

Understanding the Dynamics of REIT Equity and Debt Markets with Machine Learning



Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität
Regensburg

vorgelegt von:
HENDRIK JENETT

Berichterstatter: Prof. Dr. Wolfgang Schäfers
Prof. Dr. Daniel Rösch

Tag der Disputation: 24. April 2026

Understanding the Dynamics of REIT Equity and Debt Markets with Machine Learning

Hendrik Jenett

Table of Contents

List of Tables.....	VII
List of Figures	VIII
1 Introduction	1
1.1 Motivation and Background	1
1.2 Course of Analysis and Research Questions	4
1.3 Co-Authors, Submissions and Conference Presentations	7
1.4 References	9
2 Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach	11
2.1 Abstract	11
2.2 Introduction	12
2.3 Literature Review	15
2.3.1 REIT Performance Modeling with Traditional Methods	15
2.3.2 Financial Products Performance Modeling with Machine Learning....	16
2.3.3 Direct Real Estate Performance Modeling with Machine Learning....	17
2.3.4 Indirect Real Estate Performance Modeling with Machine Learning..	18
2.4 Data.....	18
2.5 Methodology	23
2.5.1 Machine Learning Method	23
2.5.2 Feature Importance	26
2.5.3 Feature Effects	27
2.6 Results	28
2.6.1 Global Absolute Feature Importance	28
2.6.2 Evolution of Feature Importance	31
2.6.3 Local Directional Feature Effects	37
2.7 Conclusion	44
2.8 Appendix	46
2.8.1 Appendix I – Chosen Variables and Related Literature	46
2.8.2 Appendix II – Model Performance Comparison	46
2.8.3 Appendix III – Feature Importance of Macroeconomic Variables.....	49
2.8.4 Appendix IV – Feature Importance of REIT-Specific Variables.....	50
2.9 References	52

3	Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt	58
3.1	Abstract.....	58
3.2	Introduction.....	59
3.3	Literature Review and Hypotheses Development	61
3.3.1	Location and Firm Risk.....	62
3.3.2	Diversification and REIT Equity Performance and Debt Costs	62
3.3.3	Hypotheses Development	63
3.4	Data and Variable Construction.....	64
3.4.1	REIT Bond Risk Premium	64
3.4.2	Local Real Estate Market Risk.....	65
3.4.3	Geographic Diversification	68
3.4.4	REIT-Specific and Macroeconomic Control Variables	69
3.5	Methodology	72
3.5.1	Artificial Neural Network.....	73
3.5.2	Model Setup.....	74
3.5.3	Accumulated Local Effect (ALE) Plots.....	75
3.5.4	Ordinary Least Squares Regression.....	76
3.6	Results	77
3.6.1	Results ALE Plots.....	77
3.6.2	Optimized Regression Results.....	79
3.6.3	Performance Evaluation of Applied Models	83
3.7	Conclusion.....	83
3.8	Appendix	86
3.8.1	Appendix I – Rolling Local Beta Construction and OLS Results.....	86
3.8.2	Appendix II – ANN Hyperparameter Distribution	88
3.8.3	Appendix III – ALE Plots and Regressions of Continuous Variables....	89
3.8.4	Appendix IV – ALE Plots and Regressions of Categorical Variables.....	93
3.8.5	Appendix V – Variable Transformation to Optimize Polynomial OLS .	95
3.9	References	96
4	The Non-Linear Dynamics of Disclosure Tone and REIT Bond Yields	99
4.1	Abstract.....	99
4.2	Introduction.....	100
4.3	Literature Review and Hypothesis Development	104

4.3.1	Disclosure and Bond Markets.....	104
4.3.2	Hypotheses Development.....	105
4.4	Data.....	107
4.4.1	Derivation of Tone Measures.....	107
4.4.2	Dependent Variable and Control Variables.....	110
4.5	Methodology.....	114
4.5.1	Fixed Effects Regression.....	115
4.5.2	eXtreme Gradient Boosting (XGBoost).....	115
4.5.3	Accumulated Local Effects (ALE).....	116
4.6	Results.....	118
4.6.1	First Dimension of Non-Linearity.....	118
4.6.2	Second Dimension of Non-Linearity.....	122
4.6.3	Robustness Checks.....	124
4.7	Conclusion.....	126
4.8	References.....	129
5	Conclusion	136
5.1	Executive Summary.....	136
5.2	Final Remarks.....	142
5.3	References.....	144

List of Tables

Table 2.1	Variable Description	21
Table 2.2	Descriptive Statistics.....	23
Table 2.3	Applied Performance Metrics	25
Table 2.4	Subperiod Correlation Measures	31
Table 2.5	Chosen Variables and Related Literature.....	46
Table 2.6	Evaluation of XGBoost and LASSO	47
Table 3.1	Highest and Lowest Beta CBSAs and REITs	67
Table 3.2	Variable Description.....	70
Table 3.3	Descriptive Statistics.....	72
Table 3.4	Employed Performance Metrics.....	77
Table 3.5	Regression Results	82
Table 3.6	Performance Comparison of All-Year Models	83
Table 3.7	Rolling Beta Regression Results	87
Table 3.8	Hyperparameter Distribution	88
Table 3.9	Variable Transformations	95
Table 4.1	Variable Description.....	112
Table 4.2	Descriptive Statistics.....	113
Table 4.3	Pairwise Correlations.....	114
Table 4.4	Regression Results for Aggregated Tone Measures.....	119
Table 4.5	Regression Results for NPT and Bond Yield Spread	122
Table 4.6	Regression Results for Disaggregated Tone Measures.....	124
Table 4.7	Robustness Check – Tone Measures Based on Entire Financial Report.....	125
Table 4.8	Robustness Check – Interaction Effects.....	126

List of Figures

Figure 2.1	Average Weekly Return and Volatility	22
Figure 2.2	Expanding Window Approach	25
Figure 2.3	Absolute SHAP Values – Ten Most Important Features for Return.....	29
Figure 2.4	Absolute SHAP Values – Ten Most Important Features for Volatility	30
Figure 2.5	Clustered Feature Importance for Return and Volatility.....	33
Figure 2.6	Time-Varying Feature Importance of Macroeconomic Variables.....	35
Figure 2.7	Time-Varying Feature Importance of REIT-Specific Variables.....	37
Figure 2.8	ALE Plots for Return Drivers over the Subperiods.....	39
Figure 2.9	ALE Plots for Volatility Drivers over the Subperiods.....	42
Figure 2.10	R ² Evolution over Time	47
Figure 2.11	MSE Evolution over Time	48
Figure 2.12	MAE Evolution over Time.....	48
Figure 2.13	Time-Varying Feature Importance of Macroeconomic Variables.....	49
Figure 2.14	Time-Varying Feature Importance of REIT-Specific Variables.....	50
Figure 3.1	Distribution of Betas	67
Figure 3.2	Geographic Distribution of CBSA Beta.....	68
Figure 3.3	Distribution of HHI.....	69
Figure 3.4	Exemplary ANN Architecture.....	73
Figure 3.5	Cross-Sectional Expanding Window Approach	75
Figure 3.6	ALE Plot and Univariate Regression of Local Beta and HHI.....	78
Figure 3.7	Rolling Betas over Time.....	86
Figure 3.8	ALE Plots and Regressions of Continuous Variables	89
Figure 3.9	ALE Plots and Regressions of Categorical Variables	93
Figure 4.1	Positive, Negative, and Net Positive Tone Measures over Time	110
Figure 4.2	ALE Plots for Tone Measures.....	121

1 Introduction

1.1 Motivation and Background

Compared to the direct real estate investment market, indirect or listed real estate markets are known to be liquid, fast-moving, transparent, and consequently less prone to information asymmetries (Adair et al., 2006). Over the last decades, Real Estate Investment Trusts (REITs) have emerged as a prominent indirect real estate investment vehicle, offering investors fractional exposure to the real estate market through diversified portfolios of income-generating properties. In an environment marked by fluctuating interest rates, tightening regulatory requirements, and geopolitical tensions, REITs have gained renewed importance as a vehicle for inflation-hedged income and diversified exposure (Muckenhaupt et al., 2025).

REITs represent a unique intersection of physical real estate and financial markets, offering investors equity and debt instruments that provide diversified portfolios. On one hand, they are backed by physical assets, which generate income through leases and are subject to location-specific risks, regulatory constraints, and long-term market cycles. On the other hand, REITs are listed on public exchanges, offering daily liquidity, transparency, and exposure to market-wide sentiment (Clayton and MacKinnon, 2001). These characteristics makes their analysis distinct from that of other firms. However, REITs also fund their operations through common equity and public debt instruments. Unlike private real estate, REITs typically exhibit higher return volatility, as their pricing not only reflects asset-level cash flows but also broader financial conditions such as market volatility and global capital flows (Clayton and Mackinnon, 2003).

The United States (U.S.) REIT market is the largest worldwide. According to the National Association of Real Estate Investment Trusts (NAREIT), the market capitalization of U.S. REITs has grown from \$8.7 billion in 1990 to \$1.5 trillion in September 2025, representing about 3% of the total U.S. stock market capitalization. On the public debt side, bond financing has also become a notable capital source for REITs. In 2024, U.S. REITs issued \$48.1 billion in bonds, surpassing the pre-2020 ten-year average of approximately \$40 billion (NAREIT, 2025). Due to these dynamic developments, there is considerable interest among both practitioners and academics in understanding the dynamics of U.S. REITs as an investment.

On the equity side, REIT stock returns and their underlying volatility have been studied extensively in terms of asset pricing and return prediction (Letdin et al., 2019). Fama and

French's (1992) pioneering work on the determinants of expected non-REIT stock returns has served as a foundational framework also for numerous researchers in the realm of REITs. The relevance of accurately valuing REITs has become a robust area of discussion since the last financial crisis. Consequently, a substantial body of literature focuses on evaluating predictors of REIT returns and volatility in an effort to mitigate risks and contribute to economic stability (see, e.g., Nelling and Gyourko, 1998; Chui et al., 2003; Cotter and Stevenson, 2008; Nayar et al., 2024; Zhou and Kang, 2011; Anderson et al., 2021). However, literature on asset pricing primarily employs cross-sectional, linear regression models to investigate the effects of various drivers on equity returns, implicitly assuming stable and linear relationships (see, e.g., Fama and French, 2008; Lewellen, 2015). The highly dynamic nature of financial time series often violates these assumptions, as complex interactions, threshold effects, and regime shifts are common. This dissertation addresses this issue by employing a more flexible modeling framework that captures potential non-linearities and temporal dynamics in the relationship between risk factors and REIT returns.

In terms of debt financing, investors can enhance portfolio diversification by including REIT bonds, which provide exposure to real estate while complementing other asset classes and potentially reducing overall portfolio risk. While many researchers have examined general bond determinants, little attention has been given to REIT bonds and their constituents (Freybote, 2016; Gilstrap et al., 2022; Kozak et al., 2025). This creates a noticeable gap in the literature, given the growing importance of public debt offerings by REITs and the differing claim priorities of bondholders compared to equityholders (Hardin and Wu, 2010; Brown and Riddiough, 2003). Therefore, this dissertation examines how locational factors such as local market risk and geographic diversification affect REIT bond pricing, recognizing that REIT performance and risk exposure are inherently spatial. Moreover, it explores the role of informational efficiency in REIT public debt by incorporating textual sentiment from corporate disclosures, as qualitative information can convey nuanced and forward-looking signals beyond quantitative data (see, e.g., Han and Zhou, 2014; Jin et al., 2024; Lou, 2019).

