\ B_{t+1}^{-\gamma} & \text{if there is a disaster at } t+1 \end{cases} \quad (4.23)$$

where  $g_c$  is the consumption growth rate, and  $B_{t+1}$  is the ratio of the consumption with a disaster to the consumption without a disaster.

In the rare disaster model, the probability of entering a disaster period is constant, but the magnitude  $B_{t+1}$  is a random variable. This paper uses the historical distribution of  $B_{t+1}$  and the historical probability of disaster estimated by [Barro and Ursúa \(2008\)](#) to calculate the maximum conditional Sharpe ratio. As shown in equation (4.23), the conditional expectation and standard deviation of  $M_{t+1}$  are the same for all the periods because the probability of disaster is constant. Therefore, the maximum conditional Sharpe ratio can be obtained without simulation. The paper uses the same parameters as in [Gabaix \(2012\)](#) and chooses the time interval to be a month.

## D. Term Structure Model

[Lettau and Wachter \(2011\)](#) propose a dynamic risk-based model that jointly explains term structures of equity and interest rates, which is referred to as the "term structure model" hereafter. The term structure model is built on the foundation of the [Lettau and Wachter \(2007\)](#) model but adds the inflation risk.

There are six shocks in the model—shocks to dividend growth ( $g$ ), expected dividend growth ( $x$ ), inflation ( $\pi$ ), expected inflation ( $q$ ), risk-free rate ( $r_f$ ), and the price of risk

( $pr$ ). All variables are in logarithm form. The system is

$$\begin{aligned}
g_{t+1} &= x_t + \boldsymbol{\sigma}_g \boldsymbol{\epsilon}_{t+1} \\
x_{t+1} &= (1 - \rho_x)\mu + \rho_x x_t + \boldsymbol{\sigma}_x \boldsymbol{\epsilon}_{t+1} \\
\pi_{t+1} &= q_t + \boldsymbol{\sigma}_\pi \boldsymbol{\epsilon}_{t+1} \\
q_{t+1} &= (1 - \rho_q)\bar{q} + \rho_q q_t + \boldsymbol{\sigma}_q \boldsymbol{\epsilon}_{t+1} \\
r_{f,t+1} &= (1 - \rho_r)\bar{r}_f + \rho_r r_{f,t} + \boldsymbol{\sigma}_r \boldsymbol{\epsilon}_{t+1} \\
pr_{t+1} &= (1 - \rho_{pr})\bar{p}r + \rho_{pr} pr_t + \boldsymbol{\sigma}_{pr} \boldsymbol{\epsilon}_{t+1}
\end{aligned} \tag{4.24}$$

where  $\boldsymbol{\sigma}_g$ ,  $\boldsymbol{\sigma}_x$ ,  $\boldsymbol{\sigma}_\pi$ ,  $\boldsymbol{\sigma}_q$ ,  $\boldsymbol{\sigma}_r$ , and  $\boldsymbol{\sigma}_{pr}$  are  $1 \times 6$  vectors of loading on the i.i.d. normal random shocks  $\boldsymbol{\epsilon}_t$ .  $\rho_x$ ,  $\rho_q$ ,  $\rho_r$ , and  $\rho_{pr}$  measure the persistence of respective variables.  $\mu$ ,  $\bar{q}$ ,  $\bar{r}_f$ ,  $\bar{p}r$  are the long-run means of respective variables.

The model only prices the fundamental dividend risk directly, so the stochastic discount factor ( $M_t$ ) is given by

$$M_{t+1} = \exp\left(-r_{f,t+1} - \frac{1}{2}pr_t^2 - pr_t \frac{\boldsymbol{\sigma}_d}{\sqrt{\boldsymbol{\sigma}_d \boldsymbol{\sigma}'_d}} \boldsymbol{\epsilon}_{t+1}\right) \tag{4.25}$$

The maximum conditional Sharpe ratio of the Lettau and Wachter (LW) term structure model is given by the following formula,

$$SR_{max,t}^{LW} = \left\{ e^{pr_t^2 - 1} \right\}^{\frac{1}{2}} \tag{4.26}$$

## E. Simulation Results

Simulation results are reported in the left panel of Table 4.5. The 90% quantile of the maximum conditional Sharpe ratio given by the preferred habit model (CC1999) is 0.812, meaning that there is only a 10% chance of having  $SR_{max}^{CC}$  higher than 0.812. The preferred habit model predicts  $SR_{max}^{CC}$  higher than 0.927, 1.149, and 1.388 with probabilities of 5%, 1%, and 0.1%, respectively. A maximum conditional Sharpe ratio higher than 1.612 can only be observed once in 833 years.

The long-run risk model predicts even smaller maximum conditional Sharpe ratios than the preferred habit model. The 90% quantile for the distribution of  $SR_{max}^{BY}$  is 0.649. The model predicts  $SR_{max}^{BY}$  higher than 0.669, 0.708, 0.747, and 0.778 with probabilities of

Table 4.5: Theoretical Maximum Conditional Sharpe Ratio and Selected Data

Quantile	Model				Data			
	CC1999	BY2004	G2012	LW2011	$n$	2020/03	2009/02	2002/10
90%	0.812	0.649		2.819	12	14.08	9.457	4.221
95%	0.927	0.669		3.079	24	11.68	10.78	5.624
99%	1.149	0.708	1.524	3.606	36	8.256		
99.9%	1.388	0.747		4.346	48	7.921		
99.99%	1.612	0.778		5.362				

This left panel of this table reports simulation results of the theoretical maximum conditional Sharpe ratios of the preferred habit model (CC1999), the long-run risk model (BY2004), the rare disaster model (G2012), and the term structure model (LW2011). The preferred habit model and the long-run risk model are simulated for 100000 months, the term structure model is simulated for 25000 quarters, and the theoretical upper bound of the rare disaster model is calculated analytically. The right panel reports the empirical conditional Sharpe ratios for three selected periods.  $n$  refers to the horizon measured in months. The numbers are the same as in Figure 4.5. All numbers are annualised.

5%, 1%, 0.1%, and 0.01%, respectively. The long-run risk model also predicts less volatile  $SR_{max}$  compared to the preferred habit model because the sensitive function  $\lambda(s_t)$  increases fast when the surplus consumption approaches 0.

The rare disaster model proposed by [Gabaix \(2012\)](#) (G2012) uses the time-invariant probability of disaster occurrence, so the maximum conditional Sharpe ratio is a constant. Using the historical distribution of disaster from [Barro and Ursúa \(2008\)](#),  $SR_{max}^G$  is 1.524 for all periods.

The term structure model performs the best in terms of generating high maximum conditional Sharpe ratios. The 90% quantile for the distribution of  $SR_{max}^{LW}$  is 2.819, much higher than the numbers generated from other models. The model predicts  $SR_{max}^{LW}$  higher than 3.079, 3.606, 4.346, and 5.362 with probabilities of 5%, 1%, 0.1%, and 0.01%, respectively. The term structure model has the ability to predict high Sharpe ratios because it allows the risk of price ( $pr$ ) to be very volatile.

The right-hand-side panel of table 4.5 reports the conditional Sharpe ratios of dividend strips in 3 crisis periods—March 2020 (COVID-19 pandemic), February 2009 (global financial crisis), and October 2002 (dot-com bubble). These numbers are identical to the ones graphed in Figure 4.5.

In March 2020, the conditional Sharpe ratio of 1-year dividend peaks at 14.08, almost 8 times higher than the top 0.01% simulated upper bound predicted by the preferred habit model, 17 times higher than that predicted by the long-run risk model, more than

8 times higher than that predicted by the rare disaster model, and almost 3 times as high as predicted by the term structure model. The conditional Sharpe ratios of 2- to 4-year dividend strips are also much higher than the top 0.01% theoretical upper bounds. Conditional Sharpe ratios in February 2009 and October 2002 are not as high as in the COVID-19 pandemic crisis, but they are still several times higher than the theoretical upper bound.

## F. Discussion

These high Sharpe ratios are very puzzling. Given the duality of the mean-standard deviation frontier of return and the mean-standard deviation frontier of IMRS, an asset beyond the maximum Sharpe ratio lies above the theoretical capital market line (CML). The theoretical CML is the line driven by the preference of investors rather than the line estimated by assets traded in the market. Therefore, one must reject the tested models if he or she believes that the conditional Sharpe ratios are properly measured.

High conditional Sharpe ratios in recessions imply high uncertainty of the stochastic discount factor. Equation (4.12) shows that the maximum conditional Sharpe ratio is given by the ratio of the expected standard deviation of SDF to the expectation of SDF. Since the risk-free rate is close to 0, the expectation of SDF is close to 1. A maximum conditional Sharpe ratio of 14 requires the standard deviation of SDF to be about 14. Note that SDF is bounded below by 0. Hence, the conditional distribution of  $M$  must be very skewed. There must be scenarios with the realised SDF above 15. A realised SDF of 15 means an agent is willing to trade 15 dollars last year for 1 dollar now, which is hardly possible in real life.

One way to rationalise the findings is to have a very low consumption today compared to the last year, i.e. a huge disaster. This paper simulates the rare disaster model with the historical disaster distribution given by [Barro and Ursúa \(2008\)](#). The worst disaster happened in Japan during the Second World War, with a 63.9% drop in consumption. Most disasters cost 10% to 20% of total consumption. If there is an additional disaster event in the distribution that costs 83% of consumption, the rare disaster model can explain the high Sharpe ratio. However, such an event does not exist in Barro and Ursua's survey, and it is questionable whether it is possible to exist at all. Although I can't rule out the possibility that investors believe such a disaster can happen, it is unlikely that an almost

impossible event drives the result.

Another way to solve the problem is to adjust the coefficient of relative risk aversion ( $\gamma$ ), making people feel bad about today's low consumption. The upper bound  $SR_{max}$  is sensitive to  $\gamma$  in all models except the term structure model. For example, raising  $\gamma$  from 2 to 11 in the preferred habit model can explain the empirical result. When  $\gamma$  is higher, the upper bound is higher. However, having a reasonably low coefficient of relative risk aversion is one of the goals of these models. It is a backward move to raise  $\gamma$ . Also, altering  $\gamma$  must be accompanied by changing other parameters so that other conditions are met (e.g. the equity premium), which results in further problems.

## 4.7 Conclusion

This paper raises the short-duration equity return puzzle using dividend strips. Unlike most papers in the literature, this paper investigates the forward-looking required rate of return and conditional Sharpe ratio for dividend strips instead of the backwards-looking ones. The forward-looking method makes the returns comparable to those of bonds, and more importantly, connects the empirical return with the mean-standard deviation frontier of IMRS.

This paper uses dividend strip prices from Bloomberg and [Van Binsbergen et al. \(2012\)](#) dividend forecasts from IBES to obtain the required rate of return. The dividend forecast for the S&P 500 is aggregated from consensus forecasts of individual constituent companies. In order to estimate the conditional Sharpe ratio, I use an MGARCH model (AR(1)-DCC(1,1) model) to estimate the conditional variance of dividend return.

The dividend forecast from IBES appears to be accurate. Forecast errors for 1- to 4-year horizons are 0.12%, 0.23%, 0.29%, and 0.37%, respectively, and the dividend growth predictions are unbiased. Expected dividends move closely with the realised dividend.

Since short-duration dividends are usually indicated and easily predictable, dividends should have a lower risk than the stock index. It is surprising to find that the required rates of returns of dividend strips are extremely high during crisis periods. In March 2020 (COVID-19 pandemic), returns of 1- to 4-year dividend strips are 55%, 32%, 19%, and 15%, respectively. High returns also appear in 2002 (dot-com bubble) and 2008–2009 (global financial crisis).

Conditional Sharpe ratios measure the return adjusted by the risk. The highest conditional Sharpe ratio in our sample is 14 for the 1-year dividend in March 2020, which is much higher than the average S&P 500 Sharpe ratio of 0.79. Sharpe ratios for other horizons are also very high during crisis periods. The finding is robust to measurement errors and the transaction cost.

The empirical results are not consistent with mainstream macrofinance models. The top 0.01% of conditional Sharpe ratios for the preferred habit model, the long-run risk model, the rare disaster model, and the term structure model are 1.612, 0.778, 1.524, and 5.362, respectively. These extreme Sharpe ratios of dividend strips are far beyond the theoretical upper bound of conditional Sharpe ratios predicted by these models, which is called the short-duration equity return puzzle.

The short-duration equity return puzzle is a novel test criterion for macrofinance models, in addition to the equity premium puzzle, the excess volatility puzzle, and other criteria. A model that can explain the high conditional Sharpe ratio of the dividend strip must produce sufficient conditional volatility of the stochastic discount factor while keeping the coefficient of relative risk aversion within a reasonable level. This area is left for further exploration.

## References

- Ai, H., Croce, M. M., Diercks, A. M. and Li, K. (2018), ‘News shocks and the production-based term structure of equity returns’, *The Review of Financial Studies* **31**(7), 2423–2467.
- Andries, M., Eisenbach, T. M. and Schmalz, M. C. (2019), ‘Horizon-dependent risk aversion and the timing and pricing of uncertainty’, *FRB of New York Staff Report* .
- Bansal, R., Miller, S., Song, D. and Yaron, A. (2021), ‘The term structure of equity risk premia’, *Journal of Financial Economics* .
- Bansal, R. and Yaron, A. (2004), ‘Risks for the long run: A potential resolution of asset pricing puzzles’, *The journal of Finance* **59**(4), 1481–1509.
- Barro, R. J. and Ursúa, J. F. (2008), Macroeconomic crises since 1870, Technical report, National Bureau of Economic Research.

- Bauwens, L., Laurent, S. and Rombouts, J. V. (2006), ‘Multivariate garch models: a survey’, *Journal of applied econometrics* **21**(1), 79–109.
- Beeler, J. and Campbell, J. Y. (2009), The long-run risks model and aggregate asset prices: An empirical assessment, Technical report, National Bureau of Economic Research.
- Bilinski, P. and Bradshaw, M. T. (2022), ‘Analyst dividend forecasts and their usefulness to investors’, *The Accounting Review* **97**(4), 75–104.
- Boguth, O., Carlson, M., Fisher, A. J. and Simutin, M. (2012), ‘Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia’, *manuscript, Arizona State University* .
- Campbell, J. Y. and Cochrane, J. H. (1999), ‘By force of habit: A consumption-based explanation of aggregate stock market behavior’, *Journal of political Economy* **107**(2), 205–251.
- Campbell, J. Y. and Shiller, R. J. (1988), ‘The dividend-price ratio and expectations of future dividends and discount factors’, *The Review of Financial Studies* **1**(3), 195–228.
- Cassella, S., Golez, B., Gulen, H. and Kelly, P. (2021), ‘Horizon bias and the term structure of equity returns’, *Available at SSRN 3328970* .
- Chabi-Yo, F. and Loudis, J. (2020), ‘The conditional expected market return’, *Journal of Financial Economics* **137**(3), 752–786.
- Chen, Z. (2020), ‘Inferring stock duration around fomic surprises: Estimates and implications’, *Journal of Financial and Quantitative Analysis* pp. 1–35.
- DeAngelo, H., DeAngelo, L. and Skinner, D. J. (2000), ‘Special dividends and the evolution of dividend signaling’, *Journal of Financial Economics* **57**(3), 309–354.
- Dechow, P. M., Erhard, R. D., Sloan, R. G. and SOLIMAN, M. T. (2021), ‘Implied equity duration: A measure of pandemic shutdown risk’, *Journal of Accounting Research* **59**(1), 243–281.
- Engle, R. (2002), ‘Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models’, *Journal of Business & Economic Statistics* **20**(3), 339–350.

- Epstein, L. G. and Zin, S. E. (1991), ‘Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis’, *Journal of political Economy* **99**(2), 263–286.
- Gabaix, X. (2012), ‘Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance’, *The Quarterly journal of economics* **127**(2), 645–700.
- Giglio, S., Kelly, B. T. and Kozak, S. (2021), ‘Equity term structures without dividend strips data’, *Available at SSRN 3533486* .
- Golez, B. and Jackwerth, J. C. (2020), ‘Quarter-of-a-century of dividend pricing’, *Available at SSRN 3575450* .
- Golez, B. and Jackwerth, J. C. (2022), ‘Holding period effects in dividend strip returns’, *Available at SSRN* .
- Gonçalves, A. S. (2021a), ‘Reinvestment risk and the equity term structure’, *The Journal of Finance* .
- Gonçalves, A. S. (2021b), ‘The short duration premium’, *Journal of Financial Economics* .
- Gormsen, N. J. (2021), ‘Time variation of the equity term structure’, *The Journal of Finance* **76**(4), 1959–1999.
- Gormsen, N. J. and Lazarus, E. (2021), ‘Duration-driven returns’, *Available at SSRN 3359027* .
- Hansen, L. P. and Jagannathan, R. (1991), ‘Implications of security market data for models of dynamic economies’, *Journal of political economy* **99**(2), 225–262.
- Kalay, A. and Loewenstein, U. (1986), ‘The informational content of the timing of dividend announcements’, *Journal of Financial Economics* **16**(3), 373–388.
- LeRoy, S. F. and Porter, R. D. (1981), ‘The present-value relation: Tests based on implied variance bounds’, *Econometrica: journal of the Econometric Society* pp. 555–574.
- Lettau, M. and Wachter, J. A. (2007), ‘Why is long-horizon equity less risky? a duration-based explanation of the value premium’, *The Journal of Finance* **62**(1), 55–92.

- Lettau, M. and Wachter, J. A. (2011), ‘The term structures of equity and interest rates’, *Journal of Financial Economics* **101**(1), 90–113.
- Mehra, R. and Prescott, E. C. (1985), ‘The equity premium: A puzzle’, *Journal of monetary Economics* **15**(2), 145–161.
- Moench, E. and Siavash, S. S. (2021), ‘What moves treasury yields?’, *CEPR Discussion Paper Series* .
- Newey, W. K. and West, K. D. (1987), ‘Hypothesis testing with efficient method of moments estimation’, *International Economic Review* pp. 777–787.
- Shiller, R. J. (1979), ‘The volatility of long-term interest rates and expectations models of the term structure’, *Journal of political Economy* **87**(6), 1190–1219.
- Van Binsbergen, J., Brandt, M. and Koijen, R. (2012), ‘On the timing and pricing of dividends’, *American Economic Review* **102**(4), 1596–1618.
- Van Binsbergen, J. H. and Koijen, R. S. (2017), ‘The term structure of returns: Facts and theory’, *Journal of Financial Economics* **124**(1), 1–21.
- Van Binsbergen, J., Hueskes, W., Koijen, R. and Vrugt, E. (2013), ‘Equity yields’, *Journal of Financial Economics* **110**(3), 503–519.
- Wachter, J. A. (2005), ‘Solving models with external habit’, *Finance Research Letters* **2**(4), 210–226.
- Weber, M. (2018), ‘Cash flow duration and the term structure of equity returns’, *Journal of Financial Economics* **128**(3), 486–503.
- Woolridge, J. R. (1983), ‘Dividend changes and security prices’, *The Journal of Finance* **38**(5), 1607–1615.
- Wu, D. (2020), ‘A disaster explanation of equity term structures’, *Available at SSRN 3571921* .

# Beating the Index with ETFs

## Abstract

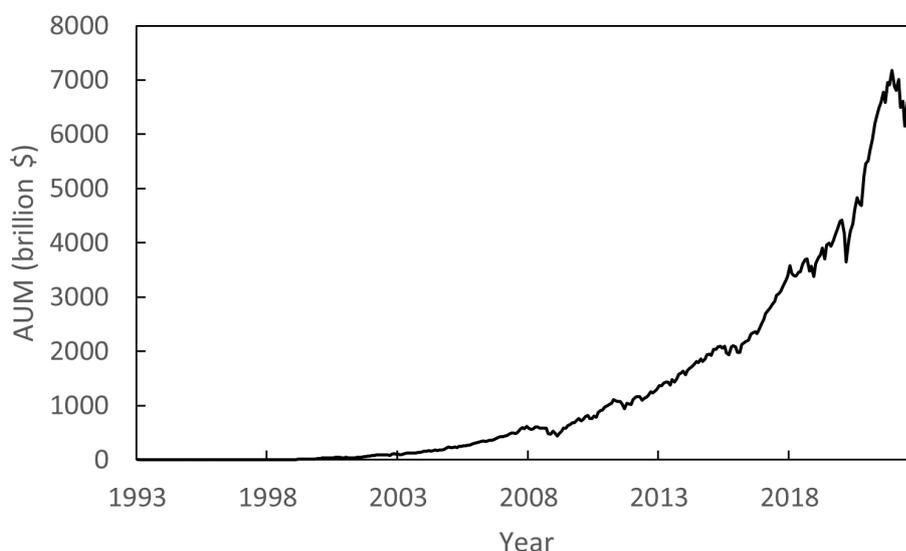
This paper uncovers a new source of tax efficiency for ETFs—using highly correlated ETFs to harvest capital losses without violating the wash-sale rule. By exploiting the tax loophole, investors can potentially earn a better return than the index. The study reveals that highly correlated ETFs have an average monthly tax-loss trading volume of 9.1% of their assets under management, which accounts for 20.7% of their total trading volume. Tax-loss harvesting is negatively related to past returns, especially for recent and negative ones. ETFs with high past volatility have higher tax-loss trading volumes, while smaller and less liquid ones have lower tax-loss trading volumes. This paper develops a parsimonious model to explain the relationship between tax-loss harvesting and past price movements. Simulations with the model predict an annual tax revenue loss of 0.52% of assets under management for highly correlated ETFs, equivalent to approximately 25 billion USD in 2021.