Thus, this dissertation deliberately considers both equity and debt side of the REIT capital structure, as their claim hierarchies lead to fundamentally different pricing dynamics. Equity investors are residual claimants who bear greater volatility but are also exposed to upside potential, whereas bondholders have senior claims, focusing primarily on downside risk, and often respond differently to market uncertainty (see, e.g., Merton, 1974; Choi et al., 2022). These contrasting perspectives suggest that the same underlying real estate fundamentals may be valued differently across markets, resulting in distinct risk premia

and transmission channels for financial shocks. Therefore, an integrated view of REIT equity and debt markets is essential to understanding how information is absorbed, how risks are priced, and ultimately, how stability in the broader financial system can be supported. This dissertation aims to provide this integrated view while focusing on detecting potential non-linearities.

In terms of addressing the non-linearity issue, recent studies applying machine learning techniques present compelling evidence of non-linear relationships in asset pricing (Gu et al., 2020; Chen et al., 2019). These methods have proven well-suited for capturing the complex patterns that traditional linear models often miss (Bianchi et al., 2021). These non-linearities have been underexplored in REIT markets, even though stock return patterns or bond risk premia of REITs may exhibit threshold behavior or diminishing impacts. The dissertation addresses this gap by applying state-of-the-art machine learning algorithms to uncover the non-linear and time-varying relationships in REIT data. More specifically, artificial neural networks (ANNs) and eXtreme Gradient Boosting (XGBoost) algorithms are employed. These algorithms are often characterized as opaque “black boxes” due to their complexity and difficulty in interpreting their internal decision-making processes (Mullainathan and Spiess, 2017). To overcome this limitation, eXplainable Artificial Intelligence (XAI) tools were developed to investigate the modeled non-linear relationships between predictors and outcomes. In this dissertation, Shapley Additive Explanations (SHAP) and Accumulated Local Effect (ALE) plots are employed to gain a deeper understanding of the dynamics of REIT equity and debt markets. These methods have been used in the context of direct real estate markets (Lorenz et al., 2023; Kraemer et al., 2023), but their use in the REIT domain remains limited (Kozak et al., 2025).

Overall, the objective of this dissertation is to advance the understanding of pricing dynamics and risk drivers in the U.S. REIT equity and debt markets by addressing methodological and informational limitations in existing research. It does so by integrating modern machine learning techniques and XAI tools to model non-linear and time-varying relationships in financial data, while also incorporating qualitative, textual data sources to mitigate potential information asymmetries. By combining structured quantitative variables with corporate disclosure, the dissertation provides novel insights into the valuation of REITs, contributing to more accurate risk assessment, enhanced return prediction, and more informed decision-making for investors, REIT managers, and policymakers.

1.2 Course of Analysis and Research Questions

This section outlines the progression of this dissertation and introduces the objectives and main research questions that will be addressed in each of the three papers. All papers focus on explaining key determinants of U.S. REIT equity and debt markets at the firm level through the application of machine learning methods. By leveraging XAI techniques, the dissertation aims to uncover non-linearities and temporal dynamics that traditional econometric models may overlook.

Paper 1: Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach

Paper 1 focuses on the simultaneous analysis of the drivers of REIT stock returns and volatility during the modern REIT era (1991–2022). More precisely, the paper identifies dynamic patterns of REITs in relation to various market cycles. The analysis is based on REIT-specific financials and macroeconomic data, which are applied to an XGBoost algorithm. To unpack the decision-making process of the XGBoost algorithm, XAI methods, including SHAP and ALE plots, are employed. The central research questions can be summarized as follows:

- Which independent variables drive REIT returns and volatility the most, and in what direction do they drive the prediction?
- Do REIT returns and volatility exhibit time-varying features, particularly in times of crisis?
- Are REIT returns and volatility shaped by non-linear relationships with their independent variables?

**Paper 2: Location Matters: Local Real Estate Market Risk and Geographic
Diversification in REIT Public Debt**

Paper 2 shifts the focus to the public debt market of REITs and investigates how locational risk factors affect REIT bond pricing. More precisely, the paper examines the impact of local real estate market risk and geographic diversification within REIT property portfolios on bond risk premia. To explore the potentially non-linear nature of these relationships but maintain statistical rigor, the paper employs a combination of machine learning techniques (i.e., artificial neural networks) and traditional regression models (i.e., fixed effects regression). This dual approach enables a more in-depth analysis of the spatial characteristics of REITs and their relevance to credit risk assessment. The research questions are therefore as follows:

- Is there a local real estate market risk premium in REIT bonds, i.e., do investors demand higher compensation for REITs exposed to more volatile property markets?
- Does the relationship between geographic diversification and REIT bond risk premia exhibit non-linearities?
- What is the interaction effect between local real estate market risk and geographic diversification on REIT bond risk premia?

Paper 3: The Non-Linear Dynamics of Disclosure Tone and REIT Bond Yields

Paper 3 investigates how the disclosure tone in REIT financial statements (i.e., 10-Q/K filings) influences bond risk premia. The analysis focuses on two non-linearity dimensions within the relationship between sentiment and bond risk premia: First, it investigates whether the relationship is non-linear, particularly in the presence of excessively positive or negative language. Second, the paper decomposes disclosure tone into a fundamentals-based and non-fundamentals-based component to assess how investors respond to tone that does not align with financial performance. This approach provides novel insights into how qualitative information is priced in REIT debt markets. The central research questions can thus be stated as follows:

- Does the relationship between disclosure tone in REIT financial statements and REIT bond risk premia follow a non-linear (quadratic) pattern?
- Do fundamentals-based and non-fundamental components of disclosure tone affect REIT bond risk premia in different ways?
- Which component of disclosure tone provides greater informational value to bond investors in determining REIT bond risk premia?

1.3 Co-Authors, Submissions and Conference Presentations

This section provides an overview of co-authors, journal submissions, publication status and conference presentations.

Paper 1: Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach

Authors:

Hendrik Jenett, Cathrine Nagl, Maximilian Nagl, S. McKay Price, Wolfgang Schaefers

Submission Details:

Journal: The Journal of Real Estate Finance and Economics

Status: Accepted (03/14/2025) and published online ahead of print
(04/01/2025)

Conference Presentations:

This paper was presented at:

- the 40th Annual Conference of the American Real Estate Society (ARES) in Orlando, USA (2024)
- the 30th Annual Conference of the European Real Estate Society (ERES) in Gdansk, Poland (2024)

Awards and Funding:

This paper was awarded the “Alan J Ziobrowski Excellence in Doctoral Research Award” at the 40th Annual Conference of the American Real Estate Society (ARES) in Orlando, USA (2024).

Paper 2: Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt

Authors:

Hendrik Jenett, Jakob Kozak, S. McKay Price, Wolfgang Schaefers

Submission Details:

Journal: The Journal of Property Investment & Finance

Status: Accepted (09/25/2025) and published online ahead of print
(11/09/2025)

Conference Presentations:

This paper was presented at:

- the 41st Annual Conference of the American Real Estate Society (ARES) in Tucson, USA (2025)
- the 31st Annual Conference of the European Real Estate Society (ERES) in Athens, Greece (2025)

Paper 3: The Non-Linear Dynamics of Disclosure Tone and REIT Bond Yields

Authors:

Hendrik Jenett, Sophia Bodensteiner, Julia Freybote, Wolfgang Schaefers

Submission Details:

Journal: Journal of Real Estate Research

Status: Under review (09/21/2025)

Conference Presentations:

This paper will be presented at:

- the 32nd Annual Conference of the Pacific Rim Real Estate Society (PRRES) in Adelaide, Australia (2026)

1.4 References

- Adair, A., McGreal, S. and Webb, J. (2006).** Diversification Effects of Direct versus Indirect Real Estate Investments in the UK. *Journal of Real Estate Portfolio Management*, 12(2), 85–90.
- Anderson, J., Anderson, R., Guirguis, H. S., Proppe, S. and Seiler, M. J. (2021).** Time-Varying Correlations of REITs and Implications for Portfolio Management. *Journal of Real Estate Research*, 43(3), 317–334.
- Bianchi, D., Buechner, M. and Tamoni, A. (2021).** Bond Risk Premiums with Machine Learning. *The Review of Financial Studies*, 34(2), 1046–1089.
- Brown, D. T. and T. J. Riddiough (2003).** Financing choice and liability structure of real estate investment trusts. *Real Estate Economics*, 31(3), 313–346.
- Chen, L., Pelger, M. and Zhu, J. (2019).** Deep Learning in Asset Pricing. Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3350138.
- Choi, J., Richardson, M. and Whitelaw, R. F. (2022).** Capital Structure Priority Effects in Durations, Stock-Bond Comovements, and Factor Pricing Models. *The Review of Asset Pricing Studies*, 12(3), 706–753.
- Chui, A. C. W., Titman, S. and Wei, K. C. J. (2003).** The Cross Section of Expected REIT Returns. *Real Estate Economics*, 31(3), 451–479.
- Clayton, J. and MacKinnon, G. (2001).** The Time-Varying Nature of the Link between REIT, Real Estate and Financial Asset Returns. *Journal of Real Estate Portfolio Management*, 7(1), 43–54.
- Clayton, J. and MacKinnon, G. (2003).** The Relative Importance of Stock, Bond and Real Estate Factors in Explaining REIT Returns. *Journal of Real Estate Finance and Economics*, 27(1), 39–60.
- Cotter, J. and Stevenson, S. (2008).** Modeling Long Memory in REITs. *Real Estate Economics*, 36(3), 533–554.
- Fama, E. F. and French, K. R. (1992).** The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427–465.
- Fama, E. F. and French, K. R. (2008).** Dissecting Anomalies. *The Journal of Finance*, 63(4), 1653–1678.
- Freybote, J. (2016).** Real estate sentiment as information for REIT bond pricing. *Journal of Property Research*, 33(1), 18–36.
- Gilstrap, C., Petkevich, A., Sezer, O. and Teterin, P. (2022).** REIT Debt Pricing and Ownership Structure. *Journal of Real Estate Finance and Economics*, 64(4), 546–589.
- Gu, S., Kelly, B. and Xiu, D. (2020).** Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273.

Han, S. and Zhou, X. (2014). Informed Bond Trading, Corporate Yield Spreads, and Corporate Default Prediction. *Management Science*, 60(3), 675–694.

Hardin, W. G. and Wu, Z. (2010). Banking Relationships and REIT Capital Structure. *Real Estate Economics*, 38(2), 257–284.

Jin, S., Lee, H. W. and Xie, Y. A. (2024). Risk disclosures in annual reports and firms' default risk. *Journal of Corporate Accounting & Finance*, 35(3), 192–212.

Kozak, J., Nagl, C., Nagl, M., Beracha, E. and Schaefers, W. (2025). Does Real Estate Determine REIT Bond Risk Premia? *Journal of Real Estate Finance and Economics*, forthcoming.

Kraemer, B., Nagl, C., Stang, M. and Schaefers, W. (2023). Explainable AI in a Real Estate Context – Exploring the Determinants of Residential Real Estate Values. *Journal of Housing Research*, 32(2), 204–245.

Letdin, M., Sirmans, C. S., Sirmans, S. S. and Zietz, E. N. (2019). Explaining REIT Returns. *Journal of Real Estate Literature* 27(1), 1–25.

Lewellen, J. (2015). The Cross-section of Expected Stock Returns. *Critical Finance Review*, 4(1), 1–44.

Lorenz, F., Willwersch, J., Cajias, M. and Fuerst, F. (2023). Interpretable machine learning for real estate market analysis. *Real Estate Economics*, 51(5), 1178–1208.

Lou, Y. (2019). Disclosure of Pending Lawsuits and Bond Terms. *Management Science*, 65(4), 1926–1947.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2), 449–470.

Muckenhaupt, J., Hoesli, M. and Zhu, B. (2025). Real estate as an inflation hedge: new evidence from an international analysis. *The North American Journal of Economics and Finance*, 80(1), 102488.

Mullainathan, S. and Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, 31(2), 87–106.

Nayar, N., Price, S. M. and Shen, K. (2024). Macroeconomic Uncertainty and Predictability of Real Estate Returns: The Impact of Asset Liquidity. *Journal of Real Estate Research*, 46(1), 82–113.

NAREIT (2025). REIT Capital Offerings. National Association of Real Estate Investment Trusts. Available at: <https://www.reit.com/data-research/reit-market-data/reit-capital-offerings>.

Nelling, E. and Gyourko, J. (1998). The Predictability of Equity REIT Returns. *Journal of Real Estate Research*, 16(3), 251–268.

Zhou, J. and Kang, Z. (2011). A Comparison of Alternative Forecast Models of REIT Volatility. *Journal of Real Estate Finance and Economics*, 42(3), 275–294.

2 Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach

2.1 Abstract

Real Estate Investment Trust (REIT) returns and volatility have been extensively studied, yet typically in isolation from each other. Given that returns and volatility are generally connected in the eyes of investors, we simultaneously analyze the drivers of REIT returns and volatility over the modern REIT era (1991–2022) using an eXtreme Gradient Boosting (XGBoost) machine learning algorithm. We enhance transparency and utility through the application of explainable artificial intelligence (XAI) techniques, particularly Shapley Additive Explanations (SHAP) and Accumulated Local Effects (ALE), which unpack the decision-making process of the model. Our analysis reveals that while no single feature consistently dominates, the influence of various drivers fluctuates significantly over time. Notably, the importance of macroeconomic indicators generally diminishes, while REIT-specific characteristics become more influential during the sample period. Furthermore, market cycles (macroeconomic shocks) cause large deviations from otherwise long-run patterns. However, during these times of economic uncertainty, drivers of risk and return correlate more strongly in comparison to times of economic stability. Lastly, we find nonlinearities in the way the drivers influence returns and volatility. These insights have significant implications for investors, policymakers, and researchers as they navigate the evolving landscape of real estate investments.

Keywords: REIT Return, REIT Volatility, Machine Learning, XAI

2.2 Introduction

Understanding the determinants of Real Estate Investment Trust (REIT) returns and the underlying drivers of their volatility is crucial for investors, policymakers, and academics, particularly in times of economic uncertainty. While a substantial body of literature has emerged over the past several decades focusing on forecasting REIT equity and identifying their critical determinants (Letdin et al., 2019), these studies have largely treated the prediction of returns and the examination of return volatility as separate endeavors. That is, existing studies mainly focus on either predicting REIT returns or on examining REIT return volatility, even though theoretically speaking, returns should be a function of risk.

This study aims to provide a more comprehensive understanding of the dynamics between REIT return and risk across different market conditions. It will utilize the same features to explain both returns and volatility, allowing for insights into their interdependence. Utilizing data from 1991 to 2022, our analysis encompasses several major crisis periods, including the Asia crisis, dot-com bubble, Global Financial Crisis (GFC), and the COVID-19 pandemic, each of which is marked by heightened market uncertainty. Our joint analysis of REIT-specific and macroeconomic characteristics provides detailed insights into the composition of returns and volatility. This study not only considers the main drivers of REIT returns and volatility overall but also examines how features evolve over time and within subperiods, shedding light on REIT market cycles.

From a methodology viewpoint, the literature on REIT returns and volatility primarily adopts linear regression models to investigate the explanatory power of certain variables (see, e.g., Chui et al., 2003; Clayton and MacKinnon, 2003; Chung et al., 2016; Giacomini et al., 2017). Predictive models have primarily used company-specific financial variables such as dividend yields, the book-to-market ratio, and capitalization rates (Below et al., 2000; Chiang, 2015; Goebel et al., 2013), along with macroeconomic indicators like inflation and short-term interest rates (Akinsomi et al., 2016; Asteriou and Begiazzi, 2013; Chatrath and Liang, 1998; Chen et al., 2014; McCue and Kling, 1994). For REIT volatility, some papers examine interlinkages based on univariate models (see, e.g., Cotter and Stevenson, 2007; Zhou and Kang, 2011), while others use multivariate models to isolate specific volatility linkages among different REIT subsectors (Cotter and Stevenson, 2006).

Recent advancements have seen the application of machine learning techniques, which might be better suited to address the non-linearity, non-stationarity, and complex interactions typical of financial time series. For instance, Serrano and Hoesli (2007) demonstrate that artificial neural networks (ANNs) outperform traditional linear models in predicting REIT returns, a finding echoed by Loo (2019) and Zhang et al. (2023b). These

studies highlight the potential for superior predictive performance of machine learning models over traditional models.

Motivated by these advancements and recognizing the nascent application of machine learning in REIT prediction, we employ an eXtreme Gradient Boosting (XGBoost) algorithm for analyzing both REIT returns and volatility. This machine learning algorithm builds predictive models by systematically combining weak predictive models, typically decision trees, into a strong learner by focusing on correcting the errors of prior models through iterative refinement. Given the data-intensive nature of machine learning processes, we focus on REITs rather than direct real estate, as granular REIT data is available in short intervals. Model estimates are based on 114,806 cross-sectional weekly observations of 151 equity REITs that are constituents of the NAREIT index and incorporate 24 potential explanatory variables. To overcome the often-cited black box nature of machine learning models, we employ explainable artificial intelligence (XAI) techniques, specifically Shapley Additive Explanations (SHAP) and Accumulated Local Effects (ALE), to decode the decision-making process of our model. Ultimately, we calculate the Pearson correlation coefficient to investigate the risk-return behavior in different market cycles. In contrast to other studies that focus on a horserace to determine superior performance across methods, our aim is to reveal the magnitude and directional, non-linear impact of individual features on REIT returns and volatility, offering a more nuanced understanding of the dynamic interplay of various determinants over time.

These novel statistical techniques provide new insight into important questions that have heretofore mainly been studied separately for REIT returns and volatility. First, we ask which variables most significantly drive REIT returns and volatility, and in what direction? Second, we ask whether REIT returns and volatility exhibit time-varying features, particularly during periods of crisis? Third, we analyze whether the predictors and dependent variables interact non-linearly.

The results reveal that no single variable consistently emerges as the most important across the entire sample period and all market conditions. Rather, the characteristics influencing REIT returns and volatility display time-varying properties, indicating that different factors are dominant in various market phases. However, there is an important returns pattern that emerges; economic indicators play a major role in REIT returns during the earlier years, while REIT-specific financial variables have become more influential over time. In contrast, volatility is primarily driven by macroeconomic conditions over the whole sample period.

Furthermore, we find that crisis periods cause substantial deviations from otherwise long-run patterns. Nonetheless, the correlation between risk and return drivers increases during

times of economic uncertainty. Variables that capture macroeconomic conditions do a better job than REIT-specific variables at explaining return and risk simultaneously. This nuanced understanding enables investors to tailor their strategies according to prevailing market conditions, enhancing portfolio resilience in the face of economic fluctuations.

Altogether, our study advances the literature in multiple ways. First, by conducting a simultaneous examination of both REIT returns and volatility, our study bridges a crucial gap in existing research. We chose this dual approach over, e.g., risk-adjusted returns, to allow for a deeper understanding of the interdependencies between these two dimensions and facilitates a more nuanced analysis of how market conditions influence REIT performance. Second, the paper highlights how the drivers of REIT returns and volatility vary over time, since many previous studies have looked at the cross-section of returns only. This offers insights into how different economic environments, including periods of crisis, affect the securitized real estate market. Third, we leverage the cutting-edge XGBoost algorithm to overcome limitations in traditional analytical methods. We effectively capture and analyze the nonlinear dynamics and interaction effects within the data, providing a more sophisticated and accurate analysis of the factors influencing REIT markets. Fourth, we demonstrate the use of XAI techniques, specifically SHAP and ALE, to address the opacity often associated with machine learning. The use of XAI enhances transparency by providing clear insights into the contributions and non-linearities of each variable in our models. This is crucial for the adoption of machine learning in economic research and practice.

The results hold important implications for investors and policymakers. Investors can use the time-varying nature of the investigated characteristics to efficiently allocate their portfolio based on market cycles and thus mitigate risks and contribute to the overall stability of the REIT market. Policymakers can analyze policy effects based on the feature importance of individual characteristics, and implement new financial policies related to recent patterns.

The remainder of this paper is structured as follows. The “Literature Review” offers an overview of the related literature. The “Data” section describes the utilized dataset and data processing. A detailed description of the XGBoost and XAI methods to explain returns and volatility is in the “Methodology” section. The empirical findings about feature importance and model performance can be found in the “Results” section. Finally, the “Conclusion” section summarizes the paper.

2.3 Literature Review

2.3.1 REIT Performance Modeling with Traditional Methods

In recent decades, there has been considerable interest among researchers and professionals in understanding the predictability of Real Estate Investment Trust (REIT) stock returns and their volatility. However, these studies mainly focused either on returns or on volatility.

Early research into REIT return prediction focuses on their sensitivity to general stock and bond markets. Liu and Mei (1992) investigate REIT return predictability and their correlation with assets like small-cap stocks, large-cap stocks, and bonds. They identify significant forecasting variables such as the treasury bill yield and the yield spread between AAA bonds and treasury bills on equity REITs. Contrary, Li and Wang (1995) find the dividend yield to be among the top drivers of REIT returns, along with the term spread and default spread. They also establish a negative relationship between inflation and returns. Graham and Knight (2000) investigate the incremental information content of net income and cash flows for REITs, finding that funds from operations (FFO) carry greater informational value than net income.