**Keywords:** ETF, Tax-Loss Harvesting, Trading Volume

## 5.1 Introduction

It has been three decades since the inception of the first exchange-traded fund (ETF) in 1993, and ETFs have steadily gained popularity over the years. As of June 2023, the US market featured 2,921 ETFs, collectively holding assets totalling 7.3 trillion US dollars. The scale of ETFs was about 30% of the scale of traditional mutual funds and kept rising.<sup>1</sup> The trend of the total asset under management (AUM) of ETFs is shown in Figure 5.1.

Figure 5.1: Total Asset Under Management of ETFs



ETFs are cannibalising the market share of traditional mutual funds. Intuitively, ETF and the mutual fund are substitutes. Both ETFs and mutual funds serve as investment vehicles for conveniently implementing trading strategies. By looking at fund flows of ETFs and mutual funds, [Agapova \(2011\)](#) finds that among every \$100 inflow to ETFs, \$22 is pulled out from traditional mutual funds. This substitution effect became more prominent after 2003, following an increase in the statutory capital gains tax rate.

The success of the ETF can be attributed to its advantages over traditional mutual funds, including (1) the flexibility to trade ([Gardner and Welch 2005](#)), (2) low expense ratio and transaction cost ([Kostovetsky 2003](#)), and (3) tax efficiency ([Colon 2017](#), [Moussawi et al. 2022](#)). With these virtues, index-tracking ETFs can follow the index closely even after accounting for taxes. Nevertheless, these inherent strengths of ETFs do not offer the

<sup>1</sup>See the Investment Company Institute (ICI) website: <https://www.ici.org/statistics>.

potential to outperform the index adjusted by risk. The optimal return from holding an index-tracking ETF should align with the index return.

This paper reveals a new source of tax efficiency for ETFs—tax-loss harvesting without triggering the wash-sale rule, with which ETFs can perform better than the index. The paper first documents the evidence of using highly correlated ETFs to harvest capital loss and builds a parsimonious model to explain the connection between tax-loss harvesting and past price movements. The model can also estimate the tax revenue loss from tax-loss harvesting with ETFs.

Investors are liable to pay tax for their realised capital gains over a certain threshold. If an investor expects to pay the capital gain tax and holds a losing security, she can replace that security with another to realise a capital loss. This action reduces her portfolio's total realised capital gains and, consequently, her capital gains tax burden. However, the Internal Revenue Service (IRS) has implemented a wash-sale rule, which disallows tax deductions for the purchase and sale of "substantially identical" securities within a 30-day period. ([Internal Revenue Service 2022](#)) As a result, the replacement security is hardly the same as the old one, making it challenging for investors to realise losses without making substantial changes to their portfolios.

ETFs offer an ideal solution to sidestep the wash-sale rule. According to IRS *Publication 550*, "Ordinarily, stocks or securities of one corporation are not considered substantially identical to stocks or securities of another corporation." ETFs, structured as trusts, limited partnerships, or open-end funds, represent distinct investment entities and should be regarded as different "corporations". However, despite this distinction, ETFs can exhibit high correlations. Buying and selling highly correlated ETFs change the portfolio little. The IRS does not provide any specific guidelines about tax-loss harvesting with ETFs. ([Fischer 2010](#)) Therefore, [Bouchey et al. \(2016\)](#) conclude, "It seems clear that to sell an ETF at a loss and then buy back the same ETF within 30 days would be a wash sale. What is less clear is whether buying back a different ETF would be a wash sale... There is no specific test to determine whether two funds are substantially identical or not."

For example, an investor holds 100 shares of SPDR S&P 500 ETF (SPY), purchased at \$470 each. The current price of SPY is \$440, and the investor wishes to realise the loss to deduct the capital gain. By selling all the ETF shares, she can realise a \$3,000 loss and offset a \$600 tax with a marginal tax rate of 20%. If she intends to hold a similar portfolio,

she can either repurchase 100 shares of SPY, voiding the tax deduction, or acquire an equivalent value of iShares Russell 1000 ETF (IWB). Since both SPY and IWB are broad-market trackers, they have almost the same return. The correlation coefficient of the two ETFs is 0.9967. The portfolio changes little after harvesting the loss. More importantly, tax-loss harvesting with ETFs resides in a regulatory grey area of US tax law. Under current regulations, selling SPY and buying IWB does not trigger the wash-sale rule. The investor can defer the capital gain tax with seemingly no cost.

In practice, the financial industry appears to benefit from this tax law loophole. Vestmark, a financial technology company providing a platform for over \$1.5 trillion assets, implies the use of tax-loss harvesting with ETFs in its advertisement, "(Tax-loss harvesting) is especially advantageous within direct indexing SMAs because the number of holdings and available substitutes are not as constrained as they would be in an actively managed portfolio."<sup>2</sup> A more explicit guideline of such a tax-avoidance method comes from Investopedia, which explains, "For example, if an investor sells the SPDR S&P 500 ETF (SPY) at a loss, they can immediately turn around and purchase the Vanguard S&P 500 ETF... The rationale is that the two ETFs have different fund managers, different expense ratios, may replicate the underlying index using a different methodology, and may have different levels of liquidity in the market. Presently, the IRS does not deem this type of transaction as involving substantially identical securities, so it is allowed, although this may be subject to change in the future as the practice becomes more widespread."<sup>3</sup>

By engaging in ETF trading for tax purposes, investors can potentially achieve superior after-tax portfolio returns compared to merely holding ETFs. This additional benefit can be considered a tax credit for ETF traders. Since ETFs can closely track their respective indices, the extra tax benefit aids ETFs in "beating" the index. Simulating with historical return data between 2011 and 2015, [Bouchey et al. \(2016\)](#) find that the theoretical maximum tax alpha generated from tax-loss harvesting with two ETFs is 1.02%, slightly lower than the maximum tax alpha of trading individual stocks (1.62%).

This paper investigates how much of the tax-loss harvesting potential of ETFs is exploited by investors. The empirical research focuses on the structural change triggered by introducing a similar ETF. In situations where there isn't another highly correlated ETF available in the market, tax-loss harvesting remains impractical due to the absence of a

---

<sup>2</sup> Vestmark's *Tax Management Capabilities*, link.

<sup>3</sup> Investopedia, *Substantially Identical Security: Definition and Wash Sale Rules*, link.

suitable substitute. However, with the launch of a second ETF highly correlated with the existing one, the tax-loss harvesting strategy suddenly becomes feasible. It is expected that the trading volume of the first ETF will increase after this structural change. The paper uses difference-in-difference (DID) regressions to identify tax-loss trades. Following the introduction of the second ETF, the monthly trading volume of the first ETF experiences a 9.2% increase relative to its Assets Under Management (AUM), representing 20.7% of the post-treatment trading volume. Event study shows that the increase in trading volume accumulates gradually over a five-year period after the treatment.

Furthermore, the paper explores the relationship between tax-loss harvesting behaviour and past price movements. Returns in the past 12 months are negatively associated with the tax-loss harvesting trading volume. Intuitively, lower returns in the past lead to higher unrealised capital losses and, hence, higher potentials of tax-loss harvesting. The return effect is stronger for negative returns than for positive returns. The effect is most substantial for the last month's return and decays as the lag becomes longer. This decaying response to returns aligns with US tax law, which permits deductions for short-term capital gains when securities losses are realised within one year of purchase. The highest marginal tax rate for short-term capital gains is 37%, considerably higher than the 20% rate for long-term capital gains. Therefore, short-term capital losses are more valuable to investors. The paper also finds that past return volatility is positively associated with tax-loss harvesting behaviour. The rationale behind this relationship is that high past volatility results in a wider distribution of tax bases, leaving more investors in the tax-loss trade region.

ETF characters are found to correlate with the tax-loss harvesting behaviour. ETFs with high bid-ask spreads tend to exhibit low tax-loss trading volumes because spreads are implicit transaction costs. Smaller ETFs demonstrate lower tax-loss trading volumes, reflecting constraints related to market depth. After controlling for size, the paper finds that the expense ratio does not explain the tax-loss harvesting behaviour.

Tax-loss harvesting depends on the trade-off between deferring tax payments and transaction costs, making market incentives an important factor in this process. The study reveals that investors are more inclined to engage in tax-loss harvesting with ETFs when the medium-to-long-term risk-free interest rate is elevated, but they exhibit reduced trading in paired ETFs when the short-term fund is tighter. In addition, there is no evidence supporting more tax-loss trades in December.

To explain the empirical findings, the paper establishes a parsimonious two-period model and solves it numerically. The model encompasses a continuum of investors with some distribution of tax basis. Investors have random initial holdings of one tradeable asset and one non-tradeable asset. In period 1, the non-tradeable asset pays a dividend, and agents trade the tradeable asset to maximise their expected utility with constant absolute risk aversion. Trades are subject to the transaction cost and the income tax. The equilibrium determines the asset price. In period 2, both assets pay dividends, and agents die. The model simulates scenarios with and without the wash-sale rule separately. The difference in trading volumes of the two scenarios is the tax-loss harvesting trading volume. The model explains the connections between the tax-loss harvesting trading volume and past return, past volatility, and the bid-ask spread.

Moreover, the model can estimate the tax revenue loss and the tax alpha from tax-loss harvesting. By tuning parameters to match the empirical estimation of the tax-loss trading volume, this paper finds that the annual tax revenue loss is about 0.52% of AUM of highly correlated ETFs, equivalent to 25 billion USD. The transaction cost has a minimal direct impact on the tax alpha but has a strong indirect impact through the trading volume.

The research is particularly essential for policymakers. The ETF has been well known for its tax efficiency. The holding tax efficiency of ETFs comes from the seemingly non-existent capital gain distribution, which heavily relies on heartbeat trades designed solely for tax purposes. The trading tax efficiency of ETFs relies on two highly correlated ETFs not identified as "substantially identical" by the IRS. While tax-loss harvesting is a legal and advantageous strategy for investors, its broader economic implications warrant scrutiny. It is crucial to acknowledge that tax efficiency does not equate to market efficiency. There are three problems with ETF tax efficiency. Firstly, ETF investors do not receive the same treatment as they would if they directly purchased individual stocks. Secondly, ETFs enjoy tax advantages over mutual funds, creating a disparity. Lastly, these tax loopholes diminish government tax revenue. Policymakers should carefully think about the welfare implications of the current ETF tax law.

The paper proceeds as follows. Section 5.2 discusses the contribution of this research to the existing literature. Section 5.3 undertakes an empirical investigation into tax-loss harvesting using ETFs. A simple tax-loss harvesting model is presented in Section 5.4. Finally, Section 5.5 serves as the conclusion of the research.

## 5.2 Related Literature

This paper mainly contributes to two streams of literature—the development of ETF and the capital gain tax. The tax-loss harvesting model is related to the literature on trading volume.

### A. ETF and Mutual Fund

The ETF gained popularity quickly, taking the market share from traditional mutual funds. (Figure 5.1) There is a long literature trying to explain the success of the ETF. [Lettau and Madhavan \(2018\)](#) argue that ETFs are usually more liquid than underlying securities. [Kostovetsky \(2003\)](#) builds a 1-period threshold model to compare the ETF and the mutual fund. He finds that the expense ratio, transaction costs, tax considerations, and qualitative factors are important in explaining the switch from mutual funds to ETFs. [Gardner and Welch \(2005\)](#) argue that ETFs offer investors greater flexibility due to their structural characteristics. Like stocks, ETFs are traded during regular trading hours, while mutual fund shares can only be redeemed after the market closes. Consequently, ETFs typically exhibit superior liquidity. Besides, it is easy to hold a leveraged or short position on ETFs, whereas short-selling mutual fund shares is not a viable option.

[Agapova \(2011\)](#) answers why ETFs and mutual funds coexist. Using CRSP and N-SAR data between 2000 and 2004, the author investigates how ETF and mutual fund flows interact. The paper shows that the ETF performs better regarding tracking errors. The ETF and the mutual fund are not perfect substitutes. The inflow to the ETF is associated with a 22% outflow from the mutual fund. The substitute effect becomes higher after the 2003 tax reform, which decreases the top-tier capital gain tax rate from 20% to 15%.

Tax efficiency is a crucial factor in the ETF's success. [Frei and Welsh \(2022\)](#) establish a rank-dependent expected utility model to find equilibrium market shares for ETFs and mutual funds. They assume agents have constant absolute risk aversion and returns are drawn from a given distribution. Simulation shows that agents will move from ETFs to mutual funds if ETFs become taxable. Besides, authorised participants (APs) have the ability to create (destroy) ETF shares by giving (withdrawing) a basket of underlying securities to the ETF sponsor. By delivering securities with the lowest basis to the redeeming AP, the ETF sponsor can circumvent capital gain distributions to investors. This practice, known

as "in-kind redemption," enjoys tax-exempt status by law. [Colon \(2017\)](#) argues that this tax exemption should be regarded as an "ETF tax subsidy". In cases where there are still positive realised capital gains despite day-to-day redemptions, ETFs can collaborate with APs to execute the so-called "heartbeat" trade to eliminate any remaining realised capital gains. A "heartbeat" trade involves a large inflow of securities followed by a similar-size outflow to the same AP days later. [Moussawi et al. \(2022\)](#) find that ETFs and mutual funds realise similar capital gains, but the capital gain distribution of ETFs is almost 0. The tax difference produces an annual tax alpha of 0.92%. The result is robust to industry, value, and size. The paper also finds that the realised capital gain drives heartbeat trades and explains migration from active funds to ETFs. By exploiting the quasi-natural experiment of the 2012 tax reform, the paper finds that high-net-worth investors switch from mutual funds to ETFs when the tax rate increases after 2012.

The holding tax efficiency of the ETF comes from the tax-exempt in-kind redemption. Looking at the returns and tax distributions of SPY and Vanguard Index 500 Fund between 1994 and 2000, [Poterba and Shoven \(2002\)](#) find the two funds have similar returns and dividends, but SPY distributes almost no capital gain due to in-kind redemption. Some papers criticise the exploitation of in-kind redemption as an ETF tax subsidy. ([Colon 2017](#), [Hodaszy 2016, 2022](#)) First documented by [Kashner \(2017\)](#), heartbeat trades are found to be used to eliminate realised capital gains of ETFs, enhancing after-tax returns by 6 bps per quarter. [Moussawi et al. \(2022\)](#) find that heartbeat trades can reduce the tax burden by 0.86% per year. [Colon \(2023\)](#) notices that Congress set limits on in-kind redemption in 2012, but it still allows non-pro-rata security distribution due to index rebalancing, which is exploited by heartbeat trades.

Although many papers document the advantages of ETFs over mutual funds, it does not mean there is unanimity that ETFs dominate mutual funds. For example, [Gastineau \(2004\)](#) finds the Russell 2000 mutual fund performs better than the ETF tracking the same index in 2001 and 2002. The author argues that mutual funds have better operational efficiency because they do not need to keep the security basket fixed for the whole trading day.

## B. Capital Gain Tax

The existence of capital gain tax has many implications for the capital market. The most obvious impact is the change in after-tax returns. [Bergstresser and Pontiff \(2013\)](#) collect tax regulations between 1927 and 2009, calculating the effective tax rates and after-tax performance for several strategies. They investigate the effective tax rates for different quantiles of income. The paper shows that the momentum strategy has the best pre-tax performance, but the value strategy has the best after-tax performance. The value and size premia are lower after deducting the tax. [Sialm \(2009\)](#) argues that the tax burden is capitalised. i.e. Investors factor in the tax impact when pricing securities. Sialm calculates effective tax rates in both time series and cross-section. Using a multifactor pricing model, the paper finds that stocks with high dividends tend to have high returns, especially for high-tax periods.

The capital gain tax encourages tax-loss harvesting behaviour. [Chaudhuri et al. \(2020\)](#) backtest the tax-loss harvesting alpha between 1926 and 2018, estimating an average tax alpha of 1.08% when the wash-sale rule does not apply and 0.82% when the wash-sale rule applies. The tax alpha is higher in recessions than expansions, peaking at 2.13% during the great depression. Besides, the tax alpha increases with return volatility, drawdown, tax rate, and cash contribution.

Some papers find the turn-of-year effect is associated with tax-loss harvesting. [Poterba and Weisbenner \(2001\)](#) show that the return in the first five trading days in January is negatively related to past losses, while past gains have no impact on the return in the new year. The effect is found to be stronger for small firms. The magnitude of the effect depends on tax laws, confirming the existence of tax-loss harvesting behaviour. [Ivković et al. \(2005\)](#) investigate the disposition and lock-in effects with taxable and tax-deferred accounts of 78000 households. They find that tax-loss harvesting happens all year round, while December witnesses above-average (16.9%) loss sales. Positions purchased within six months are more likely to be sold for tax purposes. Capital gains lock-in effect exists in taxable accounts, especially for large and long-term positions. The disposition effect also exists. [Agarwal et al. \(2014\)](#) also find that tax-loss harvesting happens all year round. [Grinblatt and Moskowitz \(2004\)](#) find that tax-loss harvesting is associated with the momentum effect. Using hedged portfolios neutral to value, size, and industry, they

find that consistent winner stocks have positive returns. Winner and loser stocks behave in opposite ways around the turn of the year, indicating the tax-loss selling. Small stocks with high turnover and low institutional ownership exhibit more pronounced turn-of-year and seasonal effects.

### C. Trading Volume

A long piece of literature focuses on why the trading volume exists. Intuitively, no trade occurs if all agents agree on the same price as new information comes. By allowing agents to be heterogeneous, [Karpoff \(1986\)](#) finds that there can be a normal trading volume with random meetings between buyers and sellers. Besides, the trading volume persists. [Holthausen and Verrecchia \(1990\)](#) classify the trading volume in response to information into two effects—*informedness* and *consensus*. They find that high informedness leads to high return volatility and trading volume, while high consensus leads to high volatility and low trading volume. [Easley and O'Hara \(1992\)](#) argue that the trading volume depends on the adverse selection of informed investors. A higher trading volume relates to a faster convergence to the equilibrium. In addition, [Odean \(1998\)](#) and [Scheinkman and Xiong \(2003\)](#) argue that overconfidence results in high trading volumes.

Many papers focus on the connection between security price movement and trading volume. [Karpoff \(1987\)](#) surveys the literature and finds that the trading volume reacts positively to both return and the absolute value of return, raising the asymmetric volume-price hypothesis. [Kim and Verrecchia \(1991\)](#) build a 2-period rational expectation model to investigate the return and volume in response to announcements. They show that price changes proportionally to the importance of information, and volume changes proportionally to the absolute value of return. The preciseness of information is positively associated with both volume and return, while the preciseness of pre-announcement information decreases them. [Wang \(1994\)](#) also proposes a model with asymmetric information, arguing that volume positively correlates to absolute return and dividend. The paper also shows that the trading volume contains no information on price. Borrowing Wang's idea, [Llorente et al. \(2002\)](#) show that the asymmetric information model can explain the empirical findings of momentum and volume. They find that hedging trades reverse while speculative ones persist following high-volume days, especially for small and high-spread firms. [Chen et al. \(2001\)](#) conduct empirical research on nine markets, finding that absolute return is

positively associated with trading volume. Moreover, return and volume Granger cause each other. Return volatility persists after controlling for the trading volume.

## 5.3 Empirical Analysis

The sample of this research includes all index-tracking ETFs listed on the Center for Research in Security Prices (CRSP) mutual fund file. A fund is included if the indicator variable *et\_flag* is "F" and *index\_fund\_flag* is "D" (full replication index fund). Return and other variables are collected from the CRSP monthly security file. The two files are paired based on security CUSIP. The federal funds rate is obtained from the Federal Reserve Bank of St. Louis website. The risk-free yield curve for horizons between 1 and 10 years is obtained from OptionMetrics. The sample covers the period between January 1993 and December 2022. Some trading volume data points appear wrongly recorded, with the monthly volume size as high as 939 times AUM. Those records are usually associated with small ETFs with short listing periods. To deal with the problem, the relative trading volume is winsorised at 99% level. The robustness check shows that the result barely changes if these records are discarded.

### A. Variable Construction

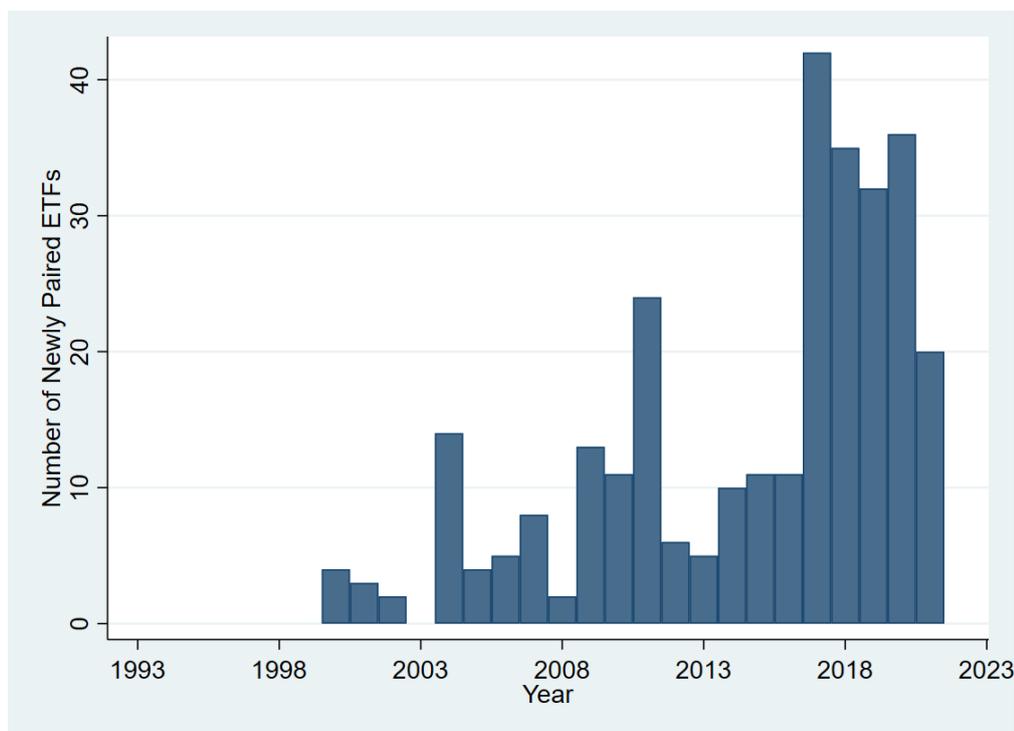
The research aims to see how introducing the first similar ETF affects the trading volume of the existing ETF. Without a similar ETF, investors cannot harvest loss without changing the portfolio substantially. Therefore, the introduction of the first similar ETF is a qualitative change, and the cutoff is clear. The definition of ETF similarity is essential. If the old and the new ETFs are not close enough, tax-loss harvesting with them results in a change in the return-risk character of the portfolio. This research defines two highly correlated ETFs as an ETF pair. ETFs  $j$  and  $k$  form a pair if they have at least 12 overlapping observations, and their returns have a correlation coefficient over 99%. If ETF  $j$  is paired with any other ETFs in month  $t$ , the tax-loss harvesting strategy is available without triggering the wash-sale rule, and the dummy variable  $paired_{j,t}$  takes the value of 1. Otherwise, the wash-sale rule applies, and  $paired_{j,t} = 0$ .  $paired$  keeps taking the value of 1 if more than 2 ETFs are highly correlated because only the introduction of the first similar ETF changes the strategy feasibility qualitatively. This method excludes the

introduction of a similar ETF in the last 12 months of the sample, so the paper excludes observations in 2022 in the following regressions.

The specification method with correlation coefficient has an advantage over grouping ETFs by index. For example, SPY tracks the S&P 500 index, and the iShares Russell 1000 ETF (ticker: IWB) tracks the Russell 1000 index. If ETFs are grouped by index, SPY and IWB are not considered to be similar. However, the returns of the two ETFs have a correlation coefficient of 0.9967. Tax-loss harvesting with SPY and IWB is feasible with a tiny mismatching risk. Following [Brown et al. \(2021\)](#), this paper primarily focuses on the correlation threshold of 99% but also explores the impacts of using different correlation thresholds. In addition, ETFs tracking the same index are investigated as an alternative identification strategy.

Figure 5.2 shows the number of newly paired ETFs by year. The first pair of ETFs appeared in the market in 2000 when the iShares Core S&P 500 ETF (ticker: IVV) started to be traded as a substitute for SPY. The number of similar ETFs increased rapidly afterwards. 42 ETFs had the first similar ETF in 2017. The introduction of the first similar ETF is cyclical. There were only two new ETF pairs in 2008 due to the global financial crisis, while the number of new ETF pairs peaked in 2004, 2011, and 2017.

Figure 5.2: Introduction of the First Similar ETF by Year



The size difference across ETFs is enormous, so the trading volume needs to be adjusted

by the ETF scale. The dependent variable,  $volume_{j,t}$ , is the proportion of outstanding shares of ETF  $j$  traded within month  $t$ . Other variables are defined as follows.  $corr99_{j,t}$  is the number of similar ETFs,  $r_{m,t}$  is the value-weighted stock market return,  $spread_{j,t}$  is the ETF spread relative to its price,  $size_{j,t}$  is the logarithm of AUM,  $fee_{j,t}$  is the average expense ratio in the past 12 months,  $r_{j,s,t}$  is the return  $s$  months before month  $t$ , and  $volatility_{j,t}$  is the realised volatility in the last month, measured as the sum of squared daily returns.

Table 5.1 presents the descriptive statistics of key variables. The full-sample mean of  $volume$  is 0.42, meaning that 42% of ETF shares are traded in the market every month. The average trading volume of paired ETFs is 0.44, slightly higher than those without similar ETFs. 27% of the observations have  $paired = 1$ . ETFs have 1.32 similar ETFs on average. Some ETFs do not have similar ETFs in the whole sample, while the most popular ETF has 59 similar ETFs. For ETFs highly correlated with others, the average number of similar ETFs is 4.84.  $r_m$  takes similar values for both sub-samples. The average spread of ETF is as low as 0.24%, but the maximum spread is 7.05 times the price. The average bid-ask spread is lower for ETFs that are highly correlated with others. The average  $size$  is 18.82, corresponding to \$149 million. The variation of  $size$  is big. ETFs with  $paired = 1$  tend to have bigger sizes. The average expense ratio in the sample is 0.44%, with the lowest fee of -0.14% and the highest fee of 4.50%. ETFs with  $paired = 1$  have an average expense ratio of 0.28%, lower than the average of ETFs with  $paired = 0$  (0.50%). The average return of ETFs ( $r$ ) is 0.45%, lower than the market return ( $r_m$ ), but ETF returns can be more extreme than the market. The realised volatility of ETFs ( $volatility$ ) is around 0.01 on average, but it can be as high as 15.07 in extreme situations.

## B. Identify Tax-Loss Harvesting

This paper uses difference-in-difference (DID) regressions to identify the marginal effect of the introduction of a similar ETF. The variable  $paired$  serves as the interaction term of the treated and the post-treatment dummies. The treated and the post-treatment dummies are captured by ETF and time fixed effects. The main regression is as follows.

Table 5.1: Descriptive Statistics

	Full Sample				<i>paired</i> = 0		<i>paired</i> = 1	
	Mean	SD	Minimum	Maximum	Mean	SD	Mean	SD
<i>volume</i>	0.42	1.00	0.00	9.95	0.41	1.01	0.44	0.97
<i>paired</i>	0.27	0.44	0	1				
<i>no_paired</i>	1.32	4.48	0	59			4.84	7.50
<i>r<sub>m</sub></i>	0.84%	4.61%	-18.48%	12.97%	0.84%	4.56%	0.87%	4.74%
<i>spread</i>	0.24%	2.24%	-99.99%	704.74%	0.28%	2.60%	0.13%	0.40%
<i>size</i>	18.82	2.34	11.36	26.84	18.34	2.09	20.11	2.46
<i>fee</i>	0.44%	0.25%	-0.14%	4.50%	0.50%	0.24%	0.28%	0.20%
<i>r</i>	0.45%	5.88%	-89.65%	159.41%	0.41%	5.99%	0.55%	5.58%
<i>volatility</i>	0.01	0.06	0.00	15.07	0.01	0.07	0.00	0.01

The table presents descriptive statistics of key variables in the paper. The number of observations is 231886, but there are 30687 missing data points for *fee*.

$$\begin{aligned}
volume_{j,t} = & \beta_0 + \beta_1 paired_{j,t} + \beta_2 no\_sim_{j,t} + \beta_3 r_{j,1,t} + \beta_4 volatility_{j,t} \\
& + \beta_5 r_{m,t} + \beta_6 spread_{j,t} + \beta_7 size_{j,t} + \beta_8 fee_{j,t} + \gamma_t + \delta_j + \epsilon_{j,t}
\end{aligned} \tag{5.1}$$

where  $\gamma_t$  refers to the time fixed effect,  $\delta_j$  represents the ETF fixed effect, and  $\epsilon_{j,t}$  is the error term. Standard errors are clustered at the ETF level.

The main regression result is reported in Column (3) of Table 5.2. The coefficient of *paired* is 0.091 with a *t*-statistic of 11. The monthly trading volume increases by 9.1% of AUM if an ETF is paired with a similar ETF, corresponding to 20.7% of the post-treatment trading volume. The coefficient of *no\_paired* is small and insignificant (0.001), showing that the increase in trading volume is not from the increased number of substitutes. The first similar ETF is essential to trading volume, but not the following ones.

Column (4) includes four more variables to capture the impacts of increasing the number of similar ETFs further. *paired2* takes the value of 1 if an ETF has two similar ETFs, *paired3* = 1 if it has three similar ETFs, and so on. The coefficient of *paired* drops slightly to 0.08. Point estimates of the coefficients decrease monotonically as the number of substitutes goes up. The introduction of a second similar ETF increases the trading volume by 3.6%, and the coefficient is significant at the 1% level. The result indicates that the second substitute is still valuable for tax-loss harvesting, although the marginal impact is much lower than the first. The reason is that the wash-sale rule has a 30-day limit. If, after harvesting a loss with two highly correlated ETFs, investors find another opportunity

to harvest the loss within 30 days, they cannot switch back to the original ETF to claim the tax benefit. A third highly correlated ETF helps these investors grasp the opportunity. The coefficient for *paired3* is marginally significant, while those for *paired4* and *paired5* are insignificant.

Table 5.2: Difference-in-Difference Regression Result

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>paired</i>	0.135*** (0.008)	0.094*** (0.008)	0.091*** (0.008)	0.080*** (0.008)
<i>paired2</i>				0.036*** (0.010)
<i>paired3</i>				0.022* (0.013)
<i>paired4</i>				0.018 (0.014)
<i>paired5</i>				0.002 (0.015)
<i>no_paired</i>	0.002* (0.001)	-0.001 (0.001)	0.001 (0.001)	-0.000 (0.001)
Control	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes
ETF FE	No	No	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.040	0.079	0.080	0.080

The table reports regression results of different specifications of Equation (5.1). Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

Columns (1) and (2) show regression results without fixed effects, which are not DID regressions. The ETF fixed effect only explains a small portion of the variation of trading volume, while the time fixed effect captures the variation of trading volume over time. The coefficients of *paired* are not very different from the main regression.

One concern of using DID is the existence of a parallel trend before the treatment. Tax-loss harvesting without triggering the wash-sale rule is only available with ETF pairs. It is pointless to trade more before introducing the first similar ETF. Therefore, there should be no increase in trading volume prior to the treatment. An ETF may respond to the scheduled launch of a new ETF in advance. By including a set of fund characters as control variables, the research mitigates, if not eliminates, the impact of the pre-treatment response.

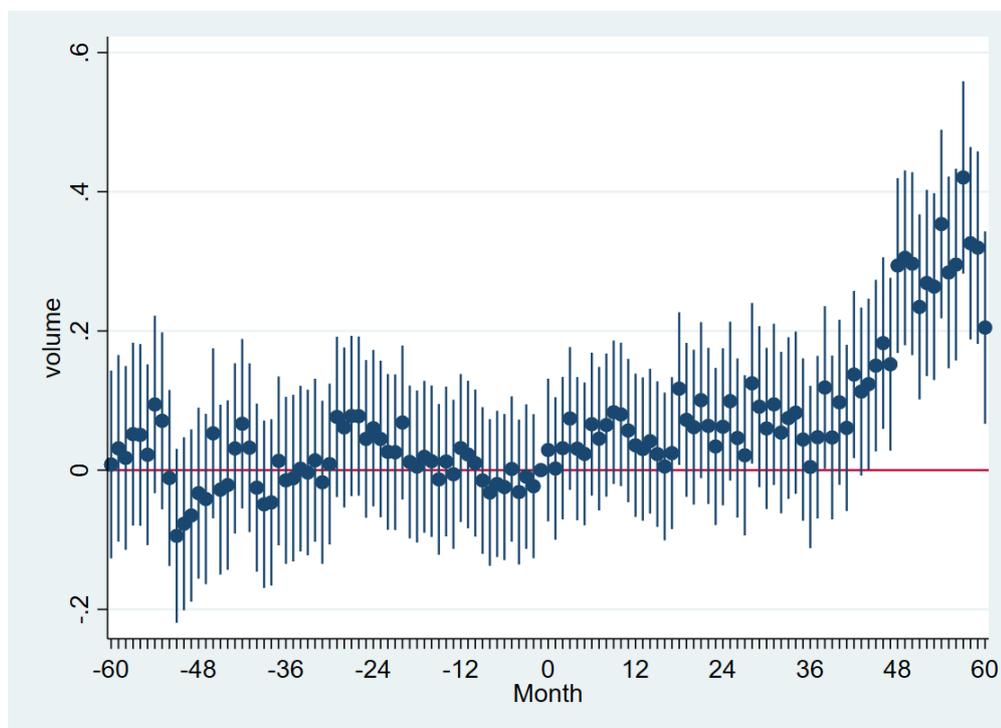
Using the event study method, this paper tests the parallel trend before the treatment.

The timeframe of the event study is five years before and after the treatment. The regression includes a series of dummy variables indicating the time relative to the treatment month.  $D_{j,\tau,t}$  takes the value of 1 if month  $t$  is  $\tau$  months after the treatment. Observations more than five years before or after the treatment are grouped into  $\tau = -61$  or  $\tau = 61$ . The base level is set to be  $\tau = -1$ , so there is no intercept term. The regression is as follows.

$$\begin{aligned} volume_{j,t} = & \sum_{\tau=-61}^{61} D_{j,\tau,t} + \beta_2 no\_sim_{j,t} + \beta_3 r_{j,1,t} + \beta_4 volatility_{j,t} \\ & + \beta_5 r_{m,t} + \beta_6 spread_{j,t} + \beta_7 size_{j,t} + \beta_8 fee_{j,t} + \gamma_t + \delta_j + \epsilon_{j,t} \end{aligned} \quad (5.2)$$

Figure 5.3 plots the  $D_{j,\tau,t}$  coefficients and respective 95% confidence intervals. As shown in the figure, none of the dummy variables before the treatment significantly differs from 0, confirming a parallel trend before treatment. The trading volume gradually increases after the introduction of the first similar ETF. All point estimates of dummies after the treatment are greater than 0, although not all are significant. Tax-loss trading volume increases rapidly after three years.

Figure 5.3: Event Study: Tax-Loss Trading Volume around Treatment



The table reports coefficients and 95% confidence intervals for  $D_{j,\tau,t}$  in Equation (5.2).  $D_{j,-1,t}$  is normalised to 0.

### C. Tax-Loss Harvesting and Price Movement

The tax-loss harvesting behaviour is strongly connected with past movements of ETF prices. If the price of ETF drops in the past months, the current unrealised capital loss is more likely to be positive, which can be harvested. If the price increases in the past months, it is more likely that the tax-loss harvesting opportunity will disappear. Realised volatility also matters. If the past volatility is high, the tax basis of investors has a broader distribution. As a result, more investors find tax-loss harvesting profitable after deducting transaction costs. The paper uses different specifications of past movements to unveil the connection.

**Past Return** Table 5.3 presents the regression results. Column (1) is the main regression, the same as Column (3) in Table 5.2. The coefficients of both  $r_1$  and *volatility* are significantly positive. The coefficient of  $r_1$  means a 10% rise of the ETF return in the last month results in an increase in the total trading volume by 3.18% of AUM. A 1% increase in last month's realised volatility leads to an increase in the relative trading volume of 0.36%. The result is consistent with the relationship between trading volumes and price movements of stocks. (Karpoff 1987, Chen et al. 2001)

Column (2) adds  $r_{2-12}$  to the list of independent variables, which is the cumulative return from 12 months ago to 2 months ago. The coefficient of  $r_{2-12}$  is 0.189, significantly positive but lower than that of  $r_1$ , showing that investors react more strongly to recent price movements.

Column (3) adds interaction terms  $r_1 \times \textit{paired}$  and  $r_{2-12} \times \textit{paired}$  to investigate the tax-loss trading volume due to past returns. The interaction terms capture the additional trading volume of being paired, not the total trading volume. The coefficients of both interaction terms are significantly negative, meaning the tax-loss trading volume is high when the loss is high (gain is low). The coefficient of  $r_1 \times \textit{paired}$  is -0.606, meaning that when facing a 10% price increase in the past month, ETFs highly correlated with others have a lower trading volume (6.06% of AUM) compared with others. More importantly, the sum of coefficients of  $r_1$  and  $r_1 \times \textit{paired}$  is -0.152, significantly different from 0. Unpaired ETFs react positively to last month's return, while paired ETFs react negatively. The coefficient of  $r_{2-12} \times \textit{paired}$  is -0.325, lower than that of  $r_1 \times \textit{paired}$  in terms of the absolute magnitude. Investors conduct more tax-loss harvesting in response to the recent price

Table 5.3: Tax-Loss Trading Volume in Response to Past Returns

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>paired</i>	0.091*** (0.008)	0.097*** (0.008)	0.124*** (0.008)	0.059*** (0.008)
$r_1$	0.318*** (0.034)	0.332*** (0.034)	0.454*** (0.036)	0.373*** (0.034)
$r_1 \times \textit{paired}$			-0.606*** (0.059)	
$r_{2-12}$		0.189*** (0.011)	0.246*** (0.011)	
$r_{2-12} \times \textit{paired}$			-0.325*** (0.019)	
<i>volatility</i>	0.360*** (0.044)	0.359*** (0.044)	0.346*** (0.044)	0.275*** (0.044)
<i>volatility</i> $\times$ <i>paired</i>				7.539*** (0.287)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.080	0.082	0.084	0.084

The table reports results for regressions with different specifications of past returns. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

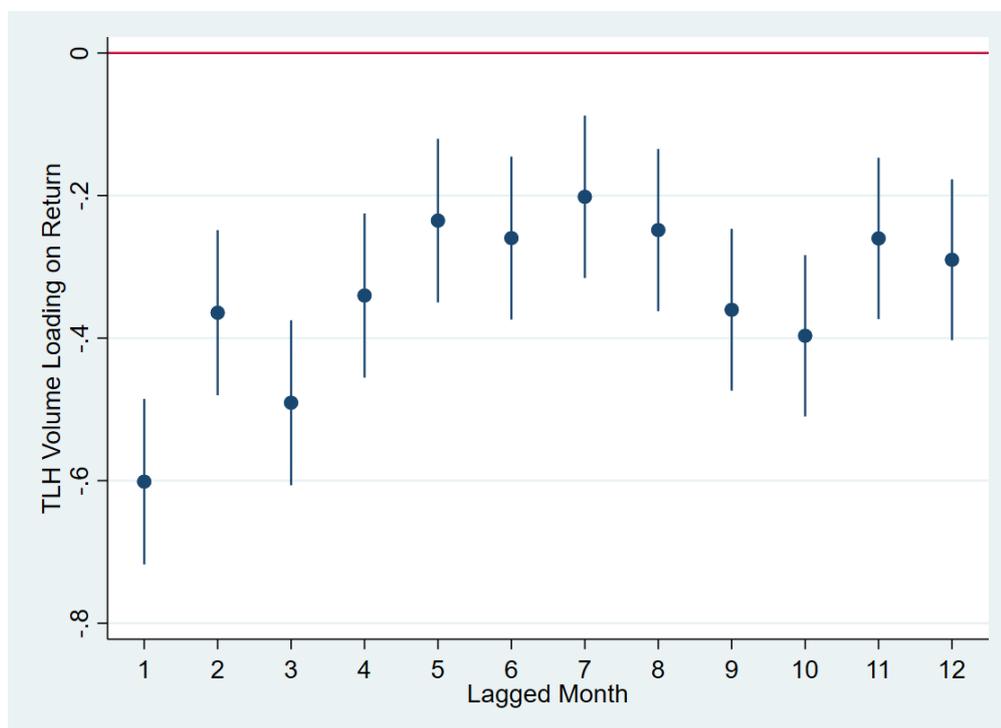
change than the older one, which is consistent with the expectation. In the US, capital gains from securities held for over one year are considered long-term capital gains, which have a lower tax rate than short-term capital gains. Recent returns are more likely to be associated with short-term gains, while returns with further lags are more likely to be associated with long-term gains. Therefore, recent losses are more "valuable" from the view of tax-loss harvesting. The sum of coefficients of  $r_{2-12}$  and  $r_{2-12} \times \textit{paired}$  is also significantly negative.

This paper delves further into the tax-loss harvesting behaviour in response to past returns with more lags. Figure 5.4 depicts the coefficients and 95% confidence intervals of the interaction terms ( $\eta_\tau$ ) in the following regression,

$$\begin{aligned}
\textit{volume}_{j,t} = & \beta_0 + \beta_1 \textit{sim}_{j,t} + \sum_{\tau=1}^{12} \theta_\tau r_{j,\tau,t} + \sum_{\tau=1}^{12} \eta_\tau r_{j,\tau,t} \times \textit{sim}_{j,t} + \beta_2 \textit{no\_sim}_{j,t} \\
& + \beta_4 \textit{volatility}_{j,t} + \beta_5 r_{m,t} + \beta_6 \textit{spread}_{j,t} + \beta_7 \textit{size}_{j,t} + \beta_8 \textit{fee}_{j,t} + \gamma_t + \delta_j + \epsilon_{j,t}
\end{aligned} \tag{5.3}$$

As shown in the figure, the coefficient decreases with lag. If the return of the last month is 10% higher, paired ETFs have a lower trading volume (6.01% of AUM) than the unpaired ones. If the return 7 months ago is 10% higher, paired ETFs only have a 2.02% lower relative trading volume. It further confirms that investors conduct more tax-loss trades in response to the recent price changes. The result is consistent with [Ivković et al. \(2005\)](#), who find the probability of selling securities decreases with the length of the holding period.