Clayton and MacKinnon (2003) examine the link between REITs, financial assets, and direct real estate returns, noting a shift in REIT market dynamics towards dependence on small-cap stocks and real estate-related factors in the 1990s. Hung and Glascock (2010) analyze REITs' momentum returns, and find liquidity to be the predominant factor influencing momentum returns. Goebel et al. (2013) present evidence that after controlling for momentum, underlying REIT characteristics such as the book-to-market ratio, institutional ownership, and illiquidity are all strongly associated with REIT returns during the 1993 to 2009 window, while firm size is not. Anderson et al. (2015) examine the effect of property type diversification and find a strong positive relationship with return on assets and return on equity.

Beyond traditional factors, Akinsomi et al. (2016) incorporate sentiment and uncertainty indicators, finding significant relationships with REIT returns. Bond and Xue (2017) highlight investment and profitability as critical factors in explaining REIT returns. Beracha et al. (2019) investigate the extent to which REIT operational efficiency is related to operational performance, total risk, and stock return, using variables such as EBITDA, interest expenses, and total revenue. More efficient REITs outperform lower efficiency REITs in terms of risk-adjusted cross-sectional stock returns. Suzuki et al. (2023) look at

capital expenditures and how they affect REIT performance and find that REITs chose properties designated for divestment based on operating expenses and rental yields.

Research on REIT volatility has gained traction over the past two decades. Cotter and Stevenson (2006) observe strong correlations in uncertainty between equity REITs and mainstream stock indices. Based on their findings, Cotter and Stevenson (2008) find volatility persistence in the REIT sector, albeit lower compared to the S&P 500 index. Zhou and Kang (2011) advocate for dynamically adjusted long-memory models to better capture REIT volatility. Li (2012) identifies systematic risk, leverage, trading volume, and unexpected inflation as drivers of REIT return volatility, with systematic risk exhibiting higher magnitude in up markets. DeLisle et al. (2013) find systematic volatility risk not priced in equity REIT stocks, suggesting their use for portfolio hedging. Based on their findings, Li (2016) documents that expected REIT returns show an insignificant relationship both with the volatility of the stock market and the volatility of the Fama–French factors. Anoruo and Murthy (2017) examine the relationship between implied volatility as measured by the VIX and REIT returns, showing that they have significant negative effect on each other across all investigated frequencies. Bonato et al. (2021) underscores the role of realized skewness and kurtosis as indicators for realized volatility. Zhou (2020a, 2020b) compares the accuracy of different realized volatility measures in forecasting REIT return volatility. Lastly, Li and Zhu (2022) find that trading volume induces return volatility and that this effect is backed by transaction costs.

The analysis of past literature reveals that research has mostly focused on forecasting REIT returns or REIT volatility individually, thus neglecting the interdependent patterns between financial return and uncertainty. Furthermore, many of these studies looks at the cross-section of REITs and establishes a relationship between predictors and REIT performance at a certain point in time. However, it is crucial to observe whether the impact of certain variables changes over time and how it behaves during crisis periods, which is why we chose a time-varying setting in our study. Our independent variables for this study are chosen based on these studies on REIT returns and volatility. An overview of all factors and their related studies can be found in Appendix I.

2.3.2 Financial Products Performance Modeling with Machine Learning

Machine learning applications have recently proven to be a robust forecasting method. In the general finance literature, there have been various attempts to apply machine learning and XAI to financial products, including stock price prediction (see, e.g., Gu et al., 2020; Chen et al., 2024; Bryzgalova et al., 2020), credit risk management (Bussmann et al., 2021), cryptocurrency investments (see, e.g., Babaei et al., 2022; Nagl, 2024), fund

performance (Kovvuri et al., 2023), and debt financing (see, e.g., Bianchi et al., 2021; Lin and Bai, 2022).

As the form of dependence between the predictors and stock returns seems to be non-linear, the application of machine learning methods has also increased. Gu et al. (2020) compare many statistical and machine learning models and show that neural networks and regression trees perform best. Bryzgalova et al. (2020) use decision trees to group similar stocks together and use this information for portfolio sorts. Their strategy triples the Sharpe ratios, compared to traditional portfolio sorts. Feng et al. (2020) use hidden states of a neural network to reduce the dimension of their portfolio sorts, which automatically allow non-linearities and interactions. They find again a superior performance compared to a traditional portfolio sort. Rossi (2018) uses a machine learning algorithm to construct mean–variance efficient portfolios and document a superior performance. Chen et al. (2024) apply a combination of machine learning algorithms to estimate an asset pricing model for excess stock returns. Recently, also risk premiums of bonds and hedge funds are targeted with machine learning. Bianchi et al. (2021) employ a battery of machine learning algorithms to forecast bond returns and find neural networks among the best performing methods. Wu et al. (2021) apply machine learning methods to forecast hedge fund returns and use them for selection. Again, neural networks are the best choice. Summarizing, there is broad evidence that the dependence of predictors and stock returns is non-linear, which results in superior performance of machine learning methods. Furthermore, Freyberger et al. (2020) document a time-varying impact of some predictors.

2.3.3 Direct Real Estate Performance Modeling with Machine Learning

In the direct real estate literature, there has also been a significant increase in machine learning and XAI applications over the past years. Kok et al. (2017) demonstrate that automated valuation models (AVMs) leveraging machine learning and hyperlocal data outperform traditional appraisals in accuracy, cost, and timeliness, offering significant potential for real estate valuation. Kraemer et al. (2023b) examine the impact of spatial training level on the accuracy of AVMs and demonstrate that machine learning models benefit from higher spatial granularity. In a different study, Kraemer et al. (2023a) introduce the use of ALE plots with an XGBoost algorithm to improve transparency and understanding of value-determining features in residential real estate markets, highlighting non-linear relationships. Lorenz et al. (2023) explore how XAI techniques enhance understanding of rent predictions by revealing nonlinear relationships among hedonic features, especially size and age of a property, while examining both feature rankings and

feature effects on the prediction to provide deeper insights into the factors influencing rental values. Stang et al. (2023) compare traditional appraisal methods with an XGBoost, showing that the machine learning method outperforms conventional techniques in accuracy and scalability. Francke and van de Minne (2024) find that combining a random effects model with different machine learning algorithms for predicting commercial real estate values can reduce the average appraisal error found in literature. Further applications of non-parametric modeling techniques in direct real estate can be found in studies such as Yoo et al. (2012), Antipov and Pokryshevskaya (2012), McCluskey et al. (2013), and Yilmazer and Kocaman (2020).

2.3.4 Indirect Real Estate Performance Modeling with Machine Learning

Since 2003, machine learning has also been used in indirect real estate research. Brooks and Tsolacos (2003) find neural networks to provide the most accurate predictions for REIT returns compared to univariate time-series models and an VAR model. Serrano and Hoesli (2007) look at these same three models and confirm the superiority of neural networks, especially when considering a large array of factors. Feng and Li (2016) create a REIT portfolio with 20% above market returns using a Support Vector Machine (SVM). Wang et al. (2016) apply SVM and back-propagation neural networks to the Singaporean REIT market. Their findings indicate that machine learning approaches outperform simple regression models in predicting REIT performance and that the choice of variables significantly impacts the simulation results. Habbab et al. (2023) examine the performance of two traditional regression methods and three machine learning methods, and confirm that the machine learning methods (i.e., XGBoost and neural network) works best for predicting REIT stock prices. Moreover, Loo (2019), and Zhang et al. (2023b) demonstrate that non-parametric machine learning models outperform linear regression models in forecasting REIT returns.

In summary, the studies on indirect real estate reviewed conduct methodological horse races. However, none of the studies utilized XAI to unpack the decision-making process behind the models. This highlights the necessity of applying XAI to explain REIT returns, particularly with an emphasis on the data's time-varying properties.

2.4 Data

The dataset used in this study consists of 114,806 weekly observations of 151 equity REITs that are constituents of the FTSE NAREIT Index from October 1990 to December 2022. We construct our sample not based on the NAREIT Index but on the individual REITs within

the index, allowing for a cross-sectional analysis. Our sample is based on three different datasets, namely (1) REIT market data, (2) REIT accounting data, and (3) macroeconomic data.

The REIT market data comes from Eikon and comprises the weekly closing price and market capitalization for each REIT. The returns are calculated as the percentage change from one week's close price to the previous weeks. The weekly volatility for each REIT is calculated based on a separate dataset comprising the daily returns, which are then merged into the weekly dataset. Since the actual volatility is unobserved, an observable proxy is needed. Following Patton (2011) and Zhou (2020a), we employ daily squared open-to-close returns as the proxy for the latent volatility, assuming a zero mean of returns. Despite being subject to noise, the daily squared open-to-close returns are unbiased and free of microstructure noise. Quarterly accounting data is sourced from Eikon too and was obtained in accordance with prior research on explaining REIT returns and volatility. We decided to account for company leverage indirectly by incorporating total debt outstanding and market capitalization, allowing us to gain separate insights into both equity and debt components. In unreported results we examined leverage in the traditional sense (as a combined construct) and obtained consistent results. Most macroeconomic variables are from the Federal Reserve Economic (FRED) monthly database, as presented by McCracken and Ng (2016). The real estate buying conditions index comes from a survey of consumers by the University of Michigan. The REIT return and volatility are the dependent variables, while the remaining 24 characteristics represent the explanatory variables. For robustness reasons, we also fit a model using risk-adjusted returns as the dependent variable to evaluate whether this changes our results. Given the fact that returns are a function of risk, we use the same set of characteristics to explain both returns and volatility.

We chose a relatively long period of 32 years for analysis since we want to cover multiple crisis periods (i.e., the dot-com bubble, GFC) to see whether and how they affected the dynamics of REIT returns and volatility. To account for survivorship bias, we include all REITs that have operated at some point during the period under investigation, regardless of whether they started operating after 1990 or ceased operations before 2022.

Various steps are undertaken to preprocess the data. The returns and volatility are lagged as we want to use the current information to predict the following week, see Gu et al. (2020). In the early years, the number of complete information sets is relatively low but increases gradually towards the end of our sample period. We handle missing values by imputing data up to one year, meaning that an existent value is being perpetuated until a new value occurs or 52 weeks have passed. REIT-specific characteristics are winsorized at

the 5th and 95th quantile to prevent outliers from distorting the results. This process is conducted on a weekly basis to ensure that future data points or REITs not included in the sample at a specific time are not incorporated. Finally, we standardize the features before feeding them into the model for training and testing. An overview of the utilized variables is provided in Table 2.1.

Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach

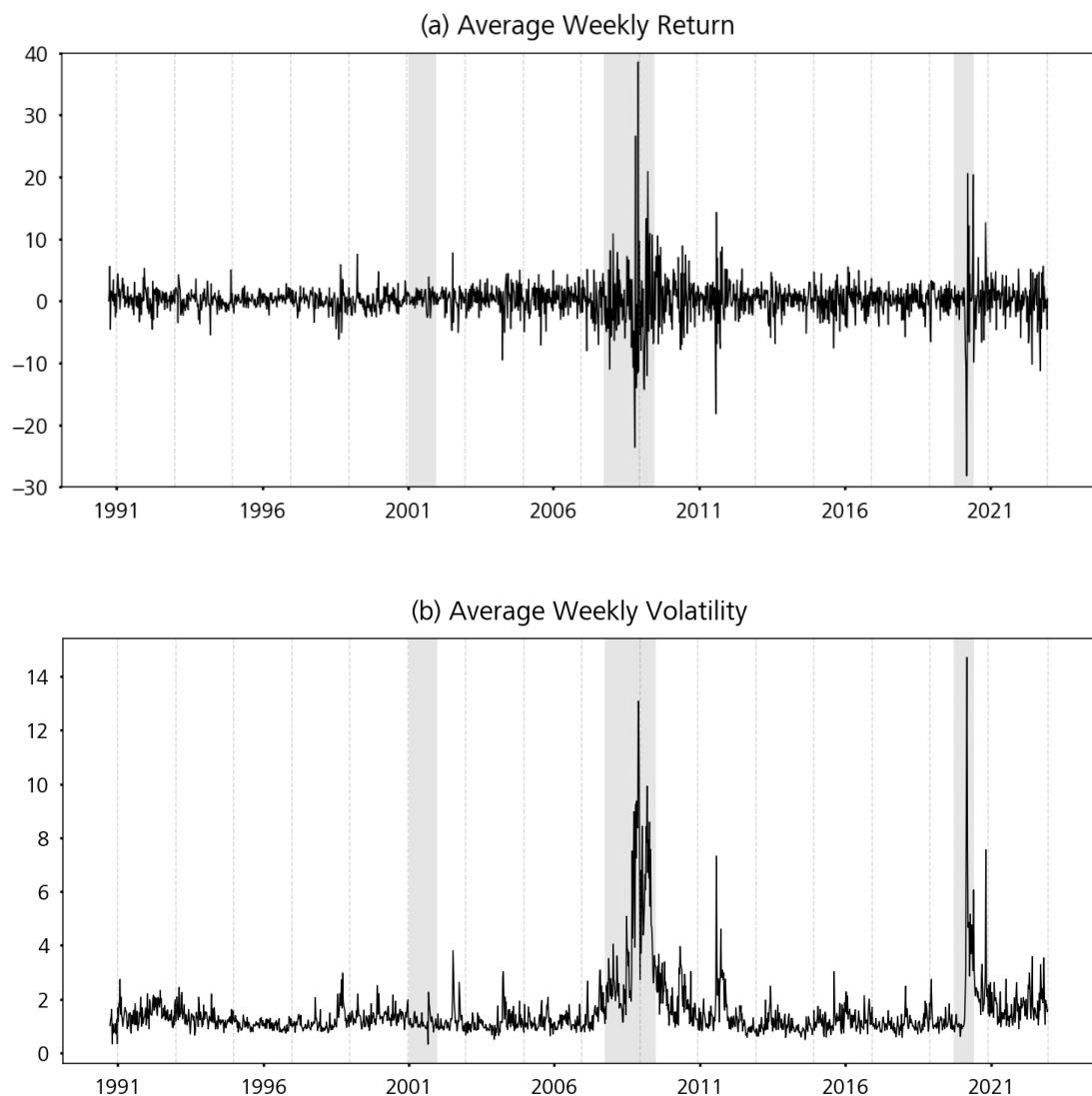
Table 2.1 Variable Description

Variable	Description	Frequency
Dependent Variables		
REIT Return	Expressed as the current adjusted price of the REIT stock minus the original adjusted price of the stock (week t-1) divided by the original adjusted price	weekly
REIT Volatility	Expressed as the standard deviation of the five daily returns for one week	weekly
Real Estate Variables		
REIT Property Type	8 REIT sectors based on property portfolio, defined by NAREIT (e.g., office, retail, diversified), reference category is diversified	
Accounting Variables		
Market Capitalization	Measure of company size, expressed as total dollar market value of a REIT's outstanding shares of stock	weekly
Total Assets	As per balance sheet	quarterly
Return on Asset	Net Income divided by total assets	quarterly
Funds From Operations	Cash flow from operations	quarterly
Normalized EBITDA	Earnings before interest, taxes, depreciation, and amortization	quarterly
Total Debt Outstanding	Dollar amount that REIT owes their lenders and creditors, including accrued interest	quarterly
Cash and short-term Investments	Cash and financial investments that can easily be converted to cash	quarterly
Interest Expense	Interest payable on any borrowings	quarterly
Total Revenue	Gross revenue	quarterly
Property, Plant & Equipment	Long-term assets in the balance sheet (e.g., real estate)	quarterly
Capex	Capital Expenditures	quarterly
Dividends	Distribution of a REIT's earnings to its shareholders	quarterly
Dividend Yield	Financial ratio that shows how much a REIT distributes relative to its stock price	weekly
Retained Earnings	Cumulative net earnings of a company after accounting for dividend payments	quarterly
Operating Income	Total revenue minus any operating expenses	quarterly
Book-to-market ratio	Market capitalization divided by Total Assets	weekly
Macroeconomic Variables		
S&P 500	Return of the S&P 500 composite index	monthly
Fed Funds Rate	Interest rate by Federal Reserve	monthly
GS10	10-Year Treasury Rate	monthly
CPI	Consumer Price Index (Inflation)	monthly
Buying Conditions	Sentiment Index expressed as % of "Good time to buy" minus % of "Bad time to buy" plus 100	monthly
VIX	CBOE volatility index	monthly
Small Cap Return	Wilshire U.S. small-cap total market return	monthly

Notes: This table provides an overview of every variable employed in this study. The first column shows the variables' names. The second column describes the variables. The last column shows the variable's available frequency. The final dataset contains weekly observations, which means that for monthly or quarterly observations, the respective value is filled in for all weeks of the respective month or quarter. We chose weekly observations because of the data noise trade-off. Daily data is usually very noisy and can negatively affect the predictions. Monthly observations often smooth out potential short-term fluctuations. We also tested the algorithm on monthly data but the results in terms of model fit were worse compared to weekly data.

Figure 2.1 (a) displays the average weekly stock return for all REITs in the sample and its temporal evolution. Given an average return of 0.14% across the entire period, the assumption of a zero mean holds for our sample. However, some anomalies can be observed during periods of economic uncertainty, where returns deviate tremendously from times of regular economic activity. The periods affected by the sample are the GFC and COVID-19 pandemic. The evidence of volatility clustering, meaning that high levels of volatility tend to be concentrated together, becomes apparent when examining Figure 2.1 (b). Table 2.2 displays the descriptive statistics of all features to help better understand the variables used in this study.

Figure 2.1 Average Weekly Return and Volatility



Notes: The plot on top displays the weekly average return of all REITs in % within the sample, while the bottom plot displays the weekly average volatility (standard deviation). The gray shaded areas illustrate periods of economic uncertainty as indicated by NBER: <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions>. The period from Q1 2001 to Q4 2001 marks the dot-com bubble burst. The period from Q4 2007 to Q3 2009 marks the Global Financial Crisis. The period from Q4 2020 to Q2 2020 marks the COVID-19 pandemic.

Table 2.2 Descriptive Statistics

Variable	Unit	Mean	Media	SD	Min	Max
REIT Return	Percentage	0.14	0.20	4.43	-50.33	73.60
REIT Volatility	Integer	1.57	1.21	1.48	0	30.52
REIT Property Type	/	/	/	/	/	/
Market Capitalization*	Integer	5,066	1,820	9,936	5.51	137,457
Total Assets*	Integer	4,749	2,688	5,467	139.13	32,674
Return on Asset	Percentage	0.67	0.61	0.77	-3.20	3.85
Funds From Operations*	Integer	2.55	2.02	1.82	-1.03	9.49
Normalized EBITDA*	Integer	93.74	51.33	111.13	-45.89	715.16
Total Debt Outstanding*	Integer	2,319	1,271	2,737	7.42	14,644
Cash and short-term Investments*	Integer	108.79	27.60	190.32	0.038	1,682
Interest Expense*	Integer	25.81	14.76	28.63	-0.073	122.01
Total Revenue*	Integer	175.60	91.66	218.68	4.32	1,240
Property, Plant & Equipment*	Integer	3,781	2,272	4,192	3.46	21,883
Capex	Integer	10.34	6.57	10.80	0.00	107.86
Dividends	Percentage	1.69	1.44	1.24	0.00	6.85
Dividend Yield	Percentage	0.06	0.05	0.04	0.00	0.53
Retained Earnings*	Integer	-446.14	-155.31	727.60	-4,201	425.87
Operating Income*	Integer	43.37	20.99	59.69	-106.03	331.67
Book-to-market ratio	Percentage	0.00	0.00	0.00	0.00	0.00
S&P 500	Percentage	0.66	1.16	3.74	-20.39	12.02
Fed Funds Rate	Percentage	1.69	1.00	1.92	0.05	8.11
GS10	Percentage	3.13	2.75	1.54	0.62	8.72
CPI	Percentage	0.22	0.21	0.30	-1.77	1.38
Buying Conditions	Index	138.50	148.00	29.87	37.00	179.00
VIX	Index	20.07	18.23	7.92	9.51	59.89
Small Cap Return	Percentage	0.88	1.04	5.12	-25.48	21.49

Notes: This table shows the summary statistics of the whole dataset from 1990 to 2022, containing equity REIT data and accompanying macroeconomic indicators. The summary of the REIT-specific variables reflects the results after winsorizing monthly at the 5th and 95th percentage quantiles. The asterisk denotes that these values are in millions. For all values above 1,000 no decimals are displayed.

2.5 Methodology

2.5.1 Machine Learning Method

The eXtreme Gradient Boosting (XGBoost) algorithm is a tree-based ensemble learning method. The origin of XGBoost can be traced to Chen and Guestrin (2016), who developed this algorithm as a response to bias and variance problems associated with single decision trees. XGBoost can solve complex issues such as classifications, pattern recognition, and regression-based tasks, exhibiting a successful track record of outperforming traditional methods and other machine learning algorithms in the field of real estate valuation (see, e.g., Kok et al., 2017; Stang et al., 2023; Kraemer et al., 2023a, 2023b)). However, the application of XGBoost in predicting REIT performance has been limited to Habbab et al. (2023), and Habbab and Kampouridis (2024). They use past REIT prices and technical analysis indicators to predict future REIT prices and create an optimal portfolio.