Figure 5.4: Tax-loss Trading Volume in Response to Past Returns



The table reports coefficients and 95% confidence intervals for the interaction term  $r_{j,\tau,t} \times \text{paired}_{j,t}$  in Equation (5.3). The horizontal axis shows the lags of returns ( $\tau$ ) in months.

**Realised Volatility** Column (4) of Table 5.3 adds the interaction term  $\text{volatility} \times \text{paired}$  to the main regression (Equation (5.1)). The coefficient of the interaction term is significantly positive. Investors conduct more tax-loss harvesting trades during volatile periods, which is consistent with [Chaudhuri et al. \(2020\)](#). The coefficient of  $\text{volatility} \times \text{paired}$  is much higher than the coefficient of  $\text{volatility}$ . If the realised volatility of the unpaired ETF increases by 1%, the trading volume goes up by 0.28% of AUM. However, for an ETF highly correlated with others, the 1% increase in volatility translates into a 7.81% higher relative trading volume.

**Asymmetric Return-Volume Connection** Karpoff (1987) and followers discover that the return-volume connection is asymmetric. If the return is positive, it is positively associated with volume. If the return is negative, it is negatively associated with volume, but the slope is flatter. Tax-loss harvesting may also appear asymmetric in response to past returns. If past returns are positive, not many investors have unrealised losses, so the response of trading volume should be close to 0. If past returns are negative, many investors have unrealised losses, and the slope should be more profound.

Table 5.4 reports the results. The regression for Column (1) replaces the last month's return ( $r_{j,1,t}$ ) in Equation (5.1) with two censored variables—last month's positive return ( $r_{j,1,t}^+$ ) and last month's negative return ( $r_{j,1,t}^-$ ). The coefficient of  $r_{j,1,t}^+$  is 1.379, significantly positive, while that of  $r_{j,1,t}^-$  is -1.002, significantly negative. More importantly, the summation of the two variables is statistically different from 0, showing the asymmetric return-volume connection and confirming the findings in the literature.

The regression in Column (2) includes two interaction terms ( $r_{j,1,t}^+ \times \textit{paired}$  and  $r_{j,1,t}^- \times \textit{paired}$ ) to capture the tax-loss trading volume. The coefficient of  $r_{j,1,t}^- \times \textit{paired}$  is -1.545. A 10% negative return is associated with a 15.45% of AUM increase in trading volume. The coefficient is more pronounced than that of  $r_1^-$  (-0.671). Surprisingly, the coefficient of  $r_{j,1,t}^+ \times \textit{paired}$  is significantly positive, although the magnitude is small (0.376). The positive coefficient is a result of missing variable bias. After controlling for the return between 12 and 2 months ago, the coefficient becomes negative. (Column (4)) Nevertheless, the coefficients of  $r_{j,1,t}^+ \times \textit{paired}$  in both columns are close to 0, confirming the previous speculation.

Column (3) adds positive and negative returns from 12 to 2 months ago ( $r_{j,2-12,t}^+$  and  $r_{j,2-12,t}^-$ ) to the regression. The coefficient of  $r_{j,2-12,t}^+$  is 0.504, much lower than that of  $r_{j,1,t}^+$  (1.353). The coefficient of  $r_{j,2-12,t}^-$  is also lower than that of  $r_{j,1,t}^-$  in terms of absolute magnitude. The results show that investors are more sensitive to recent returns, consistent with Column (2) in Table 5.3.

Column (4) includes interaction terms for all four censored past returns. The coefficient of  $r_{j,1,t}^+ \times \textit{paired}$  is -0.256. A 10% positive return in the past month results in a 2.56% less relative trading volume for paired ETFs compared with the unpaired ones. The coefficient of  $r_{j,1,t}^- \times \textit{paired}$  is -1.078, implying that a 10% negative return in the past month results in a 10.78% more relative trading volume for paired ETFs compared with the unpaired ones. The impact of the negative return is much more substantial than that of the positive return.

Table 5.4: Asymmetric Response of Trading Volume to Past Returns

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>paired</i>	0.090*** (0.008)	0.058*** (0.009)	0.099*** (0.008)	0.072*** (0.009)
$r_1^+$	1.379*** (0.056)	1.296*** (0.059)	1.353*** (0.056)	1.372*** (0.060)
$r_1^+ \times \textit{paired}$		0.376*** (0.102)		-0.256** (0.108)
$r_1^-$	-1.002*** (0.064)	-0.671*** (0.068)	-0.994*** (0.064)	-0.743*** (0.068)
$r_1^- \times \textit{paired}$		-1.545*** (0.105)		-1.078*** (0.107)
$r_{2-12}^+$			0.504*** (0.016)	0.528*** (0.017)
$r_{2-12}^+ \times \textit{paired}$				-0.098*** (0.028)
$r_{2-12}^-$			-0.229*** (0.021)	-0.131*** (0.022)
$r_{2-12}^- \times \textit{paired}$				-0.719*** (0.041)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.083	0.084	0.088	0.091

The table reports results for regressions with different specifications of past returns. "+" and "-" signs represent positive and negative past returns, respectively. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

Coefficients of  $r_{j,2-12,t}^+ \times \textit{paired}$  and  $r_{j,2-12,t}^- \times \textit{paired}$  are -0.098 and -0.719, respectively. Tax-loss harvesting is more sensitive to recent returns than old ones, no matter whether they are positive or negative. The impact of positive returns 12 to 2 months ago is close to 0, much lower than negative returns.

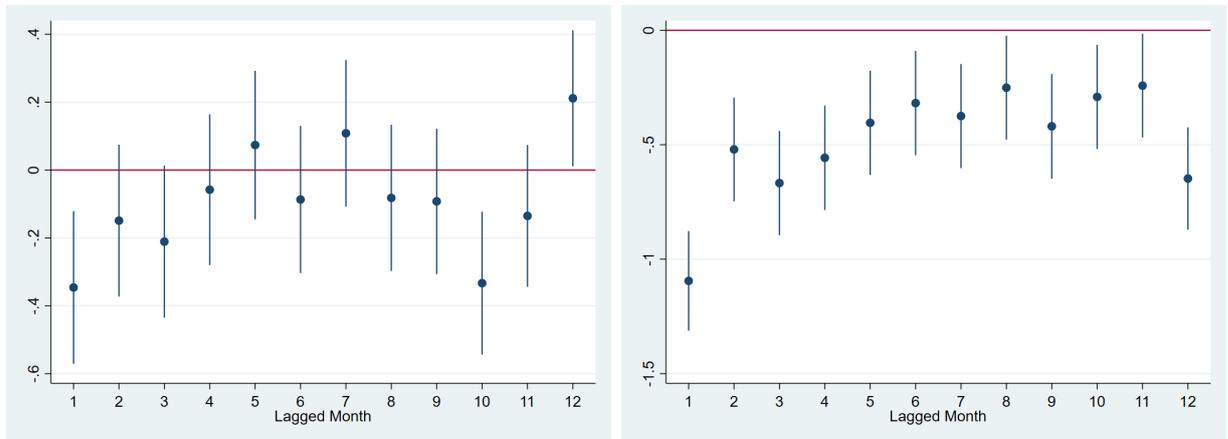
In conclusion, the results in Table 5.4 show that tax-loss harvesting has asymmetric responses to past returns. Positive returns have minor impacts on tax-loss trading volumes, while negative returns strongly negatively explain tax-loss trading volumes.

This paper delves further into the asymmetric tax-loss harvesting response to returns with lags from 1 month to 12 months. The regression is as follows,

$$\begin{aligned}
 volume_{j,t} = & \beta_0 + \beta_1 sim_{j,t} + \sum_{\tau=1}^{12} \theta_{\tau}^+ r_{j,\tau,t}^+ + \sum_{\tau=1}^{12} \eta_{\tau}^+ r_{j,\tau,t}^+ \times sim_{j,t} + \sum_{\tau=1}^{12} \theta_{\tau}^- r_{j,\tau,t}^- \\
 & + \sum_{\tau=1}^{12} \eta_{\tau}^- r_{j,\tau,t}^- \times sim_{j,t} + \beta_2 no\_sim_{j,t} + \beta_4 volatility_{j,t} + \beta_5 r_{m,t} + \beta_6 spread_{j,t} \\
 & + \beta_7 size_{j,t} + \beta_8 fee_{j,t} + \gamma_t + \delta_j + \epsilon_{j,t}
 \end{aligned} \tag{5.4}$$

Coefficients of the interaction terms ( $\eta_{\tau}^+$  and  $\eta_{\tau}^-$ ) are plotted in Figure 5.5. Bars correspond to 95% confidence intervals. Panel (a) shows tax-loss trading volume in response to past positive returns. Of the 12 lags, 9 have insignificant coefficients at the 5% level. Positive returns 1 and 10 months ago have significantly negative coefficients (-0.346 and -0.333), while the positive return 12 months ago positively explains the tax-loss trading volume with a coefficient of 0.212. The coefficients are small. To summarise, positive past returns have close to 0 impacts on the tax-loss harvesting behaviour.

Figure 5.5: Asymmetric Response of Tax-Loss Trading Volume



(a) Loading of Tax-Loss Trading Volume on Positive Return

(b) Loading of Tax-Loss Trading Volume on Negative Return

Panel (a) in this figure reports coefficients and 95% confidence intervals for the interaction term  $r_{j,\tau,t}^+ \times paired_{j,t}$  in Equation (5.3); panel (b) reports coefficients and 95% confidence intervals for the interaction term  $r_{j,\tau,t}^- \times paired_{j,t}$ . The horizontal axis shows the lags of returns ( $\tau$ ) in months.

Panel (b) plots the coefficients for past negative returns. Unlike Panel (a), coefficients for all 12 lags are significantly negative. The coefficient for last month's negative return is -1.095, implying that a 10% negative return in the last month predicts a 10.95% higher relative trading volume for paired ETFs than unpaired ones. The magnitude is much higher

than  $\eta_1^+$ . As the lag increases,  $\eta_\tau^-$  shows a converging pattern towards 0. The coefficient for the negative return 11 months ago is -0.242, 22% of the magnitude as the negative return 1 month ago.  $\eta_{12}^-$  deviates from the decreasing pattern with a higher absolute magnitude than  $\eta_{11}^-$ .

Results in Figure 5.5 show that the pattern in Figure 5.3 is driven by tax-loss trading volume in response to negative past returns, not positive ones. Investors are more sensitive to recent negative returns than remote negative returns, but they do not exhibit a similar pattern for positive returns.

More importantly, the figures provide direct evidence related to tax-loss harvesting. To begin with, coefficients of positive returns are close to 0, while those of negative ones are all significantly negative. Investors are more sensitive to negative returns than positive returns because tax-loss harvesting relies on unrealised losses. Secondly, the decreasing pattern in Panel (b) is consistent with tax codes. Short-term capital gains are taxed at a higher rate (up to 37%) than long-term ones (up to 20%), meaning losses are more valuable if realised early. Finally, the deviation of  $\eta_{12}^-$  from the decreasing pattern coincides with the cutoff between long- and short-term capital gains—12 months. Investors have the incentive to sell losing securities before they become long-term assets. [Constantinides \(1984\)](#) shows that the real option of realising losses short-term is "extremely valuable".

## D. Heterogeneous Analysis

Fund characters are expected to correlate with the tax-loss harvesting behaviour. The paper includes interaction terms of *paired* and fund characters in Equation (5.1) in regressions, finding that the fund liquidity is associated with the tax-loss trading volume.

Results are reported in Table 5.5. Column (1) is the main regression, the same as Column (3) in Table 5.2. Coefficients of *size*, *spread*, and *fee* are significantly negative. Since *volume* is measured as the proportion of shares exchanged in the market, it is intuitive that large ETFs have relatively small *volume*. A 10% increase in the asset under management is associated with a lower relative (higher absolute) trading volume of 1.34% (8.53%). The bid-ask spread is an implicit transaction cost that discourages investors from trading. A 1% increase in the bid-ask spread decreases the trading volume by 1.15% of the AUM. A high expense ratio raises the cost of holding the ETF, making the ETF less attractive to investors, so the trading volume is lower. A 10 bps increase in the expense

ratio is associated with a 4.69% lower relative trading volume.

Table 5.5: Tax-Loss Trading Volume and Fund Characters

Dependent Variable: <i>volume</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
<i>paired</i>	0.091*** (0.008)	0.949*** (0.054)	0.096*** (0.008)	0.034** (0.014)	0.949*** (0.061)	1.020*** (0.062)
<i>size</i>	-0.134*** (0.002)	-0.122*** (0.002)	-0.134*** (0.002)	-0.133*** (0.002)	-0.122*** (0.002)	-0.121*** (0.002)
<i>size</i> × <i>paired</i>		-0.043*** (0.003)			-0.043*** (0.003)	-0.046*** (0.003)
<i>spread</i>	-1.153*** (0.195)	-1.094*** (0.195)	-1.058*** (0.197)	-1.137*** (0.195)	-1.094*** (0.195)	-0.910*** (0.197)
<i>spread</i> × <i>paired</i>			-3.866*** (1.131)			-7.343*** (1.148)
<i>fee</i>	-46.867*** (2.639)	-46.654*** (2.637)	-47.064*** (2.639)	-49.637*** (2.698)	-46.637*** (2.703)	-47.026*** (2.704)
<i>fee</i> × <i>paired</i>				15.076*** (3.064)	-0.094 (3.217)	0.074 (3.217)
Control	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779	183779	183779
$R^2$	0.080	0.081	0.080	0.080	0.081	0.082

The table reports results for regressions with different specifications of interaction terms. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

Column (2) includes the independent variable *size* × *paired*, which has a significantly negative coefficient. It does not mean that investors make fewer tax-loss trades with large ETFs. Since *volume* is a relative definition, the coefficient should be compared against -1 instead of 0. The interpretation should be that investors make less tax-loss trades proportional to AUM with large ETFs. A 10% increase in the asset under management decreases the relative tax-loss harvesting trading volume by 0.43%, but the overall effect is an increase in the absolute tax-loss trading volume by 9.53%.

Column (3) includes *spread* × *paired* in the regression. The coefficient is negative with significance at the 1% level. Since spread is an implicit transaction cost, it is clear that a high spread reduces tax-loss harvesting. A 1% increase in the bid-ask spread decreases tax-loss trading volume by 1.09% of the AUM. The absolute magnitude of the *spread* × *paired* coefficient is larger than that of *spread*. Compared with regular trades, investors are more sensitive to the spread when making tax-loss trades. The intuition is that tax-loss harvesting involves two trades simultaneously, making the impact more pronounced.

Column (4) adds the interaction term  $fee \times paired$  to the regression. The coefficient is significantly positive. However, after controlling for size, the coefficient becomes insignificant. (Column 5) The negative correlation between  $fee$  and  $size$  can explain the empirical results. The insignificant coefficient is in line with the expectation. Tax-loss harvesting is about trading, not holding ETFs. Variables such as  $spread$  are related to trading, so their interaction terms are significant. Since the expense fee occurs when holding the ETF, it should only affect the total trading volume, not the tax-loss trading volume.

To avoid the omitting variable bias, Column (6) reports the regression with all the independent variables used in Columns (1) to (5). The results above are preserved qualitatively.

## E. Market Incentive

This paper also delves into market incentives on tax-loss trading behaviour, including interest rates and months in the year. Interest rates play two roles in tax-loss harvesting decision-making. On the one hand, since the tax benefit from loss-selling is the tax payment deferral, nominal risk-free interest rates should be positively correlated with tax-loss harvesting. On the other hand, interest rates reflect the future monetary tightness. With positive transaction costs, investors reduce tax tradings when the fund is tight. This paper uses the medium-to-long-term nominal risk-free interest rate ( $i_h$ ) and the fed fund rate ( $i_{FF}$ ) to capture the two effects. This paper also explores whether there is an excessive trading volume in December.

Table 5.6 reports regression results with different specifications of market incentives. Column (1) includes the one-year risk-free rate and the fed fund rate in regression (5.1). The coefficient of  $i_1$  is 0.180, significantly positive. A 10% increase in the one-year risk-free rate raises the average trading volume by 1.8% of AUM. The coefficient of  $i_{FF}$  is -0.395, significantly negative, showing that the short-run monetary tightness decreases tradings in general. A 10% increase in the fed fund rate reduces the average trading volume by 3.95% of AUM.

To investigate the impacts on tax-loss harvesting, Column (2) includes interest rates and their interaction terms with  $paired$  in the regression. Coefficients of  $i_1$  and  $i_{FF}$  are significant and close to those in Column (1). The coefficient of  $i_1 \times paired$  is 0.075, significantly positive, meaning that the tax-loss trading volume is positively associated with the one-year interest rate. A 10% rise in  $i_1$  increases the tax-loss trading volume by 0.75% of

Table 5.6: Tax-Loss Trading Volume and Market Conditions

Dependent Variable: <i>volume</i>						
Interest Rate Horizon	(1)	(2)	(3)	(4)	(5)	(6)
$h =$	1 Year	1 Year	5 Years	10 Years	$PC_1$	$PC_1$
$i_h$	0.180*** (0.056)	0.153*** (0.056)	0.031 (0.120)	-0.079 (0.235)	0.089 (0.328)	0.104 (0.328)
$i_h \times paired$		0.075*** (0.010)	0.019*** (0.005)	0.011*** (0.004)	0.041*** (0.011)	
$i_{FF}$	-0.395*** (0.069)	-0.363*** (0.069)	-0.228* (0.128)	-0.115 (0.246)	-0.243 (0.177)	-0.254 (0.177)
$i_{FF} \times paired$		-0.085*** (0.010)	-0.026*** (0.005)	-0.018*** (0.004)	-0.027*** (0.005)	
<i>December</i>						-0.014 (0.123)
<i>December</i> $\times$ <i>paired</i>						0.006 (0.012)
Control	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes	Yes	Yes
N	183779	183779	182205	182205	182205	182205
$R^2$	0.080	0.080	0.081	0.081	0.081	0.080

The table reports results for regressions with different specifications of risk-free interest rates and their interaction terms with *paired*.  $PC_1$  refers to the first principal component of risk-free interest rates between 1 and 10 years. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

AUM. On the contrary, the coefficient of  $i_{FF} \times paired$  is significantly negative (-0.085). A 10% increase in the fed fund rate reduces the tax-loss trading volume by 0.85% of AUM. Columns (3) to (5) explore interest rates with 5- and 10-year horizons and the first principal component of risk-free rates between 1 and 10 years ( $PC_1$ ). All the interaction terms are significant and have the same signs as in Column (2).

Column (6) investigates the turn-of-year effect by including a dummy variable for December and the interaction term *December*  $\times$  *paired*. The coefficient of *December* is significantly positive, showing that trading volume increases in December. The coefficient of the interaction term is small (0.6% of AUM) and insignificant. This paper does not find evidence of more intensive tax-loss harvesting with ETFs in December.

However, it does not mean the result is inconsistent with the literature. Although [Poterba and Weisbenner \(2001\)](#) and [Ivković et al. \(2005\)](#) find evidence of tax-loss harvesting with stocks in December, stocks differ greatly from ETFs in tax-loss harvesting. The trading strategy in this paper involves highly correlated ETFs with correlation coefficients higher

than 99%, meaning that there is only a minimal market risk to switch to another ETF. On the contrary, it is almost impossible to find highly correlated stocks without violating the wash-sale rule. As a result, stock investors prefer to make the tax-loss decision later in the year to avoid unnecessary changes to their portfolios, but not ETF investors.

Besides, since capital losses can be carried forward forever, there is no discount to realise losses early. With a low bid-ask spread, it is optimal for investors to harvest losses as soon as they present because the opportunity could vanish if the price increases. This idea is in line with the practice of the investment industry. For instance, Vestmark claims it takes "opportunities to capture tax losses when they present themselves."<sup>4</sup>

## F. Robustness Check

This research conducts several robustness tests, including adopting robust DID estimators, exploring an alternative definition of similar ETFs, varying the correlation threshold, running a placebo test, and checking the impact of winsorisation.

**Robust DID** The difference-in-difference setting with ETF and time fixed effects provides a straightforward estimate of the average treatment effect on the treated (ATT). However, the DID estimator with two-way fixed effects (TWFE) can be biased if (1) there are multiple periods and (2) treatments are assigned in different periods. By decomposing the TWFE DID estimator, [Goodman-Bacon \(2021\)](#) finds that the estimator is generally biased if either the parallel assumption fails or treatment effects are not constant. The rationale is that TWFE includes treated observations earlier in the sample in the control group for events later in the sample. [Baker et al. \(2022\)](#) argue that the bias of TWFE DID prevails in the finance literature.