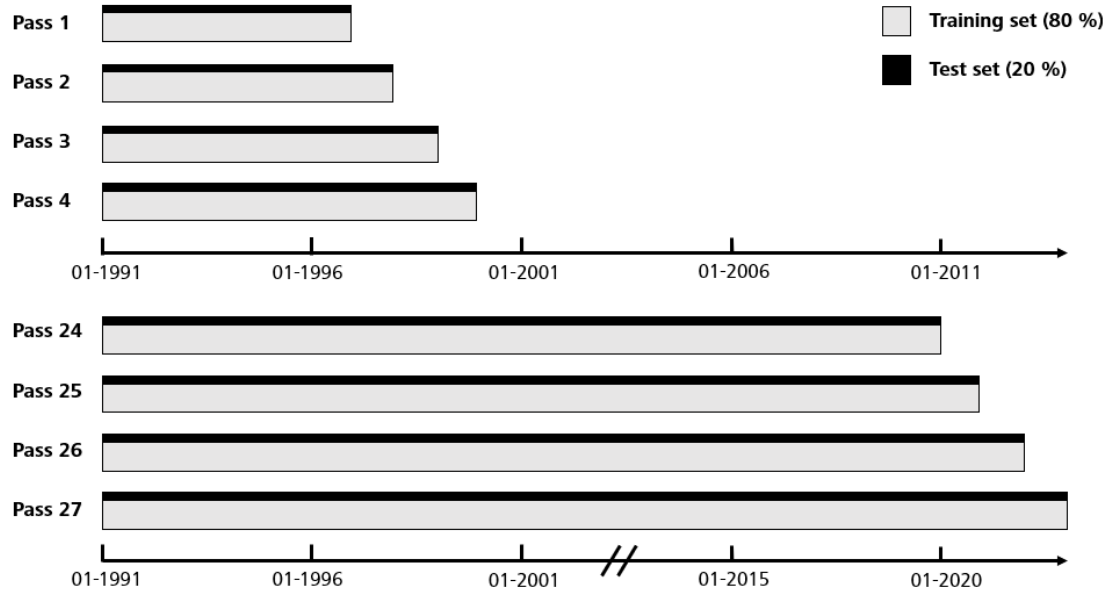
Tree-based ensemble learning algorithms combine the results of several decision trees to produce superior performance compared to using a single decision tree. The basic idea behind ensemble learning algorithms is that individual so-called weak learners, in our case, single decision trees, can be combined to achieve one strong learner. In the case of the XGBoost, the following technical expression can be stated:

$$h(y|x) = \sum_{i=1}^M u_m h_m(y|x), \quad (2.1)$$

where h is the strong learner, h_m are the individual weak learners and u_m is used to weight the weak learners. M is the number of single trees, x is the full features space and y the response variable. Boosting is a type of ensemble learning in which the weak learners h_m are trained successively. The algorithm starts with one model and uses the errors made to improve subsequent trees. Gradient boosting uses the so-called gradient descent algorithm by adding new trees to minimize the loss of the model. The eXtreme Gradient Boosting is a computationally effective and highly efficient version of gradient boosting. The advantage of XGBoost compared to parametric and semi-parametric models is that it can recognize complex patterns like non-linearities or higher-order interaction terms within a large amount of data, thus reducing the need for manual optimization. For more information about tree-based methods, ensemble learning and gradient boosting, the interested reader is recommended to read Hastie et al. (2009).

The dataset is split into two subsets: (1) The training data is used for initial model estimation, where every input is labeled with a respective output so that the algorithm learns patterns in the data; (2) The test data is used for estimation and performance evaluation. A hyperparameter search is conducted at the beginning of each year. We implement an expanding window approach based on Gu et al. (2020) to evaluate both return and volatility outcomes. The expanding window process is illustrated in Figure 2.2. During the first iteration, the data from the first five years is utilized for training (80% of the five-year data), while the remaining 20% are reserved for testing. With every iteration, the training dataset expands by one year. The main advantage of this approach is that the algorithm tests on unseen data and can provide unbiased, robust results.

Figure 2.2 Expanding Window Approach



We employ a LASSO (Least Absolute Shrinkage and Selection Operator) regression to benchmark the performance of our machine learning algorithm. This linear model addresses overfitting and allows obtaining a better bias-variance trade-off by adding a penalty term to the traditional linear regression model (Tibshirani, 1996). However, LASSO only models the expectation, i.e., the return of a sequence, but not the uncertainty respectively the volatility. Therefore, we benchmark only the return estimates against LASSO and use other metrics to validate the estimates volatility based on XGBoost.

Previous studies on forecasting REIT returns have proven the superiority of nonparametric models over conventional regression methods, mainly due to their ability to address the non-linearity and non-stationary data properties (Brooks and Tsolacos, 2003; Habbab et al., 2023; Loo, 2019; Serrano and Hoesli, 2007). Table 2.3 shows the different performance metrics used to evaluate the forecasts the XGBoost and the LASSO regression produced.

Table 2.3 Applied Performance Metrics

Metric	Formula	Description
R^2	$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\bar{y})^2}$	Explained variance of the dependent variable. A higher value indicates a better model fit
Mean Squared Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$	Sum of squared deviations from actual to predicted values. A lower value indicates a better model fit.
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $	Sum of absolute deviations from actual to predicted values. A lower value indicates a better model fit.

Despite the fact that XGBoost seems to work well in many cases, its black box nature often limits practitioners' willingness to apply them to real-life scenarios. XAI is a branch of research that addresses the obscure nature of how the model learns patterns in the data and detects causal relationships between the variables. To the best of our knowledge, this study is the first to open the black box of an XGBoost for modeling REIT returns and volatility. Following Lorenz et al. (2023), we employ model-agnostic tools to assess both feature importance and feature effects. Feature importance reveals and ranks the individual importance of variables for return and volatility prediction. Feature effects graphically illustrate the relationship between return and volatility and the influencing variables.

2.5.2 Feature Importance

We use Shapley Additive Explanations (SHAP), developed by Lundberg and Lee (2017), to evaluate feature importance. SHAP is a local post-hoc model-agnostic technique based on a coalitional game theory method named Shapley values. It allows us to derive the impact of individual features on the model's estimates and their evolution over time. Some of the main advantages are that SHAP values provide fair and consistent results. Fairness relates to extracting each feature's fair contribution to the overall effect using the SHAP values. Consistency implies that the total sum of all features SHAP values corresponds to the difference between the model estimates and the expected average outcome. Additionally, SHAP values account for non-linearities and higher-order interactions in the data, making them especially valuable for complex machine-learning tasks (Lundberg and Lee, 2017).

The SHAP value of a feature represents its influence on the payout, obtained by summing its weighted contributions across all possible combinations of feature values. Mathematically, the SHAP value of an observation i and feature j can be represented as follows:

$$\phi_{i,j} = \sum_{S \subseteq \{x_{i,1}, \dots, x_{i,p}\} \setminus \{x_{i,j}\}} \frac{|S|! (p - |S| - 1)!}{p!} (f(S \cup \{x_{i,j}\}) - f(S)), \quad (2.2)$$

where $\phi_{i,j}$ denotes the Shapley value for feature observation i , S stands for a coalition, and $x \in \mathbb{R}^{p \times n}$ depicts the set of all features containing of p variables and n observations. The final part of the equation calculates the marginal contribution of feature i on the predictions. Here, $f(S \cup \{x_{i,j}\})$ represents the XGBoost model trained with feature i present, while another XGBoost model $f(S)$ is trained with the feature withheld. Since the effect of withholding a feature relies on the presence of other features in the model, the preceding differences are computed across all possible subsets $S \subseteq \{x_{i,1}, \dots, x_{i,p}\} \setminus \{x_{i,j}\}$.

Finally, the result is weighted based on the number of features present in the coalition S . SHAP values are calculated individually for each observation in the training dataset, hence cannot directly be interpreted on a global level. They can be interpreted as the influence of an individual feature on the prediction made by the machine learning algorithm compared to the average prediction. To achieve global, model wide interpretability of SHAP values, individual observations can be averaged for each feature, allowing a ranking of feature importance.

2.5.3 Feature Effects

SHAP values only give us insights into which features contribute the most to the prediction. Accumulated Local Effects (ALE) plots, on the contrary, shed light on how a single feature contributes to the prediction and can therefore detect non-linear relationships. These plots provide insights into how an independent variable influences the average prediction of the dependent variable across all its values in the dataset. A drawback of other commonly used methods, such as Partial Dependency Plots (PDP) by Friedman (2001), is that they assume that variables are uncorrelated, which is rarely the case in real-world applications in finance and real estate. ALE plots, introduced by Apley and Zhu (2020), address this issue by allowing for correlated features. As a result, ALE plots offer a more robust approach to understanding the effects of features on predictions. They are also gaining increasing attention in the direct real estate literature, as evidenced by studies such as Kraemer et al. (2023a) or Lorenz et al. (2023).

The total range of observed values for each independent variable $x_r \in \mathbb{R}^{N \times 1}$, is split into K bins. We define $Z_{r,k}$ as the $\frac{k}{K}$ quantile of its empirical distribution with $Z_{r,0}$ being the minimum and $Z_{r,K}$ the maximum value of Z_r . Moreover, assume that $S_{r,k}$ defines the set of values within the left open interval from $Z_{r,k-1}$ to $Z_{r,k}$ with $n_{r,k}$ being the number of observations within the interval $S_{r,k}$. We define $k(x_r)$ as an index that returns the bin for a value of x_r . Consequently, the (uncentered) accumulated local effect is given by:

$$g_{ALE}(x_r) = \sum_{k=1}^{k(x_r)} n_{r,k}^{-1} \sum_{i \in S_{r,k}} [f(Z_{r,k}, X_{\setminus r,i}) - f(Z_{r,k-1}, X_{\setminus r,i})], \quad (2.3)$$

where $x_{\setminus r} \in \mathbb{R}^{N \times P-1}$ denotes the set of features without the feature r of P variables and $f(\cdot)$ the network's prediction. The quantity in the squared brackets delineates the prediction of $f(\cdot)$ if the observation i is replaced with $Z_{r,k}$ in the minuend and the prediction with $Z_{r,k-1}$ instead of the true value i in the subtrahend. These differences are summed over every observation in $S_{r,k}$. Since this has to be done for each bin k , $g_{ALE}(x_r)$

denotes the sum of the inner sums weighted by the number of observations in each bin (Apley and Zhu, 2020). The centered ALE main effect has a zero mean with respect to the marginal distribution of x_r and can be written as:

$$\Theta_{ALE}(x_r) = g_{ALE}(x_r) - N^{-1} \sum_{i=1}^N g_{ALE}(x_{r,i}). \quad (2.4)$$

Centering the individual ALEs ensures that the final ALE plot remains interpretable. In this context, the plot consistently illustrates the impact of a feature as a function of the average prediction of the XGBoost across the entire feature space. The visual analysis of ALE plots allows us to detect any non-linear relationships between the independent variables and the predicted return and volatility.

2.6 Results

Our study aims to analyze the time-varying properties of REIT returns and their underlying volatility by opening the black box nature of our machine learning model and interpreting whether and how return and volatility drivers reveal information about market cycles. First, the most predictive characteristics for REIT return and volatility are presented. Second, we analyze selected characteristics over time to understand their patterns comprehensively. Third, individual feature effects on the prediction are illustrated and analyzed. In contrast to other publications, the focus of this study is not to determine whether machine learning outperforms a traditional regression. Yet, a brief discussion on model performance can be found in Appendix II.

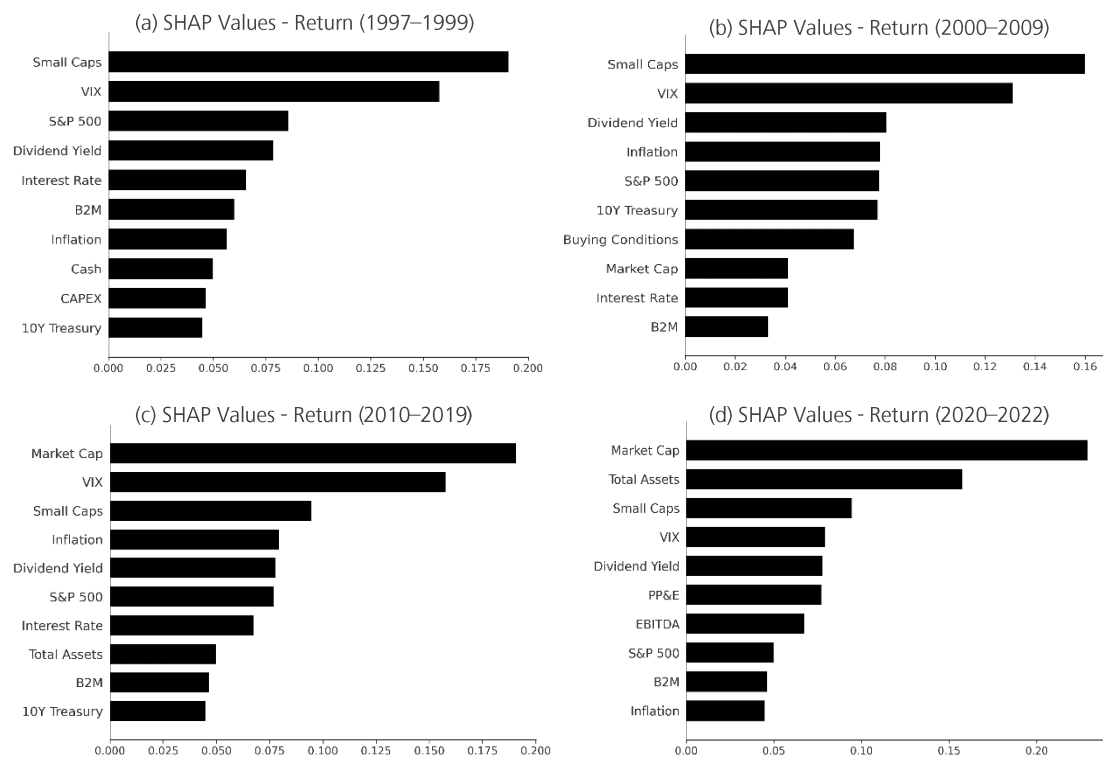
2.6.1 Global Absolute Feature Importance

Our analysis spans a relatively long period from 1997 until 2022, during which different market conditions have been predominant and multiple crises have occurred. To account for the various market cycles and guarantee a profound analysis, we chose to split the sample into four subperiods, namely (1) 1997 to 1999: Asia crisis, high-interest regime, (2) 2000 to 2009: Dot-Com boom, GFC, (3) 2010 to 2019: low-interest regime, real estate boom, and (4) 2020 to 2022: COVID-19 pandemic, the war in Ukraine, high-interest regime.

The absolute SHAP values of the ten most influential features for return explanation for each subperiod are illustrated in Figure 2.3. The key finding is that macroeconomic factors are the predominant drivers of REIT returns, especially in the earlier years. While the displayed REIT-specific features only contributed about 25% to returns in the late 1990s,

nowadays, around 60% of the return is driven by REIT-specific variables. The increase in REIT-specific variables start in 2010 and can be attributed to the gradual maturation of the REIT market. The market has become much more transparent and liquid over the years, therefore investors base their decisions more on factors like operational efficiency, financial performance and asset class strength. There seems to be a strong direct link between REITs and small-cap stocks. However, the relationship has decreased over time and has only been about half as strong in more recent years. The S&P 500 index has a lower but also relatively high overall impact and remains mostly stable throughout the observed time frames. These findings contradict Zhang et al. (2023a), who state that REITs behave little like stocks. The CBOE Volatility Index (VIX) also substantially influences returns in the first three periods, while it loses significance in the 2020 to 2022 period. Similar patterns of declining relevance can be observed for inflation and interest rates. On the REIT-related side, the dividend yield has a relatively strong and time-persistent impact on financial returns. This finding expands on Chiang (2015), who established a positive relationship only between 1993 to 2011. While the size variables market cap and total assets could be neglected in the first period, its influence heavily increased since 2010, when REIT size has become the predominant driver of returns. The impact of the book-to-market ratio remains consistent, except for the period between 2000 and 2009.

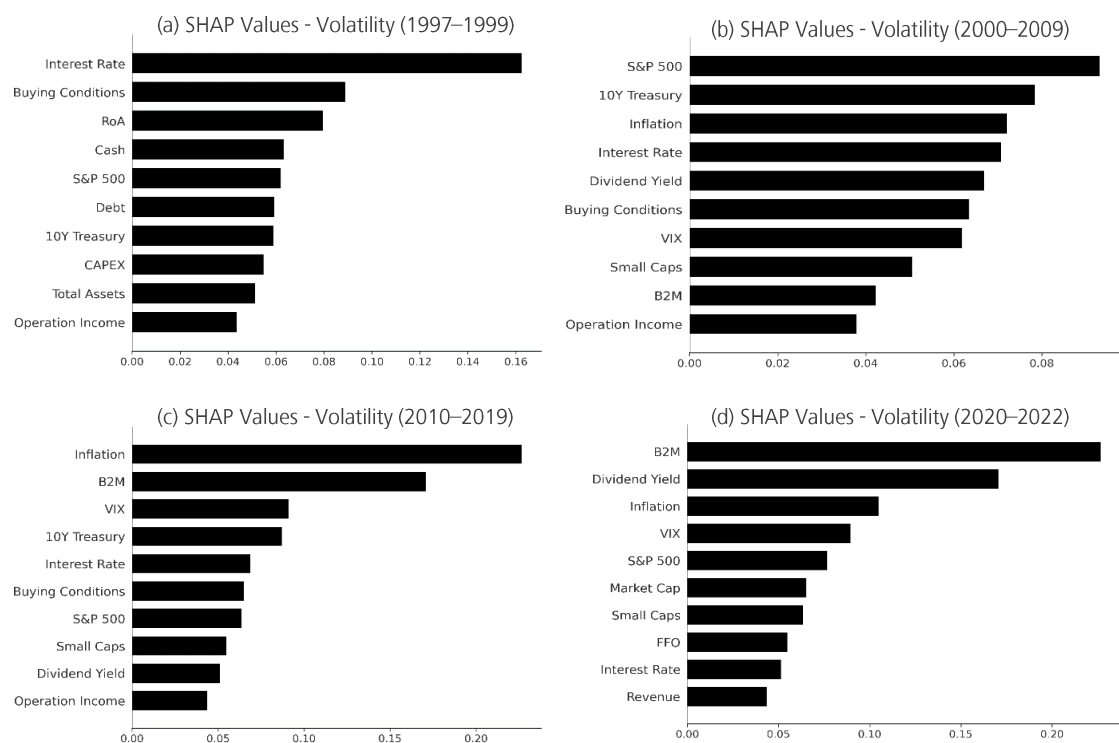
Figure 2.3 Absolute SHAP Values – Ten Most Important Features for Return



The SHAP values for the volatility are shown in Figure 2.4. Overall, the macroeconomic environment plays a crucial role among the top ten volatility drivers. Strikingly, many dominant variables in the first subperiod do not appear in any subsequent periods. This can be explained by the uncertainty in the market during the Asia crisis. The period was marked by high-interest rates, which were the dominant factor for the uncertainty of REITs. After 2000, the interest rate has lost significance but still plays a mentionable role until today. On the contrary, inflation has driven return uncertainty since 2000, with a peak in the third subperiod. The impact of sentiment about real estate buying conditions has declined over the years and is no longer a determining factor. Looking at REIT-specific factors for the 2000 to 2009 and 2010 to 2020 periods, only dividend yields, and the book-to-market ratio contributed to the volatility. The impact of the book-to-market ratio has previously been confirmed by DeLisle et al. (2013) for the 1996 to 2010 period. As opposed to that, 56% of the volatility can be traced to accounting variables in the last subperiod, highlighting the increased importance of company-specific values.

The most important features of the model using risk-adjusted returns closely align with those extracted from Figure 2.3 and Figure 2.4, providing supporting evidence for the robustness of these results.

Figure 2.4 Absolute SHAP Values – Ten Most Important Features for Volatility



In order to dive deeper into the interdependencies between return and risk drivers within the subperiods, we calculate the Spearman correlation for all subperiods and drivers, as indicated in Table 2.4. The results reveal that the correlation between risk and return drivers increases during periods of economic crisis (subperiods 2 and 4). This heightened correlation during the dot-com bubble, the Global Financial Crisis, and the COVID-19 pandemic indicates that economic uncertainty strongly align risk and return drivers. In contrast, more stable economic periods (subperiods 1 and 3) show lower correlations, reflecting a decoupling of risk and return influences. Economic uncertainty amplifies the sensitivity of REITs to macroeconomic variables, as policy responses like interest rate cuts or fiscal stimulus broadly impact both returns and volatility. Market sentiment often shifts uniformly, leading to synchronized reactions to economic indicators. Financial stress and liquidity constraints further tighten the relationship between these drivers, as investor behavior becomes more risk-averse, causing simultaneous fluctuations in returns and volatility. These findings underscore the importance of considering both risk and return in a unified framework, especially during times of economic instability.

Table 2.4 Subperiod Correlation Measures

Subperiod	Spearman Correlation
1997–1999	0.61
2000–2009	0.85
2010–2019	0.70
2020–2022	0.78

2.6.2 Evolution of Feature Importance

The next question to be answered is whether REIT features exhibit time-varying patterns, particularly in times of economic uncertainty. As alluded to previously, the analysis of the four subperiods reveals insights about significant changes in the decisive factors for REITs. We now want to visualize the annual changes of macroeconomic and REIT-specific return and volatility drivers to better understand the market dynamics and relationship between risk and return.