The regressions in this paper also suffer from this problem. However, the bias should be small. On the one hand, Figure 5.3 shows that the parallel trend assumption holds, eliminating one source of bias. On the other hand, there are 2511 ETFs in the sample, but only 298 events are identified. The majority of the control group is either not treated or not yet treated. Besides, as shown in Figure 5.2, there is an increasing trend of events per period, further reducing the proportion of observations that are treated and included in the control group.

---

<sup>4</sup>[Vestmark's Tax Management Capabilities](#), link.

To better understand the potential bias of the TWFE estimator, this paper also employs DID estimators that are robust to the staggered DID setting. Many propensity-score-based robust DID estimators have been proposed since the discovery of the problem. (For example, [Arkhangelsky et al. \(2021\)](#).) This paper adopts a widely used robust DID estimator developed by [Callaway and Sant’Anna \(2021\)](#), which contains three steps—(1) identification of disaggregated causal parameters by treatment groups, (2) aggregation of these parameters to form summary measures, and (3) estimation and inference. Under appropriate assumptions, the robust DID estimator is unbiased.

However, the robust estimator has its problems as well. The second step (aggregation) requires estimating the propensity scores of being treated with the control variables. Unfortunately, the estimation of propensity scores is very imprecise in this research. The reason is that launching a new ETF (treatment) is pre-determined. It takes at least several months<sup>5</sup> to go through administration processes before launching a new ETF. As a result, the control variables in Regression (5.1) have little power to predict propensity scores. In addition, there are many treatment groups (102) compared to the number of events (298), making the first step (identification) estimation very imprecise. For these reasons, this method is only used for robustness checks.

There are several different ways to aggregate the ATT of individual groups. Each comes with different model assumptions and properties. ([Callaway and Sant’Anna 2021](#)) This paper considers three aggregation methods that are robust to misspecification of the formula to estimate propensity scores—inverse probability tilting ([Graham et al. 2012](#)), outcome regressions ([Heckman et al. 1997](#)), and doubly robust DID ([Sant’Anna and Zhao 2020](#)).

Table 5.7 reports the ATT with different robust DID estimators. Columns (1) to (3) use only never-treated ETFs as the control group; columns (4) to (6) include both never-treated and not-yet-treated ETFs in the control group. All six estimates of ATT are positive and comparable with the TWFE DID estimate in Table 5.2 (0.091) in terms of magnitude. The estimated ATT with never-treated ETFs are close to those with not-yet-treated ETFs. The result implies that the treated sample does not affect the estimation of ATT much.

However, since propensity scores are estimated imprecisely, ATT estimates with dif-

---

<sup>5</sup>The length of launching an ETF varies case by case. The ETF sponsor needs a period to develop a viable investment strategy. Then, the SEC approves or disapproves the new ETF in 45 to 90 days. After the approval, it takes several weeks for the ETF to be listed on an exchange. More details can be found on the ICI website.

Table 5.7: ATT with Robust DID Estimators

Dependent Variable: <i>volume</i>						
Aggregation Method	(1)	(2)	(3)	(4)	(5)	(6)
	Never Treated			Not Yet Treated		
	IPT	REG	DR	IPT	REG	DR
ATT	0.389*** (0.131)	0.169 (—)	0.037 (—)	0.435*** (0.163)	0.168 (—)	0.037 (—)
Control	Yes	Yes	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779	183779	183779

The table reports average treatment effects on the treated (ATT) with [Callaway and Sant’Anna \(2021\)](#) robust DID estimator and different aggregation weighting methods. Columns (1) to (3) use only never-treated ETFs as the control group; columns (4) to (6) include both never-treated and not-yet-treated ETFs in the control group. IPT refers to the inverse probability tilting method of aggregation weights ([Graham et al. 2012](#)); REG represents using the weighting method based on outcome regressions ([Heckman et al. 1997](#)); and DR means doubly robust DID method ([Sant’Anna and Zhao 2020](#)). Standard errors are clustered by ETF and reported in parentheses. (—) means that the standard error cannot be properly estimated. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

ferent aggregation methods are very distinctive from each other. Using only never-treated ETFs as the control group, aggregation with inverse probability tilting (IPT) method yields an ATT of 0.389, while aggregation methods based on outcome regressions (REG) and doubly robust DID (DR) report the ATT of 0.169 and 0.037, respectively. Besides, aggregation methods based on outcome regressions and doubly robust DID cannot properly estimate standard errors. IPT is the only one that can estimate the standard error, which shows that the ATT is significantly different from 0 at the 1% level.

To conclude, the robust DID estimator should not be used to estimate the ATT from introducing the first similar ETF because propensity scores cannot be properly estimated. However, the results indicate that the ATT is significantly positive, and the TWFE DID estimate has a correct magnitude.

**Alternative Definition** Investors may be very sensitive to the mismatching risk when harvesting the loss. In extreme cases, they may not trade S&P 500 and Russell 1000 ETFs for tax-loss harvesting, even though they have a correlation coefficient of 0.9967. However, they can always use two ETFs that track the same index to conduct tax-loss harvesting. For example, investors can trade Vanguard S&P 500 ETF (ticker: VOO) and SPDR S&P 500 ETF (SPY) simultaneously. To test this hypothesis, the paper defines the treatment in an alternative way and re-runs DID regressions. The price-volume connection is also

explored.

Instead of using the introduction of the first similar ETF as the treatment, this part defines the treatment as the introduction of the first ETF that tracks the same index. The variable *same* takes the value of 1 if there is at least one other ETF that tracks the same index. Otherwise, *same* = 0. The variable *same* serves as the difference-in-difference term in regressions. The regression equation is as follows,

$$\begin{aligned} volume_{j,t} = & \beta_0 + \beta_1 same_{j,t} + \beta_2 no\_same_{j,t} + \beta_3 r_{j,1,t} + \beta_4 volatility_{j,t} \\ & + \beta_5 r_{m,t} + \beta_6 spread_{j,t} + \beta_7 size_{j,t} + \beta_8 fee_{j,t} + \gamma_t + \delta_j + \epsilon_{j,t} \end{aligned} \quad (5.5)$$

where  $no\_same_j$  is the total number of ETFs tracking the same index as ETF  $j$ , and other variables are defined the same as Equation (5.1).

Table 5.8 Column (1) reports the regression result of Equation (5.5). The coefficient of the DID variable is 0.001, which is insignificant and close to 0. There is no difference before and after the introduction of the first ETF that tracks the same index. The result is not contradictory to the findings in Table 5.2. By definition, an observation with *same* = 1 must have *paired* = 1, but the reverse is not necessarily true. Some tax-loss harvesting opportunities with small mismatching risks are classified into the group *same* = 0, making the difference in trading volume between *same* = 0 and *same* = 1 small. Even though investors can use ETFs tracking the same index to harvest loss, the "treatment" is too late to capture the change in trading volume. The result reveals that investors exploit the tax-loss harvesting opportunity with highly correlated ETFs.

Column (2) includes the past return 12 to 2 months ago in the regression. The coefficients for  $r_1$  and  $r_{2-12}$  are 0.329 and 0.184, respectively. The point estimates are close to regressions with *paired* (0.332 and 0.189). The difference between the two coefficients is significant, showing again that *volume* is more sensitive to recent returns.

Column (3) adds two interaction terms to the regression,  $r_1 \times same$  and  $r_{2-12} \times same$ , capturing the tax-loss trading volume. The coefficient of  $r_1 \times same$  is -0.216, insignificant and much less than -0.606 in Table 5.2 in terms of the absolute value. The coefficient of  $r_{2-12} \times same$  is -0.154, significant at the 1% level, but it is still less pronounced than -0.325 in Table 5.2. The reason for only  $r_{2-12} \times same$  being significant is the power of test.  $r_{2-12}$  is the sum of returns in 11 months, which has a better power of test and a lower

Table 5.8: Regression Results with Alternative Definition of Treatment

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>same</i>	0.001 (0.019)	0.000 (0.019)	0.017 (0.020)	-0.026 (0.019)
$r_1$	0.315*** (0.034)	0.329*** (0.034)	0.335*** (0.034)	0.318*** (0.034)
$r_1 \times same$			-0.216 (0.144)	
$r_{2-12}$		0.184*** (0.011)	0.188*** (0.011)	
$r_{2-12} \times same$			-0.154*** (0.048)	
<i>volatility</i>	0.362*** (0.044)	0.360*** (0.044)	0.360*** (0.044)	0.357*** (0.044)
<i>volatility</i> $\times$ <i>same</i>				7.825*** (0.906)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.080	0.081	0.081	0.080

The table reports results for regressions with different specifications of past returns. The treatment is defined as the introduction of the first new ETF that tracks the same index as the incumbent one instead of the introduction of the first similar ETF. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

standard error than  $r_1$  by construction. To conclude, the regression exhibits a negative link between tax-loss harvesting and past return, but the evidence is not as clear as using highly correlated ETFs.

Column (4) investigates the impact of realised volatility on the tax-loss trading volume by including the interaction term between *volatility* and *same*. The coefficient of the interaction term is 7.825, significantly positive and close to 7.539 in Table 5.2. It is much higher than the coefficient of *volatility* (0.357), showing that the tax-loss harvesting trading volume is more sensitive to *volatility* than the overall trading volume. Using *same* instead of *paired* does not change the result qualitatively.

To summarise, this part of the research explores the impact of using ETFs tracking the same index as the identification strategy. With ETFs tracking the same index, the major findings are preserved qualitatively, but the evidence of tax-loss harvesting is less clear. It means that investors commit to tax-loss harvesting with highly correlated ETFs, not just ETFs tracking the same index.

**Correlation Threshold** The definition of *paired* relies on a seemingly arbitrary threshold—ETFs  $j$  and  $k$  form a pair if  $\text{corr}(r_j, r_k) > 0.99$ . If the threshold is too high, many tax-loss harvesting trades are not included in the post-treatment group. If the threshold is too low, there can be more confounding observations. This subsection explores the impact of using different thresholds. Similar to  $\text{paired}_{j,t}$ ,  $\text{paired95}_{j,t} = 1$  if another ETF has a correlation coefficient with ETF  $j$  higher than 0.95. Same as before, they need to have at least 12 overlapping observations.  $\text{paired995}_{j,t}$  is defined similarly with a correlation coefficient higher than 0.995.

Table 5.9 shows the results with alternative definitions of *paired*. Column (1) is the main regression. Columns (2) and (3) report the tax-loss trading volumes estimated with *paired95* and *paired995*. *paired95* tends to include more ETF pairs than *paired*, some of which may not be suitable substitutes. The coefficient of *paired95* is 0.103, statistically indifferent from that of *paired*. Using a lower threshold does not help to identify more tax-loss harvesting behaviour. *paired995* excludes some ETF pairs that have similar returns but do not follow the same index. Its coefficient is 0.065, significantly greater than 0 but lower than 0.091 for *paired*. The result indicates that some tax-loss trades are conducted with ETFs that have a correlation coefficient of around 0.99. The choice of threshold in the same research is appropriate. To conclude, the result is robust to the definition of ETF pair.

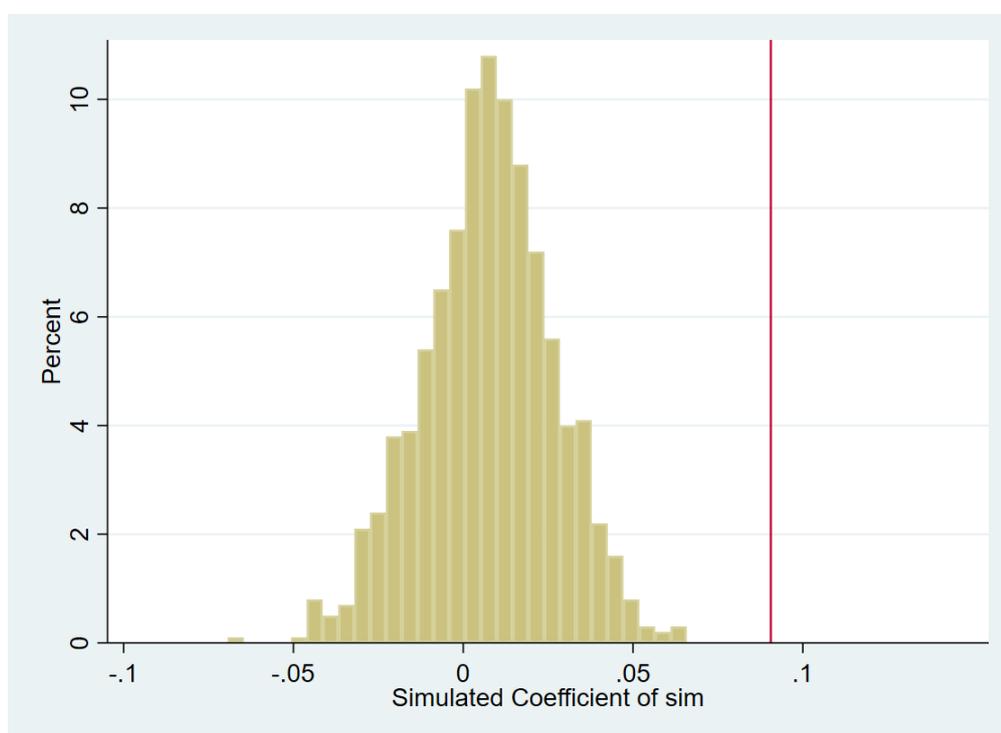
Table 5.9: Robustness Checks

Dependent Variable: volume					
	(1)	(2)	(3)	(4)	(5)
<i>paired</i>	0.091*** (0.008)			0.104*** (0.007)	0.071*** (0.025)
<i>paired95</i>		0.103*** (0.007)			
<i>paired995</i>			0.065*** (0.010)		
Control	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes	Yes
N	183779	183779	183779	183054	183779
$R^2$	0.080	0.080	0.080	0.074	0.025

The table reports robustness check results. Column (1) is the main regression. (Equation (5.1)) Columns (2) and (3) use different definitions of *paired*. Column (4) discards observations with top 1% *volume* instead of winsorising. Column (5) uses the sample without winsorising *volume*. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

**Placebo Test** A non-parametric placebo test is used to rule out the hypothesis of random treatment. Acknowledging the asynchronicity of the introduction of the first similar ETF, the research randomly assigns the treatment to every ETF. The coefficient of *paired* is estimated from Equation (5.1). The process repeats 1000 times so that the approximate distribution of the estimated coefficient is obtained. Brown bars in Figure 5.6 show the distribution, while the red line is the point estimate from the sample. All brown bars are around 0 and to the left of the red line. In other words, the estimated coefficient of *paired* has a  $p$ -value of 0.000, and the treatments are not random.

Figure 5.6: Frequency Histogram of Placebo Test



The placebo test randomly assigns treatments to ETFs 1000 times. The figure shows the frequency histogram of the simulated coefficient of *paired* from the main regression. (Equation (5.1)) The red line is the point estimate from the sample.

**Winsorisation** This paper also examines the impact of winsorising the trading volume. Column (4) shows the regression that discards observations with top 1% *volume* instead of winsorising. The coefficient of *paired* is 0.104, qualitatively the same as in Column (1). Column (5) uses the original variable *volume* without winsorising. The result is subject to the impact of outliers. The  $R^2$  statistic drops to 0.025 from 0.080. The coefficient of *paired* is 0.071, significantly different from 0 at the 1% level. In conclusion, winsorising the relative trading volume does not qualitatively change the result.

## 5.4 Tax-Loss Harvesting Model

This section presents a simple equilibrium model to explain the connection between tax-loss trading volume and price movements. The existence of capital gain tax brings challenges to models. On the one hand, the agent can make tax-loss harvesting decisions in every trading period, and the tax-deferral benefit needs to be calculated backwards from the ending period. Due to the high dimension of the choices, it is almost impossible to obtain a closed-form solution. On the other hand, the optimal tax-loss harvesting strategy requires recording the exact tax basis of every transaction so that securities with the highest tax basis can be sold first. Record keeping further increases the complexity of the problem, so it is not easy to solve the model numerically as well.

[Dammon et al. \(2001\)](#) build a model to investigate the optimal consumption and investment strategies with the capital gain tax. With a random life expectancy, agents maximise the expected utility of lifetime consumption and terminal wealth. There are two assets in the model—risky and risk-free assets. Short sales are not allowed. The model contains many decision-making periods, from 20 to 100 years old. To keep the model trackable, the paper chooses to use the moving average tax basis of the asset instead of the exact tax basis. The model is solved numerically. [Dammon et al. \(2004\)](#) update the model of [Dammon et al. \(2001\)](#) to allow the co-existence of two accounts—taxable and tax-deferral accounts. Short sales are permitted. They find a strong tendency to invest the equity in the taxable account and the taxable bond in the tax-deferral account. Tax consideration outweighs the liquidity requirement.

Using a model similar to Dammon et al., [Garlappi et al. \(2001\)](#) look at the value of tax-loss harvesting. There are two risky assets and one risk-free asset in the model. They argue that the tax-deferral feature of securities can be viewed as an option, with a value of 5-10% of the wealth. The option value decreases with volatility, risk aversion, and correlation between risk assets. [DeMiguel and Uppal \(2005\)](#) use non-linear programming to solve for the optimal portfolio with the exact tax basis. Due to the high dimensionality problem, the paper only simulates the model with up to 10 periods and 2 assets. Short sales are not allowed in the model. They find that the tax burden can be as low as 1% with the optimal tax-loss harvesting strategy. The paper also compares the welfare loss by approximating the tax basis with the moving average, finding that the certainty equivalence loss is 1%.

## A. Model Setup

The tax-loss harvesting model in this paper contains a continuum of agents with different endowments and tax bases. Each agent  $i$  maximises the expected utility of the ending wealth ( $W^i$ ). There is no consumption in the setting. Agents have constant absolute risk aversion (CARA) utility with the same coefficient of risk aversion  $\lambda = 1$ .

The model has only two periods for parsimony, but it can explain the empirical findings well. There are three assets in the model—the tradeable risky asset, the non-tradeable risky asset, and the risk-free asset. The risk-free asset has an exogenous return  $r_f$ . In the first period, only the non-tradeable asset pays a dividend ( $H_1$  for each share); in the second period, both risky assets pay liquidating dividends ( $H_2$  and  $D_2$  for each share). Dividends of the two risky assets are correlated and normally distributed. Mathematically,

$$\begin{pmatrix} D_t \\ H_t \end{pmatrix} \sim N \left( \begin{pmatrix} \bar{D} \\ \bar{H} \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & \sigma_D \sigma_H \rho \\ \sigma_D \sigma_H \rho & \sigma_H^2 \end{pmatrix} \right) \quad (5.6)$$

where  $t = 2$  in the specific setting.

Each agent is assigned an initial endowment with the two risky assets. The endowment of the non-tradeable follows a normal distribution  $Z^i \sim N(\bar{Z}, \sigma_Z^2)$ . To track the tax basis of investor  $i$ , the model adopts the stock variable  $N_{s,t}^i$  to represent the holding in period  $t$  of the tradeable asset bought in period  $s$ . For example,  $N_{0,1}^i$  is the amount of the tradable asset bought in period 0 (initial assignment) that remains in  $i$ 's investment account in the first period. The initial endowment follows a normal distribution,  $N_{0,0}^i \sim N(\bar{N}, \sigma_N^2)$ .

Agents have different tax bases for the tradeable asset. In practice, if the past return volatility of a security is high, the distribution of tax basis is wide. The reason is that prices are determined by transactions in the market. This paper controls the standard deviation of the tax basis to reflect past volatility. The past return can be controlled by changing the mean of the tax basis. The initial tax basis of investor  $i$  follows

$$\frac{P_0^i}{1 + s/2} \sim N(P_0, \sigma_P^2) \quad (5.7)$$

where  $s$  is the bid-ask spread, and  $P_0$  and  $\sigma_P^2$  are tuning parameters.

In period 1, agents make decisions on their portfolios. Agents are allowed to borrow or lend money at the risk-free rate  $r_f$ . They choose to sell  $N_{0,0}^i - N_{0,1}^i$  units and buy  $N_{1,1}^i$

units of the tradeable asset. The model does not allow short sales, so the choice variables must satisfy the following condition,

$$\forall s \leq t, 0 \leq N_{s,t}^i \leq N_{s,t-1}^i \leq \dots \leq N_{s,s}^i \quad (5.8)$$

In the 2-period case,  $s = 0, 1, 2$  and  $t = 1, 2$ . Agents must pay the tax for the dividend and the realised capital gain at the same tax rate  $\tau$ .

If the wash-sale rule applies,  $N_{s,t-1}^i - N_{s,t}^i$  and  $N_{t,t}^i$  cannot be greater than 0 at the same time. Mathematically, the condition can be written as follows.

$$\sum_{s=0}^{t-1} (N_{s,t-1}^i - N_{s,t}^i) \times N_{t,t}^i = 0 \quad (5.9)$$

The condition simplifies to  $(N_{0,0}^i - N_{0,1}^i) \times N_{1,1}^i = 0$  in period 1.

In period 2, agents liquidate all assets and consume all their wealth. Since no one lives beyond this period, assets have 0 carry-on values. They must pay the tax for dividends from both risky assets.

The market equilibrium determines the price of the tradeable asset in period 1 ( $P_1$ ). The equilibrium condition is

$$\sum_i \left[ \sum_{s=0}^{t-1} (N_{s,t-1}^i - N_{s,t}^i) \right] = \sum_i N_{t,t}^i \quad (5.10)$$

In period 1, the condition becomes  $\sum_i (N_{0,0}^i - N_{0,1}^i) = \sum_i N_{1,1}^i$ .

**Optimal Trading Strategy** Agents trade for two reasons—rebalancing the portfolio and tax-loss harvesting. Rebalancing trades occur regardless of the wash-sale rule, while tax-loss trades are possible only when the wash-sale rule does not apply. Therefore, this paper separately discusses two scenarios, with and without the wash-sale rule, and measures the tax-loss trading volume as the difference in trading volumes in the two scenarios.

When making tax-loss harvesting decisions, agents consider the trade-off between the benefit of deferring the capital gain tax and the transaction cost. Since there is no consumption in the model, agents maximise the ending wealth given the benefit and cost of tax-loss harvesting.

The discussion starts with the scenario where the wash-sale rule does not apply. Under

the current 2-period setting, agents only need to make decisions in period 1. Choice variables are  $N_{0,1}^i$  and  $N_{1,1}^i$ . The maximisation problem is

$$\begin{aligned} \max_{N_{0,1}^i, N_{1,1}^i} \mathbb{E}_1 \left[ -e^{-\lambda W^i} \right] \\ \text{s.t. } 0 \leq N_{0,1}^i \leq N_{0,0}^i \\ 0 \leq N_{1,1}^i \end{aligned} \quad (5.11)$$

The number 1 below the expectation sign means that the expectation is conditional on information up to period 1. The ending wealth  $W^i$  is defined below.

If the agent realises more loss than dividend income, she can have a tax credit to deduct future income. The tax credit in period 1 can be expressed as

$$TC_1^i = \max \left( (N_{0,0}^i - N_{0,1}^i)[P_0^i - P_1(1 - \frac{s}{2})] - H_1 Z^i, 0 \right) \quad (5.12)$$

where  $(N_{0,0}^i - N_{0,1}^i)$  is the number of assets sold in period 1,  $[P_0^i - P_1(1 - \frac{s}{2})]$  is the loss per share, and  $H_1 Z^i$  is the dividend income.

The amount of cash held by agent  $i$  in period 1 is

$$C_1^i = \begin{cases} (1 - \tau)H_1 Z^i + (N_{0,0}^i - N_{0,1}^i)[(1 - \tau)P_1(1 - \frac{s}{2}) + \tau P_0^i] - N_{1,1}^i P_1(1 + \frac{s}{2}) & \text{if } TC_1^i = 0 \\ H_1 Z^i + (N_{0,0}^i - N_{0,1}^i)P_1(1 - \frac{s}{2}) - N_{1,1}^i P_1(1 + \frac{s}{2}) & \text{if } TC_1^i > 0 \end{cases} \quad (5.13)$$

The first term  $H_1 Z^i$  is the dividend income, the second term is the cash from selling the tradeable asset, and the last term  $N_{1,1}^i P_1(1 + \frac{s}{2})$  is the payment for buying the tradeable asset. If  $TC_1^i > 0$ , the agent realises more loss than income, so she does not need to pay any tax in this period.

The ending wealth  $W^i$  comes from the following equation,

$$W^i = \min \left( (1 - \tau)[D_2(N_{0,1}^i + N_{1,1}^i) + H_2 Z^i] + \tau \left[ TC_1^i + N_{0,1}^i P_0^i + N_{1,1}^i P_1(1 + \frac{s}{2}) \right], \right. \\ \left. D_2(N_{0,1}^i + N_{1,1}^i) + H_2 Z^i \right) + C_1^i(1 + r_f) \quad (5.14)$$

The two terms in the  $\min(\cdot, \cdot)$  function represent scenarios with and without tax payment in period 2.

Uncertainty of the maximisation problem only comes from  $D_2$  and  $H_2$ , which are unknown in period 1. Agents make decisions as if  $P_1$  is observed on the market. The endogenous variable  $P_1$  is tuned to meet the market clearing condition. (Equation (5.10)) The trading volume is measured by either trading or selling volumes relative to the total number of asset shares.

$$V_t = \frac{\sum_i N_{t,t}^i}{\sum_i N_{0,0}^i} = \frac{\sum_i [\sum_{s=0}^{t-1} (N_{s,t-1}^i - N_{s,t}^i)]}{\sum_i N_{0,0}^i} \quad (5.15)$$

When the wash-sale rule applies, the optimal trading strategy remains unchanged except for the utility maximisation problem, which becomes

$$\begin{aligned} \max_{N_{0,1}^i, N_{1,1}^i} \mathbb{E}_1 \left[ -e^{-\lambda W^i} \right] \\ \text{s.t. } 0 \leq N_{0,1}^i \leq N_{0,0}^i \\ 0 \leq N_{1,1}^i \\ (N_{0,0}^i - N_{0,1}^i) \times N_{1,1}^i = 0 \end{aligned} \quad (5.16)$$

The additional condition means the number of shares bought and sold by investor  $i$  cannot be greater than 0 simultaneously.

The trading volume with the wash-sale rule is labelled  $V_t^{WS}$ . When an ETF starts to be paired with another one, it is equivalent to being exempt from the wash-sale rule. Therefore, the tax-loss trading volume is the difference between trading volumes without and with the wash-sale rule (excess trading volume).

$$V_t^{TLH} = V_t - V_t^{WS} \quad (5.17)$$

## B. Simulation

Agents in the model have different tax bases and, hence, different thresholds for engaging in tax-loss harvesting. This feature makes it difficult to solve the model analytically. Therefore, the model is solved by simulations. 1000 agents are randomly drawn to simulate the continuum. When calculating the expected utility, the simulation algorithm draws 1000 pairs of  $(D_2, H_2)'$  and calculates the ending wealth instead of solving for a closed-form solution. The algorithm separately simulates the model with and without the wash-sale

rule and calculates the excess trading volume.

**Parameter Choice** Table 5.10 reports the parameters used in the simulation. The tax rate is chosen to be 20%, the same as the current highest marginal tax rate for long-term capital gains in the US. The risk-free rate is set to be 2.4%, the average federal funds rate between 1993 and 2022. The bid-ask spread is set to 0.24%, the same as the average spread of 0.24% in the sample. Explicit transaction costs, such as commission fees, can be included in the model using a higher spread.  $\sigma_D$  is chosen to be 0.2, close to the annualised standard deviation of value-weighted market returns.  $\bar{H}$  and  $\bar{Z}$  take the value of 1 without loss of generality.  $\bar{D}$  is set to be 1 so that there is enough taxable income in period 1.  $\bar{N}$  takes the value of 2 so that most investors can maintain risk-return optimal portfolios.

Table 5.10: Parameters

Preference		
$\lambda=1$		
Payoff		
$\tau = 20\%$	$r_f = 2.4\%$	$s = 0.24\%$
$\bar{D} = 1$	$\bar{H} = 1$	$\rho = 0.3$
$\sigma_D = 0.2$	$\sigma_H = 0.05$	
Endowment		
$\bar{N} = 2$	$\bar{Z} = 1$	
$\sigma_N = 0.9$	$\sigma_Z = 0.2$	

The table summarises the parameters of the tax-loss harvesting model. Parameters that control the distribution of initial tax basis vary in simulations, so they are not reported in the table. In the baseline scenario, these parameters are  $P_0 = 0.87$  and  $\sigma_P = 0.2$ .

In addition to the parameters mentioned above, the model controls the distribution of tax basis to reflect the impacts of past returns and volatility by varying  $P_0$  and  $\sigma_P$ . (Formula (5.7)) The baseline parameters of the distribution are  $P_0 = 0.87$  and  $\sigma_P = 0.2$ , corresponding to the average past return and volatility in Table 5.1.

Other parameters are tuned such that the simulated tax-loss trading volume accounts for about 20.7% of total trading volume in the baseline scenario, the same as estimated in Section 5.3. Given the endowment parameters, about 1.3% of agents have negative endowments of the tradeable asset. To prevent violating the short-sale restriction, negative endowments are replaced by 0, which has minimal impact on simulation results.

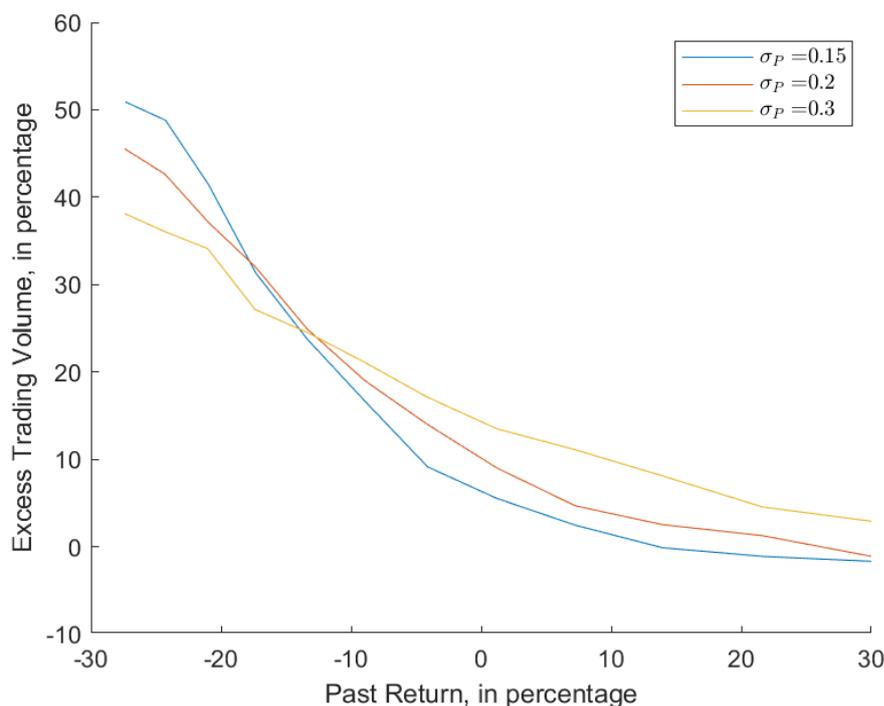
**Return and Volatility** Figure 5.7 depicts the tax-loss trading volume with different distributions of the initial tax basis, which are determined by two tuning parameters— $P_0$

and  $\sigma_P$ . The distribution is wider when  $\sigma_P$  is bigger. The past return is measured by the following formula,

$$r_1 = \frac{P_1}{P_0} - 1 \quad (5.18)$$

The past return is higher when  $P_0$  is lower. This paper chooses  $P_0$  ranging from 0.7 to 1.25 and  $\sigma_P$  from 0.15 to 0.3 for Figure 5.7.

Figure 5.7: Tax-Loss Trading Volume and Price Movements



The figure reports the difference between trading volumes without and with the wash-sale rule, measured as a percentage of the total number of outstanding shares. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with  $\sigma_P = 0.15, 0.2,$  and  $0.3,$  respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

The yellow curve, which has the highest standard deviation of the tax basis  $\sigma_P = 0.3,$  is at the top of the figure for returns over  $-13\%.$  The red curve shows that the tax-loss trading volume with  $\sigma_P = 0.2$  is lower for returns close to and above 0. As shown by the blue curve, the volume with  $\sigma_P = 0.15$  is even lower in that range. When the return drops below  $-18\%,$  the rankings reverse. Scenarios with lower volatility have higher tax-loss trading volumes. Since most returns are concentrated around 0, the tax-loss trading volume is increasing with the past realised volatility on average. The result is consistent with the finding in Table 5.3.

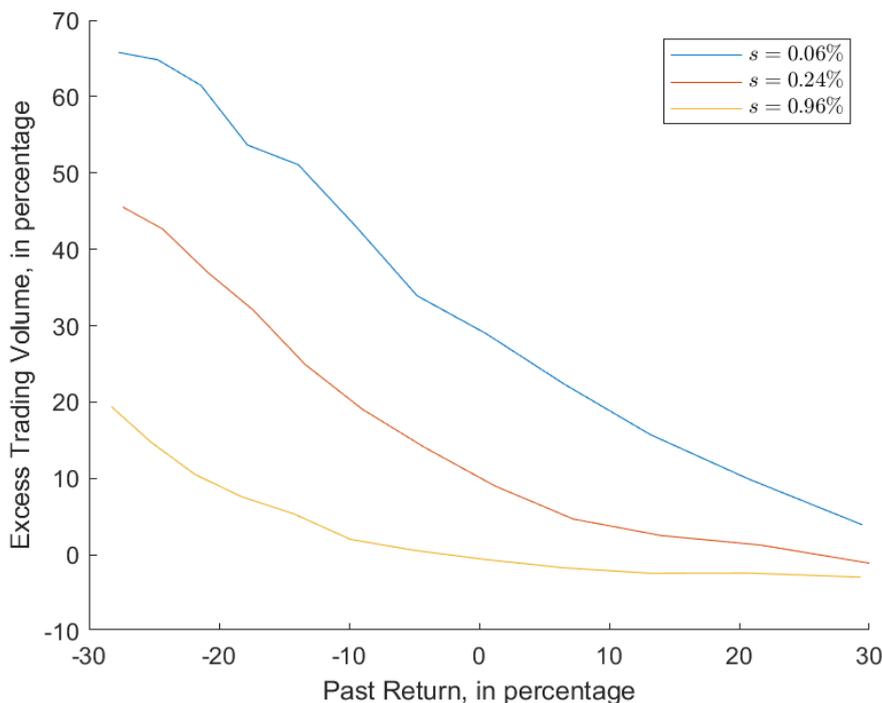
All three curves are negatively sloping, which is consistent with the negative coefficient of  $r_1 \times \text{paired}$  in Table 5.3. Tax-loss harvesting behaviour increases with past capital loss. In the scenario with  $\sigma_P = 0.2$ , the simulation finds that the tax-loss trading volume is virtually 0 when  $r_1 \geq 25\%$ , meaning that almost no one takes tax-loss trades. This is because most people have a lower tax basis than the current price and do not have unrealised capital losses. When the return drops to 0%, about half of the agents have a higher tax basis than the current price. The tax-loss trading volume increases to about 10% of the asset under management. However, not all losing agents choose to harvest the loss. Buying and selling the security cost agents the bid-ask spread  $s$ . If the tax deferral benefit is lower than the spread, agents are in the lock-in region and choose not to trade. When the return drops to the negative region, such as -25%, most agents have unrealised capital gains, so the excess volume increases to 43%. The other two curves show similar patterns.

The curves are all convex, meaning that the speed of the increase in tax-loss harvesting is positively associated with the loss. The convex curves are consistent with the empirical finding of asymmetric tax-loss harvesting behaviour in response to past returns. Focusing on the curve with  $\sigma_P = 0.2$ , the difference of tax-loss trading volumes between  $r_1 = 25\%$  and  $r_1 = 0\%$  is 10%, while the difference between  $r_1 = 0\%$  and  $r_1 = -25\%$  is 33%. Negative returns are strongly associated with the tax-loss trading volume, considerably more substantial than positive returns.

**Bid-Ask Spread** The bid-ask spread is one of the critical factors determining tax-loss harvesting. If the spread is too high, only investors with high unrealised losses find it profitable to harvest the loss. Figure 5.8 plots the tax-loss trading volume in response to the past return with spreads  $s = 0.06\%$ ,  $0.24\%$ , and  $0.96\%$ . The curve with a spread of  $0.24\%$  is the same as in Figure 5.7.  $\sigma_P$  takes the baseline value, 0.2.

The blue curve, calculated with the lowest spread of  $0.06\%$ , is at the top of the figure. The curve is downward-sloping and convex, the same as the others. As the spread goes up, the tax-loss trading volume decreases monotonically. With the past return of 0, a spread of  $0.06\%$  has a tax-loss trading volume of 29%. The number drops to 10% for a higher spread of  $0.24\%$ , as shown by the red curve. With a spread of  $0.96\%$ , the tax-loss trading volume gets further down to almost 0% (yellow curve). The result shows the importance of transaction costs in tax-loss harvesting. A small increase in the bid-ask spread can wipe

Figure 5.8: Tax-Loss Trading Volume and Bid-Ask Spread



The figure reports the difference between trading volumes without and with the wash-sale rule, measured as a percentage of the total number of outstanding shares. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with spreads  $s = 0.06\%$ ,  $0.24\%$ , and  $0.96\%$ , respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

out tax-loss trades almost completely in some scenarios. The monotonic pattern applies to all returns and is consistent with the empirical result in Table 5.5, which shows a negative connection between tax-loss harvesting and bid-ask spread.

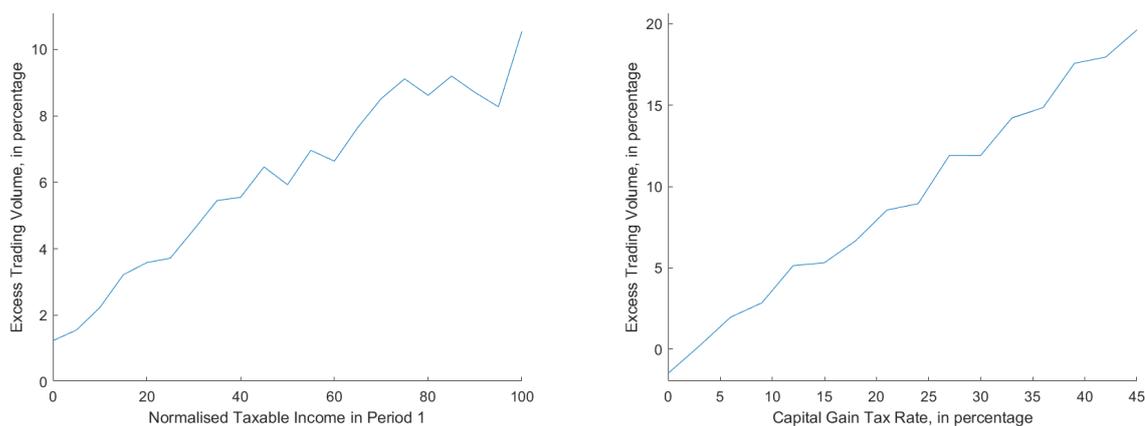
The three curves in the figure are not parallel, showing that the spread effect depends on the past return. The effect is stronger for lower returns than for higher returns. The distance between blue and red curves is close to that between red and yellow curves, but the difference in spread in the former case is  $0.18\%$ , while that in the latter case is  $0.72\%$ . The non-proportional change in trading volume shows that the spread effect is not linear. When the spread is close to 0, a small increase in spread discourages tax-loss trading at a considerable magnitude. When the spread is high, the spread effect becomes much lower.

**Additional Predictions** The tax-loss harvesting model can make additional predictions that are not discussed in the empirical analysis. This subsection focuses on two variables—taxable income and capital gain tax rate. The analysis holds parameters  $P_0 = 0.87$  and

$\sigma_P = 0.2$ . Other parameters are given in Table 5.10.

Tax-loss harvesting depends on the amount of taxable income. If there is no taxable income, there is no tax payment to deduct and no point to harvest loss. Panel (a) in Figure 5.9 depicts the connection between taxable income and the excess trading volume. The taxable income in period 1 in the original setting is normalised to 100. The horizontal axis represents 0% to 100% of the taxable income.

Figure 5.9: Additional Predictions of the Tax-loss Trading Model



(a) Excess Trading Volume and Taxable Income      (b) Excess Trading Volume and Tax rate

Panel (a) reports the excess volume as a function of the taxable income in period 1. Panel (b) reports the excess volume as a function of the capital gain tax rate. The original taxable income (Figure 5.7) in period 1 is normalised to 100. The excess volume is the difference between trading volumes without and with the wash-sale rule, measured as a percentage of the total number of outstanding shares. Initial tax basis parameters  $P_0 = 0.87$  and  $\sigma_P = 0.2$ . Other parameters are listed in Table 5.10. Past return is calculated from Equation (5.18). Each point is calculated with 1000 agents.

Without surprise, the figure shows that the tax-loss trading volume is increasing with the taxable income. When the taxable income is 0, there is no tax to deduct. Investors do not trade due to the positive spread. The simulated excess trading volume with no taxable income is slightly above 0 because of random sampling. The tax-loss trading volume increases with the taxable income. A taxable income of 100 only witnesses an excess volume of 10%. Overall, the shape of the curve is roughly linear. The intuition is that investors want to deduct all taxable incomes whenever possible.

The capital gain tax is vital to tax-loss harvesting. Since the capital gain tax rate does not change frequently in real life, the model provides valuable counterfactual analysis. Panel (b) graphs the connection between capital gain tax rate and tax-loss harvesting. The excess trading volume increases with the tax rate, with no surprise. The relationship is

close to linear. When the tax rate is 0, the tax-loss harvesting volume is seemingly 0. The number with a tax rate of 20% is 8%, and that with a tax rate of 40% is 18%.

### C. Tax Revenue Loss

The model captures key features of the empirical results. It predicts that the tax-loss harvesting behaviour is negatively associated with the past return. The negative return substantially impacts the tax-loss trading volume, while the positive one has little impact. Besides, the model shows that the tax-loss trading volume increases with past volatility and decreases with the bid-ask spread.

This paper applies the model to estimate the tax revenue loss from tax-loss harvesting with ETFs. Following the previous discussion, the model has only two periods to keep parsimony.

The discussion starts with the scenario where the wash-sale rule does not apply. Investors realise capital losses to deduct tax in period 1. The loss generated from selling an asset includes the bid-ask spread, so for investor  $i$ , the capital loss is

$$L_1^i = [P_0^i - P_1(1 - \frac{s}{2})](N_{0,0}^i - N_{0,1}^i) \quad (5.19)$$

The equation only considers the number of shares the investor sells but not the purchase of new shares because the latter does not impact the tax payable in period 1.

Similarly, when the wash-sale rule applies, the capital loss is  $L_1^{i,WS}$ . The tax revenue loss of the government relative to the asset value is

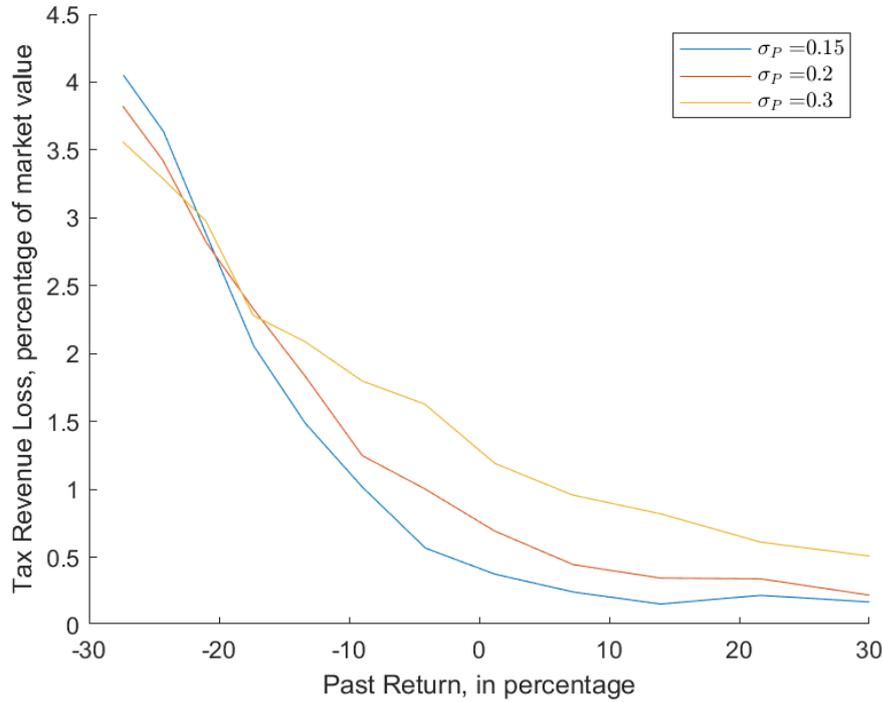
$$RL_1 = \frac{\sum_i \tau(L_1^i - L_1^{i,WS})}{\sum_i N_{0,0}^i P_1} \quad (5.20)$$

However, since transaction costs exist,  $RL_1$  is not the net benefit investors get. The tax alpha should be adjusted by the bid-ask spread.

$$\alpha_\tau = RL_1 - sV_1^{TLH} \quad (5.21)$$

**Results** Figure 5.10 depicts the relationship between the tax revenue loss and the bid-ask spread, which is similar to Figure 5.7. The horizontal axis is the return calculated from Equation (5.18), and the vertical axis is the tax revenue loss relative to the market value.

Figure 5.10: Tax Revenue Loss and Price Movements



The figure reports the simulated tax revenue loss due to tax-loss harvesting, measured as a percentage of assets under management. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with  $\sigma_P = 0.15$ ,  $0.2$ , and  $0.3$ , respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

There is a monotonic pattern that tax revenue loss is positively associated with past volatility. For returns above  $-23\%$ , the yellow curve with the highest past volatility ( $\sigma_P = 0.3$ ) is on the top of the figure. The curve with  $\sigma_P = 0.2$  is in the middle, and the one with  $\sigma_P = 0.15$  has the lowest tax revenue loss. The reason is that higher past volatility widens the spread of tax basis and, hence, the tax-loss trading volume. All three curves are downward-sloping and convex, the same as in Figure 5.7. The pattern is also mainly due to the tax loss trading volume.

The similarity between Figures 5.10 and 5.7 show that the tax revenue loss is closely connected with tax-loss trading volume, holding parameters constant. The model is tuned so that the baseline scenario ( $P_0 = 0.87$  and  $\sigma_P = 0.2$ ) has a tax-loss trading volume relative to the total trading volume close to the empirical finding ( $20.7\%$ ). Since the parameters in Table 5.10 are annualised, the tax revenue loss estimated with the baseline parameters should indicate the potential annual tax revenue loss in real life.

The simulated tax revenue loss under the baseline scenario is  $0.52\%$  of the market

value per year. Using the total assets under management of all paired ETFs in 2021, the estimated annual tax revenue loss is about 25 billion USD. The estimation is comparable with the back-test maximum tax alpha of 1.02% estimated by [Bouchey et al. \(2016\)](#), who simulate the historical return data between 2011 and 2015 with a bid-ask spread of 0.

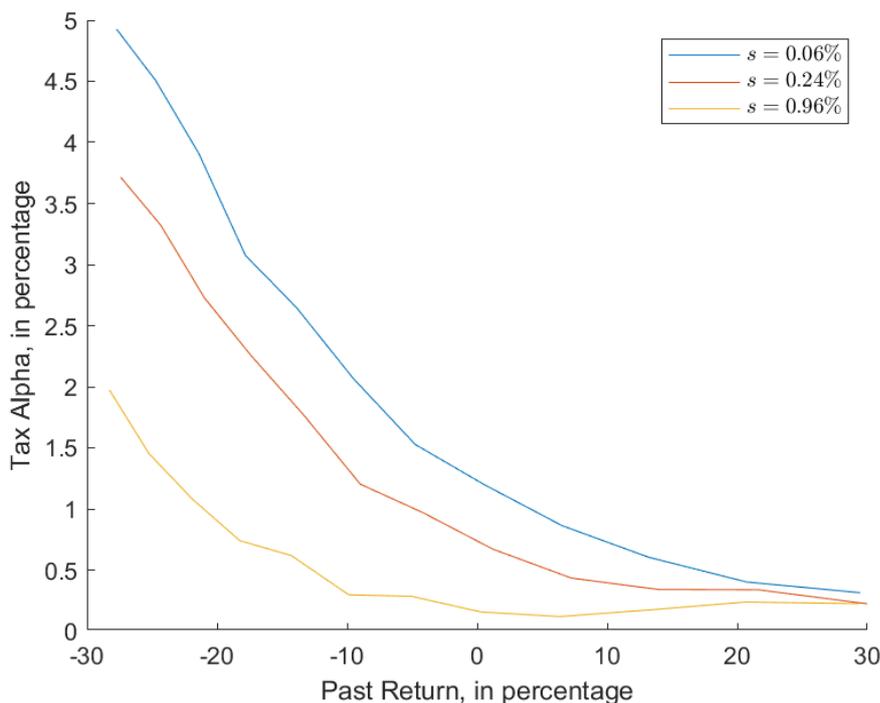
From investors' perspective, the tax benefit of harvesting capital loss is lower than the tax revenue loss due to transaction costs. Under the baseline scenario ( $P_0 = 0.87$  and  $\sigma_P = 0.2$ ), the transaction cost is 1 basis point (bp), bringing the tax alpha down to 51 bps from 52 bps. As shown in Equation 5.21, the transaction cost proportionally increases with the tax-loss trading volume. With a return of 0, the transaction cost rises to 2 bps. If the return decreases to -25%, the cost goes up to 10 bps. The estimated tax alpha with various past returns and volatility is plotted in the appendix in Table 13. Compared with the tax benefit from tax-loss harvesting, the transaction cost only accounts for a very small portion of the tax revenue loss.

Although the direct effect of the bid-ask spread on the tax revenue loss is minimal, it does not mean that the spread is unimportant. The bid-ask spread also has an indirect effect on the tax revenue loss through tax-loss trading volume. When the transaction cost increases, the number of investors with profitable tax-loss harvesting positions drops. As a result, fewer investors get involved in tax-loss harvesting, reducing the trading volume and the tax revenue loss.

The indirect effect is much stronger than the direct one. Figure 5.11 plots the tax revenue loss with three different spread levels. Under the baseline scenario ( $P_0 = 0.87$  and  $\sigma_P = 0.2$ ) with a bid-ask spread of 0.24%, the tax alpha is 51 bps, and the transaction cost is 1 bp. If the spread increases by three times to 0.96%, the tax alpha drops to 12 bps, and the transaction cost becomes seemingly 0. The indirect effect is 39 bps, much more substantial than the direct one (1 bp). If the spread decreases to a quarter of the original level (0.06%), the tax alpha becomes 93 bps, while the transaction cost remains 1 bp. In this case, the increase in trading volume is offset by the decrease in the spread, so the total transaction cost  $sV_1^{TLH}$  is almost unchanged. The change of tax alpha is entirely due to the indirect effect.

Curves in Figure 5.11 are all downward-sloping and convex, similar to those in Figure 5.8 plotting the tax-loss trading volume. The similarity means the tax alpha is closely connected with the tax-loss trading volume. The bid-ask spread is important in explaining

Figure 5.11: Tax Alpha and Bid-Ask Spread



The figure reports the simulated tax alpha resulting from tax-loss harvesting, measured as a percentage of the total number of outstanding shares. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with spreads  $s = 0.06\%$ ,  $0.24\%$ , and  $0.96\%$ , respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

the tax alpha indirectly through the trading volume, but not directly.

## 5.5 Conclusion

This paper makes several contributions to the finance literature and the US tax policy.

First, the paper shows that the tax efficiency of the ETF comes from not only owning it but also trading it. Owning an ETF helps investors track the index closely, while trading ETFs allows investors to reduce the taxable income of the whole portfolio, making ETFs more valuable than the index. Through buying and selling highly correlated ETFs, investors effectively avoid the wash-sale rule that disallows tax deduction with trades only for tax purposes.

Second, the paper finds evidence that investors exploit the trading efficiency of ETFs with a large magnitude. The paper uses DID regressions to show that paired ETFs have a monthly trading volume of 9.1% of AUM than unpaired ones, corresponding to 20.7%

of total trading volume. The result is robust to the definition of paired ETFs and the treatment of trading volume. Past returns are negatively associated with the tax-loss trading volume. The effect is more substantial for recent and negative returns than older and positive ones. The paper also finds that the negative return 12 months ago has a strong impact on tax-loss harvesting, which coincides with the threshold classifying short-term and long-term capital gains. The realised volatility is positively correlated with the tax-loss trading volume, reflecting the impact of the tax basis distribution.

In addition, this paper explores the impacts of ETF characters and market incentives on tax-loss harvesting. The research shows that the bid-ask spread negatively correlates with the tax-loss trading volume, while the size positively correlates with tax-loss harvesting. After controlling for size, the expense ratio has no impact on the tax-loss trading volume. This paper also documents the evidence that medium-to-long-term risk-free rates are positively connected with tax-loss harvesting behaviour. In contrast, the short-run monetary tightness, represented by the fed fund rate, harms tax-loss trades. There is no evidence supporting a higher tax-loss harvesting volume of ETFs in December.

Furthermore, this paper proposes a parsimonious model to explain the connections between tax-loss harvesting and price movements. The model predicts that the past return is negatively associated with tax-loss harvesting behaviour. The marginal impact of return is more substantial when the return is lower. The model also predicts a positive relationship between the realised volatility and tax-loss harvesting. The spread is negatively correlated with tax-loss trading. The marginal effect is stronger for lower spreads. These results are consistent with empirical findings. The model also makes additional predictions—higher taxable income and tax rate are positively associated with tax-loss trading volume, but the marginal effects are decreasing.

Last but not least, by simulating the model, the paper estimates the tax revenue loss from the government's perspective and the tax alpha from the investors' perspective. Under the baseline scenario tuned to match annualised market return and volatility, the estimated annual tax revenue loss is 0.52% of the market value of paired ETFs, equivalent to about 25 billion USD in 2021. The transaction cost has a minimal direct impact on the tax alpha, but it has a substantial indirect impact through the profitability of tax-loss harvesting and, hence, the tax-loss trading volume.

The trading tax efficiency of ETFs is essential to policymakers for the following reasons.

To begin with, it is unfair to tax investors of ETFs and other securities differently. Since the wash-sale rule can be avoided using paired ETFs but not other securities, the current tax policy effectively subsidises ETFs. Besides, tax-loss harvesting with ETFs takes out the capital gain tax from other assets, resulting in a potential tax revenue loss of tens of billions of dollars annually. Regulators are encouraged to explicitly define the conditions that make ETFs "substantially identical" and prevent wash sales with highly correlated ETFs accordingly.

## References

- Agapova, A. (2011), ‘Conventional mutual index funds versus exchange-traded funds’, *Journal of Financial Markets* **14**(2), 323–343.
- Agarwal, V., Gay, G. D. and Ling, L. (2014), ‘Window dressing in mutual funds’, *The Review of Financial Studies* **27**(11), 3133–3170.
- Arkhangelsky, D., Imbens, G. W., Lei, L. and Luo, X. (2021), ‘Double-robust two-way-fixed-effects regression for panel data’, *arXiv preprint arXiv:2107.13737* .
- Baker, A. C., Larcker, D. F. and Wang, C. C. (2022), ‘How much should we trust staggered difference-in-differences estimates?’, *Journal of Financial Economics* **144**(2), 370–395.
- Bergstresser, D. and Pontiff, J. (2013), ‘Investment taxation and portfolio performance’, *Journal of Public Economics* **97**, 245–257.
- Bouchev, P., Brunel, J. L. and Li, T. (2016), ‘The role of etfs in active tax management’, *The Journal of Wealth Management* **19**(3), 75–86.
- Brown, D. C., Cederburg, S. and Towner, M. (2021), ‘Dominated etfs’, *Available at SSRN 3694592* .
- Callaway, B. and Sant’Anna, P. H. (2021), ‘Difference-in-differences with multiple time periods’, *Journal of econometrics* **225**(2), 200–230.
- Chaudhuri, S. E., Burnham, T. C. and Lo, A. W. (2020), ‘An empirical evaluation of tax-loss-harvesting alpha’, *Financial Analysts Journal* **76**(3), 99–108.

- Chen, G.-m., Firth, M. and Rui, O. M. (2001), ‘The dynamic relation between stock returns, trading volume, and volatility’, *Financial Review* **36**(3), 153–174.
- Colon, J. M. (2017), ‘The great etf tax swindle: The taxation of in-kind redemptions’, *Penn St. L. Rev.* **122**, 1.
- Colon, J. M. (2023), ‘Unplugging heartbeat trades and reforming the taxation of etfs’.
- Constantinides, G. M. (1984), ‘Optimal stock trading with personal taxes: Implications for prices and the abnormal january returns’, *Journal of Financial Economics* **13**(1), 65–89.
- Dammon, R. M., Spatt, C. S. and Zhang, H. H. (2001), ‘Optimal consumption and investment with capital gains taxes’, *The Review of Financial Studies* **14**(3), 583–616.
- Dammon, R. M., Spatt, C. S. and Zhang, H. H. (2004), ‘Optimal asset location and allocation with taxable and tax-deferred investing’, *The Journal of Finance* **59**(3), 999–1037.
- DeMiguel, V. and Uppal, R. (2005), ‘Portfolio investment with the exact tax basis via nonlinear programming’, *Management Science* **51**(2), 277–290.
- Easley, D. and O’Hara, M. (1992), ‘Adverse selection and large trade volume: The implications for market efficiency’, *Journal of Financial and Quantitative Analysis* **27**(2), 185–208.
- Fischer, G. M. (2010), ‘New twists on an old plot: investors look to avoid the wash sale rule by harvesting tax losses with exchange-traded funds’, *Wash. UL Rev.* **88**, 229.
- Frei, C. and Welsh, L. (2022), ‘How the closure of a us tax loophole may affect investor portfolios’, *Journal of Risk and Financial Management* **15**(5), 209.
- Gardner, R. and Welch, J. (2005), ‘Increasing after-tax return with exchange-traded funds’, *Journal of Financial Planning* **18**(6), 30.
- Garlappi, L., Naik, V. and Slive, J. (2001), ‘Portfolio selection with multiple assets and capital gains taxes’, *Available at SSRN 274939* .
- Gastineau, G. L. (2004), ‘The benchmark index etf performance problem’, *The Journal of Portfolio Management* **30**(2), 96–103.

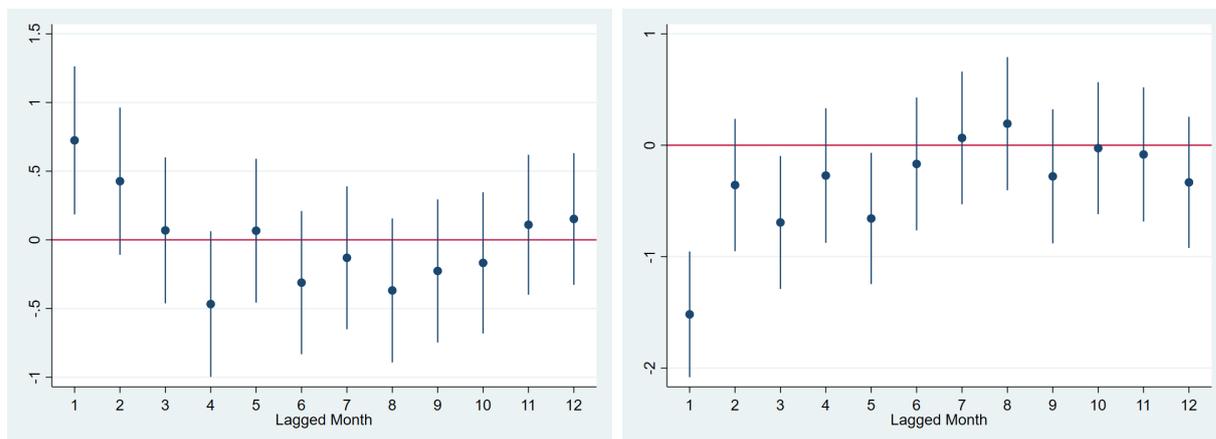
- Goodman-Bacon, A. (2021), ‘Difference-in-differences with variation in treatment timing’, *Journal of Econometrics* **225**(2), 254–277.
- Graham, B. S., de Xavier Pinto, C. C. and Egel, D. (2012), ‘Inverse probability tilting for moment condition models with missing data’, *The Review of Economic Studies* **79**(3), 1053–1079.
- Grinblatt, M. and Moskowitz, T. J. (2004), ‘Predicting stock price movements from past returns: the role of consistency and tax-loss selling’, *Journal of Financial Economics* **71**, 541–579.
- Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), ‘Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme’, *The review of economic studies* **64**(4), 605–654.
- Hodaszy, S. Z. (2016), ‘Tax-efficient structure or tax shelter: Curbing etfs’ use of section 852 (b)(6) for tax avoidance’, *Tax Law.* **70**, 537.
- Hodaszy, S. Z. (2022), ‘Exchange-traded funds use section 852 (b)(6) for tax avoidance, not just tax deferral: So why is this loophole still open?’, *Tax Lawyer* **75**(3), 489.
- Holthausen, R. W. and Verrecchia, R. E. (1990), ‘The effect of informedness and consensus on price and volume behavior’, *Accounting Review* pp. 191–208.
- Internal Revenue Service (2022), Publication 550: Investment income and expenses (including capital gains and losses), Technical report, Department of the Treasury.
- Ivković, Z., Poterba, J. and Weisbenner, S. (2005), ‘Tax-motivated trading by individual investors’, *American Economic Review* **95**(5), 1605–1630.
- Karpoff, J. M. (1986), ‘A theory of trading volume’, *The journal of finance* **41**(5), 1069–1087.
- Karpoff, J. M. (1987), ‘The relation between price changes and trading volume: A survey’, *Journal of Financial and Quantitative Analysis* **22**(1), 109–126.
- Kashner, E. (2017), ‘The heartbeat of etf tax efficiency’.  
**URL:** <https://insight.factset.com/the-heartbeat-of-etf-tax-efficiency>

- Kim, O. and Verrecchia, R. E. (1991), ‘Trading volume and price reactions to public announcements’, *Journal of accounting research* **29**(2), 302–321.
- Kostovetsky, L. (2003), ‘Index mutual funds and exchange-traded funds’, *Journal of Portfolio Management* **29**(4), 80.
- Lettau, M. and Madhavan, A. (2018), ‘Exchange-traded funds 101 for economists’, *Journal of Economic Perspectives* **32**(1), 135–154.
- Llorente, G., Michaely, R., Saar, G. and Wang, J. (2002), ‘Dynamic volume-return relation of individual stocks’, *The Review of financial studies* **15**(4), 1005–1047.
- Moussawi, R., Shen, K. and Velthuis, R. (2022), ‘The role of taxes in the rise of etfs’, Available at SSRN 3744519 .
- Odean, T. (1998), ‘Volume, volatility, price, and profit when all traders are above average’, *The journal of finance* **53**(6), 1887–1934.
- Poterba, J. M. and Shoven, J. B. (2002), ‘Exchange-traded funds: A new investment option for taxable investors’, *American Economic Review* **92**(2), 422–427.
- Poterba, J. M. and Weisbenner, S. J. (2001), ‘Capital gains tax rules, tax-loss trading, and turn-of-the-year returns’, *The Journal of Finance* **56**(1), 353–368.
- Sant’Anna, P. H. and Zhao, J. (2020), ‘Doubly robust difference-in-differences estimators’, *Journal of Econometrics* **219**(1), 101–122.
- Scheinkman, J. A. and Xiong, W. (2003), ‘Overconfidence and speculative bubbles’, *Journal of Political Economy* **111**(6), 1183–1220.
- Sialm, C. (2009), ‘Tax changes and asset pricing’, *American Economic Review* **99**(4), 1356–1383.
- Wang, J. (1994), ‘A model of competitive stock trading volume’, *Journal of Political Economy* **102**(1), 127–168.

## Appendix

### A. ETFs Tracking the Same Index

Figure 12: Asymmetric Response of Tax-Loss Trading Volume



(a) Loading of Tax-Loss Trading Volume on Positive Return

(b) Loading of Tax-Loss Trading Volume on Negative Return

The treatment is defined as the introduction of the first ETF that tracks the same index as ETF  $j$  ( $same_{j,t}$ ). Panel (a) in this figure reports coefficients and 95% confidence intervals for the interaction term  $r_{j,\tau,t}^+ \times same_{j,t}$  in Equation (5.3); panel (b) reports coefficients and 95% confidence intervals for the interaction term  $r_{j,\tau,t}^- \times same_{j,t}$ . The horizontal axis shows the lags of returns ( $\tau$ ) in months.

Table 11: Asymmetric Response of Trading Volume to Past Returns

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>same</i>	-0.002 (0.019)	-0.053*** (0.021)	-0.005 (0.019)	-0.079*** (0.023)
$r_1^+$	1.380*** (0.056)	1.352*** (0.056)	1.352*** (0.056)	1.334*** (0.056)
$r_1^+ \times same$		1.064*** (0.245)		0.694*** (0.250)
$r_1^-$	-1.008*** (0.064)	-0.963*** (0.065)	-1.000*** (0.064)	-0.962*** (0.065)
$r_1^- \times same$		-1.796*** (0.275)		-1.449*** (0.277)
$r_{2-12}^+$			0.498*** (0.016)	0.498*** (0.016)
$r_{2-12}^+ \times same$				0.080 (0.065)
$r_{2-12}^-$			-0.233*** (0.021)	-0.220*** (0.021)
$r_{2-12}^- \times same$				-0.962*** (0.134)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.083	0.083	0.087	0.088

The treatment ( $same_{j,t}$ ) is defined as the introduction of the first ETF that tracks the same index as ETF  $j$ . The table reports results for regressions with different specifications of past returns. "+" and "-" signs represent positive and negative past returns, respectively. Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

## B. Correlation Threshold

Table 12: Tax-Loss Trading Volume with Correlation Threshold of 0.95

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>paired95</i>	0.103*** (0.007)	0.107*** (0.007)	0.122*** (0.007)	0.085*** (0.007)
$r_1$	0.318*** (0.034)	0.332*** (0.034)	0.627*** (0.056)	0.336*** (0.034)
$r_1 \times \textit{paired95}$			-0.401*** (0.061)	
$r_{2-12}$		0.190*** (0.011)	0.317*** (0.017)	
$r_{2-12} \times \textit{paired95}$			-0.178*** (0.018)	
<i>volatility</i>	0.362*** (0.044)	0.360*** (0.044)	0.353*** (0.044)	0.071 (0.046)
$\textit{volatility} \times \textit{paired95}$				3.849*** (0.156)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.080	0.081	0.082	0.081

The table reports results for regressions with different specifications of past returns. The DID term  $\textit{paired95} = 1$  if the current ETF has a correlation coefficient higher than 0.95 with at least one other ETF. Otherwise,  $\textit{paired95} = 0$ . Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

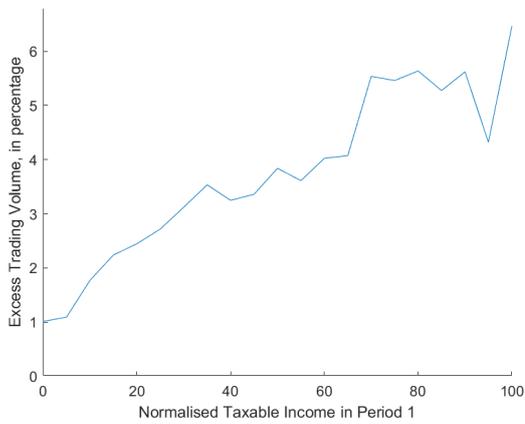
Table 13: Tax-Loss Trading Volume with Correlation Threshold of 0.995

Dependent Variable: <i>volume</i>				
	(1)	(2)	(3)	(4)
<i>paired995</i>	0.065*** (0.010)	0.069*** (0.010)	0.091*** (0.010)	0.041*** (0.010)
$r_1$	0.316*** (0.034)	0.329*** (0.034)	0.378*** (0.035)	0.339*** (0.034)
$r_1 \times \textit{paired995}$			-0.455*** (0.075)	
$r_{2-12}$		0.185*** (0.011)	0.207*** (0.011)	
$r_{2-12} \times \textit{paired995}$			-0.232*** (0.025)	
<i>volatility</i>	0.361*** (0.044)	0.359*** (0.044)	0.355*** (0.044)	0.329*** (0.044)
$\textit{volatility} \times \textit{paired995}$				6.459*** (0.390)
Control	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
ETF FE	Yes	Yes	Yes	Yes
N	183779	183779	183779	183779
$R^2$	0.080	0.081	0.082	0.081

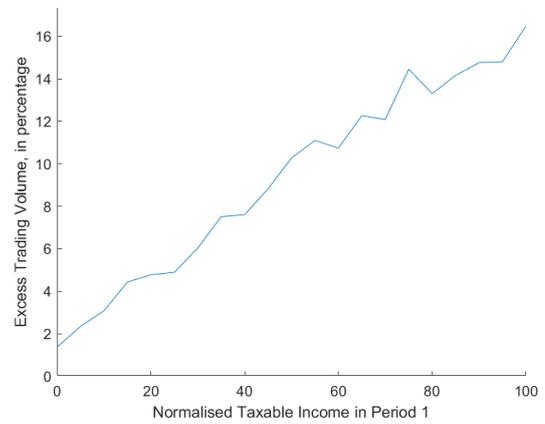
The table reports results for regressions with different specifications of past returns. The difference-in-difference term  $\textit{paired995} = 1$  if the current ETF has a correlation coefficient higher than 0.995 with at least one other ETF. Otherwise,  $\textit{paired995} = 0$ . Standard errors are clustered by ETF and reported in parentheses. \*, \*\*, and \*\*\* indicate significance at 10%, 5%, and 1% levels, respectively.

### C. Additional Predictions

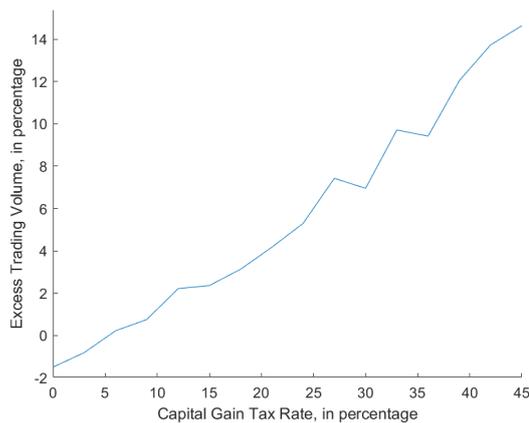
Figure 13: Additional Predictions of the Tax-loss Trading Model



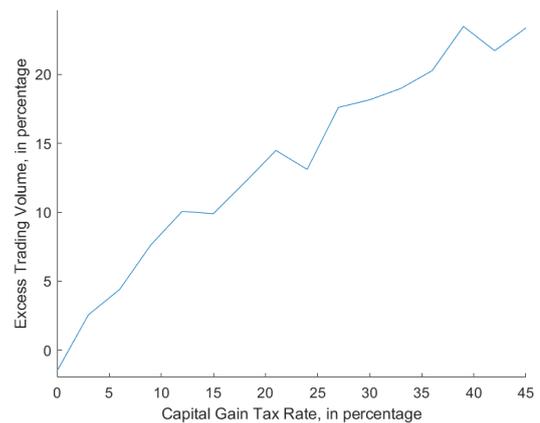
(a) Various Taxable Income,  $\sigma_P = 0.15$



(b) Various Taxable Income,  $\sigma_P = 0.3$



(c) Various Tax rate,  $\sigma_P = 0.15$

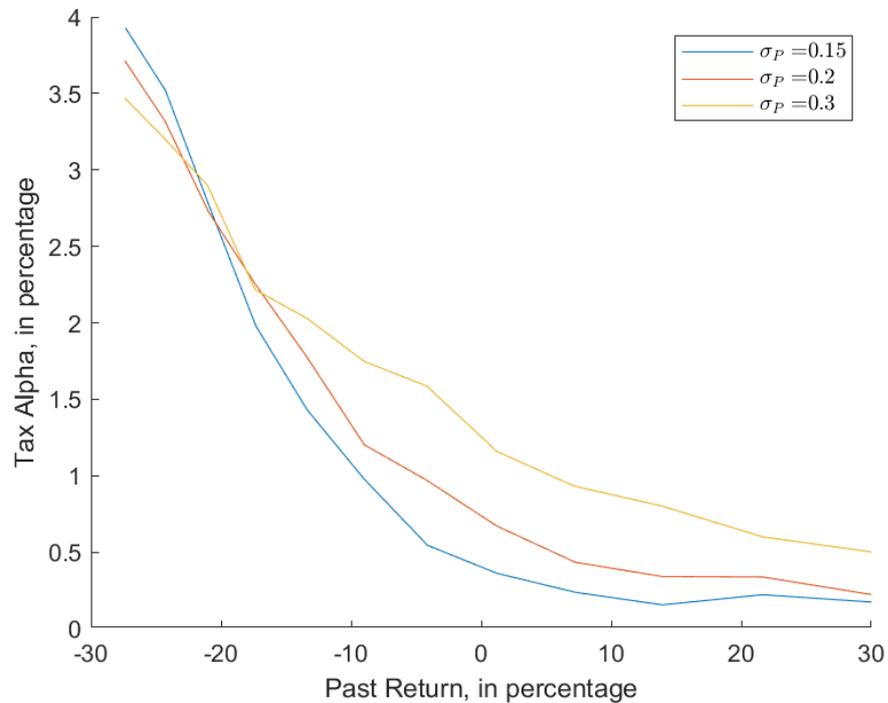


(d) Various Tax rate,  $\sigma_P = 0.3$

Panels (a) and (b) report the excess volume as a function of the taxable income in period 1. Panels (c) and (d) report the excess volume as a function of the capital gain tax rate. Initial tax basis parameter  $P_0 = 0.87$ . Other parameters are listed in Table 5.10. The original taxable income (Figure 5.7) in period 1 is normalised to 100. Each point is calculated with 1000 agents.

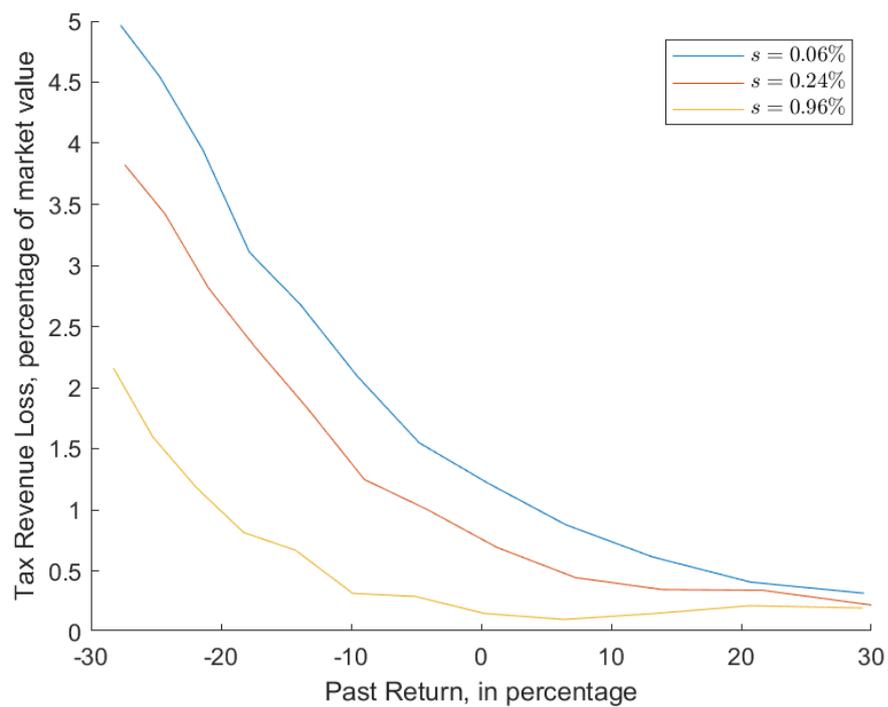
## D. Tax Alpha and Tax Revenue Loss

Figure 14: Tax Alpha and Price Movements



The figure reports the simulated tax alpha resulting from tax-loss harvesting, measured as a percentage of assets under management. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with  $\sigma_P = 0.15$ ,  $0.2$ , and  $0.3$ , respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

Figure 15: Tax Revenue loss and Bid-Ask Spread



The figure reports the simulated tax revenue loss due to tax-loss harvesting, measured as a percentage of the total number of outstanding shares. Past return is calculated from Equation (5.18). Blue, red, and yellow curves show results with spreads  $s = 0.06\%$ ,  $0.24\%$ , and  $0.96\%$ , respectively. Points in the same curve are obtained by changing  $P_0$ , ceteris paribus. Other parameters are listed in Table 5.10. Each point is calculated with 1000 agents.

## Part III

### Remark and Conclusion

# 6

## Remark and Conclusion

### 6.1 Summary of Essays

This thesis provides insights into two topics in asset pricing—the term structure of equity returns and tax-loss harvesting with ETFs. Chapter 1 briefly introduces the scope of the research and the development of consumption-based models and factor models. Chapter 2 discusses the related literature of this thesis, providing comments on research areas of the equity market, mutual funds, and capital gain taxation.

Chapter 3 presents a series of findings concerning the term structures of equity returns and the term structure of equity return volatility. Firstly, realised volatility can positively predict the dividend strip return. The volatility feedback effect of the dividend strip decreases with the horizon, while that of the stock index is insignificant. Secondly, both expected and unexpected realised volatilities contribute to the monotonic pattern of the volatility feedback effect. Expected volatility positively predicts 1-year dividend strip returns and negatively predicts medium-term dividend strip returns. Thirdly, realised and implied volatilities exhibit a negative association with the equity term premium, indicating a pro-cyclical nature of the equity term structure. With different sets of information, long-duration and short-duration volatilities explain the term structure slope of equity returns. Notably, the predictive power is more pronounced for the long horizon compared to the short horizon, albeit reversing in the very long horizon. Furthermore, the decomposition of cyclicalities underscores that the pro-cyclical term structure is primarily attributable to the high relative sensitivity of short-duration volatility. Finally, simulations suggest that the rare disaster model should be rejected, while the preferred habit model and the long-run

risk model lack empirical support in explaining the observed cyclicalities.

Chapter 4 raises the short-duration equity return puzzle using dividend strips. Departing from conventional approaches, this chapter employs a forward-looking methodology to analyse the required rate of return and conditional Sharpe ratio for dividend strips, linking empirical returns with the mean-standard deviation frontier of IMRS. The study demonstrates the unpredictably high required rates of returns of dividend strips during crisis periods, such as the COVID-19 pandemic, the dot-com bubble, and the global financial crisis. In March 2020 (COVID-19 pandemic), returns of 1- to 4-year dividend strips are 55%, 32%, 19%, and 15%, respectively. It also reveals exceptionally high conditional Sharpe ratios for dividend strips during crises. The highest conditional Sharpe ratio in our sample is 14 for the 1-year dividend in March 2020, which is too high to be explained by mainstream consumption-based models, including the preferred habit model, the long-run risk model, the rare disaster model, and the term structure model.

Chapter 5 shows a new source of ETF tax efficiency: tax-loss harvesting. By buying and selling highly correlated ETFs, investors effectively avoid the wash-sale rule, which disallows tax deductions with trades only for tax purposes. Besides, this chapter finds evidence that investors exploit the trading efficiency of ETFs with a large magnitude, corresponding to 20.7% of total trading volume. Thirdly, the study explores how ETF characteristics and market incentives impact tax-loss harvesting behaviours, highlighting associations with bid-ask spreads, ETF size, and macroeconomic factors. Additionally, this chapter develops a parsimonious model that accurately captures the relationships between tax-loss harvesting and price movements. Lastly, by simulating the model, the study estimates substantial annual tax revenue losses attributable to tax-loss harvesting—0.52% of the market value of paired ETFs, equivalent to about 25 billion USD in 2021.

## 6.2 Research Implications

The original research chapters of this thesis have important implications for both asset pricing research and government policy.

## A. Academic Implications

Chapter 3 contributes mainly to two strands of literature. The first one is the volatility feedback effect. To the best of my knowledge, this is the first paper investigating the volatility feedback effect of dividend strips and mapping out the term structure. Besides, the paper also contributes to the equity return term structure literature by showing that the term structure of equity returns is pro-cyclical, supporting [Bansal et al. \(2021\)](#) and [Ai et al. \(2018\)](#). Notably, the empirical finding of cyclical volatility provides a brand new tool to test asset pricing models.

Follow-up research can make the below potential improvements. First, this paper only considers a short sample from 2015 to 2021. It would be interesting to include extended data from the past with proprietary datasets. Second, later research can use more sophisticated methods to construct the expected future realised volatility, potentially with VAR. Finally, if the data for long-horizon variance swap is available, one can construct risk-neutral implied variance and variance premium to investigate the term structure of implied volatility.

The biggest contribution of Chapter 4 is proposing the short-duration equity return puzzle—the conditional Sharpe ratios of short-horizon dividend strips are too high to be explained during crises. This puzzle is a novel criterion to consider for consumption-based asset pricing models, in addition to the equity premium puzzle, the excess volatility puzzle, and other criteria. Besides, the research also contributes to the term structure of equity return literature by innovatively using forward-looking professional forecasts.

Follow-up research can focus on the development of a new asset pricing model that can explain the short-duration equity return puzzle. A model that can explain the high conditional Sharpe ratio of the dividend strip must produce sufficient conditional volatility of the stochastic discount factor while keeping the coefficient of relative risk aversion within a reasonable level. A potential solution is to include behavioural factors in the model, emphasising the irrational dividend expectation during economic crises.

Chapter 5 makes two main contributions to the literature. Firstly, it is the first empirical research that documents the exploitation of a tax loophole in the US—tax-loss harvesting with ETFs. The finding provides a brand new angle on the tax efficiency of the ETF literature. This chapter also contributes to the trading volume literature by developing a tax-loss trading model.

Follow-up research can focus on the optimal tax-loss harvesting rule with correlated assets and transaction costs. The proposed rule links the tax-loss harvesting behaviour with the correlation coefficient of asset returns, guiding the authority to regulate the ETF market. Future research can also explore the micro-level holding data for individual investors, which helps estimate the tax revenue loss more precisely.

## B. Policy Implications

The original research chapters also have important policy implications.

Chapter 3 shows that short-duration equity returns are more sensitive to volatility than long-duration ones. Therefore, the volatility of long-duration stocks contributes less to the overall market during market downturns. The optimal regulatory strategy for markets with price limits should allow long-duration stocks to be more volatile than short-duration ones. The extremely high required rate of return and conditional Sharpe ratio in Chapter 4 also remind policymakers to focus on the extreme movements of short-duration equity prices. Since mainstream consumption-based asset pricing models cannot rationalise the short-duration equity return puzzle, policymakers should be aware of behavioural factors that may play a role.

Chapter 5 indicates that the trading tax efficiency of ETFs is important to policymakers for the following reasons. To begin with, it is unfair to tax investors of ETFs and other securities differently. Since the wash-sale rule can be avoided using paired ETFs but not other securities, the current tax policy effectively subsidises ETFs. Besides, tax-loss harvesting with ETFs takes out the capital gain tax from other assets, resulting in a potential tax revenue loss of tens of billions of dollars annually. Regulators are encouraged to explicitly define the conditions that make ETFs "substantially identical" and prevent wash sales with highly correlated ETFs accordingly.

## 6.3 Conclusion

Overall, this thesis contributes to the research area of asset pricing, advancing the understanding of equity returns and portfolio management. Its empirical findings provide novel test criteria for asset pricing models, namely the predictable cyclicity of the equity return term structure and the short-duration equity return puzzle, which guide the development

of asset pricing theories. The empirical evidence of tax-loss harvesting with ETFs unveils a new source of tax efficiency of ETFs, furthering the research of portfolio management. Besides, this thesis makes theoretical contributions by developing a tax-loss harvesting model with transaction costs.

## References

- Ai, H., Croce, M. M., Diercks, A. M. and Li, K. (2018), ‘News shocks and the production-based term structure of equity returns’, *The Review of Financial Studies* **31**(7), 2423–2467.
- Bansal, R., Miller, S., Song, D. and Yaron, A. (2021), ‘The term structure of equity risk premia’, *Journal of Financial Economics* .