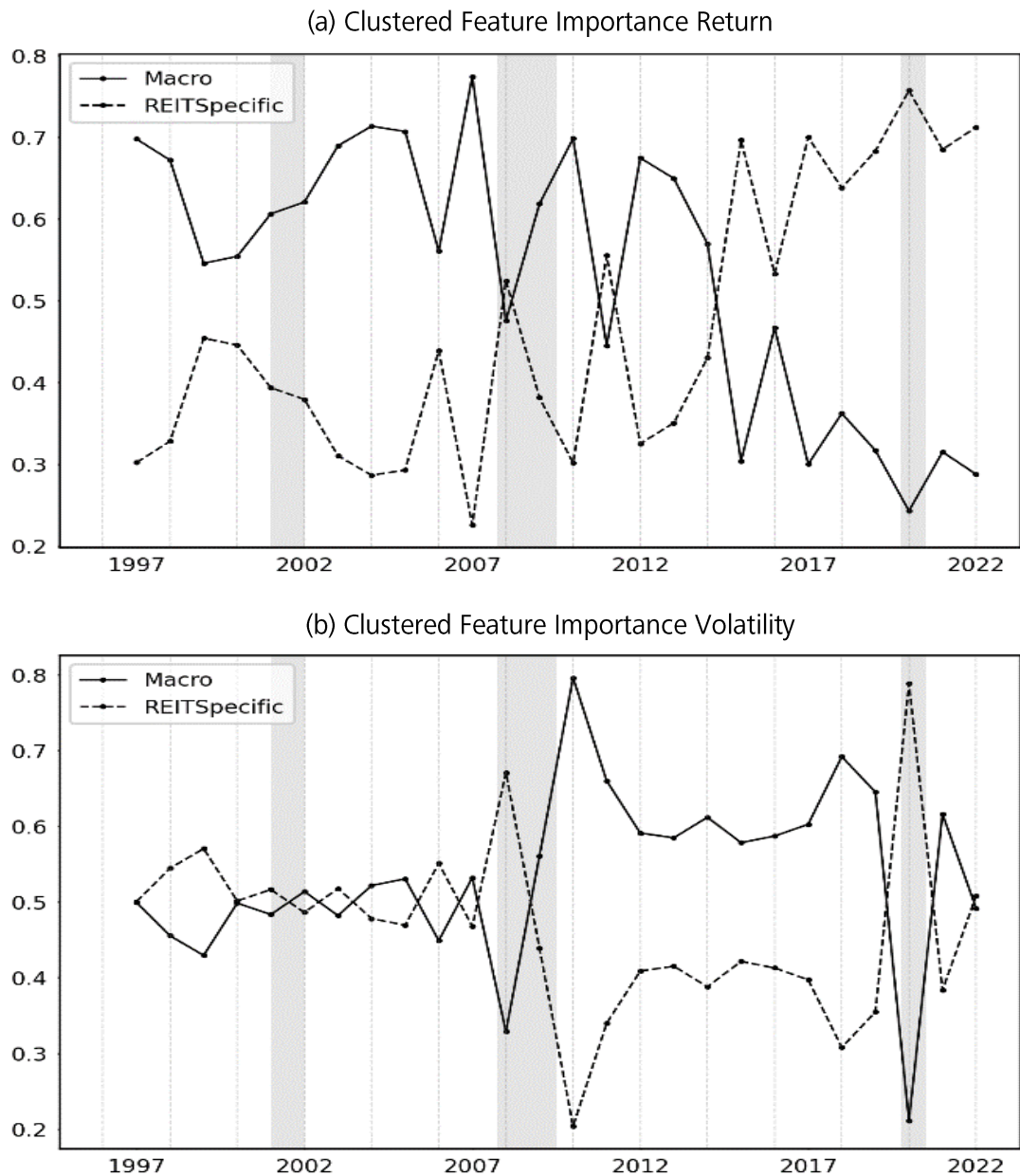
For a broader view of the evolution of determinants, we cluster all variables into a macroeconomic and a REIT-specific group. Figure 2.5 displays the sum of absolute feature importance of the clustered variables over time for (a) return and (b) volatility estimation. The solid line depicts the macroeconomic variables, while the dashed line shows all REIT-specific factors. The sum of feature importance for both clusters each year equals 1, meaning that a more substantial influence of macroeconomic constituents coincides with lower REIT-specific constituents' impact, and vice versa. This direct dependency allows an in-depth analysis of the sensitivity of REITs in different market states. In accordance with

the subperiods results, we see a steady upward trend of company-based characteristics over the years for returns, with 2014 marking the years where REIT variables have been the dominating force in the composition of returns. In the earlier stages of the REIT market (1990s and 2000s), REITs were a relatively new asset class with limited investor familiarity and liquidity. This made returns more sensitive to broad economic trends. The financial crisis led to a reassessment of risk, particularly in real estate markets. Investors increasingly focused on REIT-specific metrics such as asset quality and balance sheet strength, as they became more critical for assessing resilience in the face of systematic shocks. Moreover, advances in financial reporting standards and data accessibility likely increased the importance of firm-level factors. With about 70% of returns driven by company variables in the last four years, our findings suggest that the REIT market is becoming separate from the macro environment.

Looking at the crisis periods shaded in gray, we observe substantial deviations, especially in the GFC period. In the lead-up to the crisis, returns were significantly influenced by economic factors. However, as the crisis unfolded in 2008, there was a sharp decline, followed by a robust recovery in the following years. These anomalies can be attributed to the housing market crash that marked the GFC's start, causing massive interest rate cuts and rising unemployment rates.

The overall cluster variation in the importance for volatility is much lower compared to returns, providing evidence for a more equal interplay between macroeconomy and REIT-related features as well as the predictability of risk. However, especially during economic downturns, volatility prediction is a complex task, as seen by the REITs-specific spikes during the GFC and COVID-19 periods. During crisis periods, REIT-specific accounting variables may have a higher short-term impact on volatility because investors focus more on the immediate financial health and stability of individual REITs, such as liquidity, leverage, and market cap. Economic indicators contributed to REIT volatility excessively after the crisis and during the boom period, as the sensitivity of investors to small movements in economic policies increased.

Figure 2.5 Clustered Feature Importance for Return and Volatility



Following the broader analysis of the two clusters, we examine the time-varying behavior of selected macroeconomic variables with striking evolutions in Figure 2.6. The solid line represents the feature's impact on returns, while the dashed line indicates the feature's impact on volatility. An overview of all remaining macroeconomic variables can be found in Appendix III.

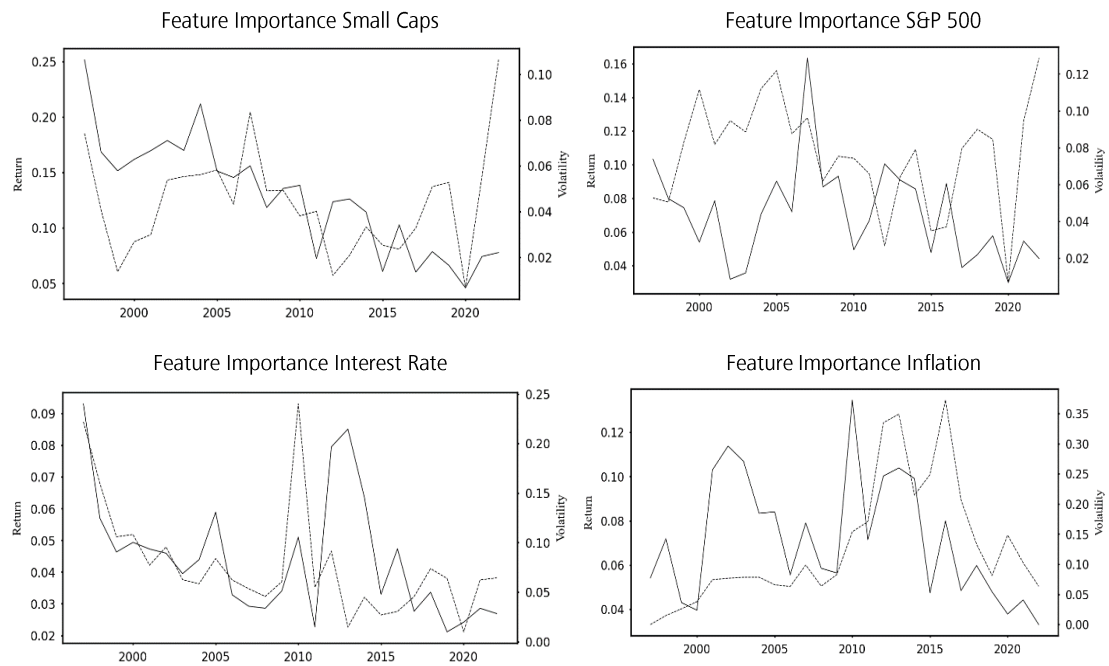
Small-cap stocks' influence on REIT returns and volatility has gradually decreased over time, again confirming the maturing of the REIT market and its increasing independence from broader equity market movements due to their income-generating nature. With their consistent dividend payouts, REITs have been viewed more as bond-like investments rather than high-volatility equity instruments. The growth and diversification of REITs across

property sectors may have further insulated REIT performance from the volatility typically associated with small-cap stocks. This adds a new perspective on the results from Clayton and MacKinnon (2003), who analyzed data until 1998. The S&P 500 index, however, maintained a relatively constant impact on returns and volatility over time, except for a notable spike in 2007. Since some of the REITs are also constituents of the S&P 500 index, the results seem justifiable in terms of correlation. The overall lower importance can be attributed to investors' perception of REITs as a distinct asset class, as confirmed for the years since 2020 by Zhang et al. (2023a). Furthermore, investors allocate capital based on individual factors such as property fundamentals, lease structures, and dividend yields rather than the general stock market movements.

The interest rate's impact on both returns and volatility peaked around the GFC, reflecting the heightened uncertainty and financial market turmoil during this period. The sharp recovery in these impacts can be attributed to the Federal Reserve's monetary policy actions aimed at restoring confidence and supporting economic recovery. Post-crisis, while interest rates continued to influence both metrics, their effects diminished, highlighting stabilization efforts and investor adaptation to the new economic environment. Inflation emerged as the most significant driver globally, particularly during the recovery period from 2010 to 2019, impacting both returns and volatility. Policies like quantitative easing and a low-interest-rate environment contributed to heightened inflation expectations, increasing sensitivity to inflation-related dynamics. Investors are much more sensitive to information about inflation and interest changes when rates are low.

In many cases, the impacts of the key macroeconomic variables on returns and volatility tend to coincide across the observed period, which confirms the general hypothesis that the same drivers influence return and risk. However, the absolute magnitude of impact differs: only small-cap stocks and the S&P 500 index have an overall higher impact on returns. In comparison, all remaining economic features have a higher impact on volatility. This might be because small-cap and large-cap indices show immediate market reactions that affect returns, while the other variables have a long-run impact and first affect the perceived risk. The impact of interest rates on volatility precedes the impact on return by roughly a year during crisis periods, likely because initial changes in interest rates quickly increase market uncertainty and risk perceptions, which then affect REIT returns as the broader economic implications and investor adjustments gradually unfold. Overall, while there is a notable degree of co-movements between return and volatility impacts, the exact timing and magnitude can vary, reflecting the complex dynamics of how these macroeconomic variables interact with the REIT market. The average Spearman correlation coefficient of all economic factors between return and volatility is 0.22.

Figure 2.6 Time-Varying Feature Importance of Macroeconomic Variables



Notes: The solid line denotes the return feature importance and the dashed line the volatility feature importance.

Looking at REIT-specific drivers in Figure 2.7 reveals further valuable insights into the dynamics between risk and return in the REIT market. In accordance with the subperiod investigation, the market cap has become a vital factor for returns, driven by the significant growth and consolidation of the market over the years. Investors may increasingly favor larger REITs due to perceived liquidity, stability, and scale advantages. The impact of the market cap on volatility peaks around periods of widespread market turmoil. During times like this, investors seek safety in larger, more established companies with solid balance sheets and stable cash flows. A REITs total assets reveal very similar patterns, while the peaks may reflect the changing perception of asset quality and management effectiveness.

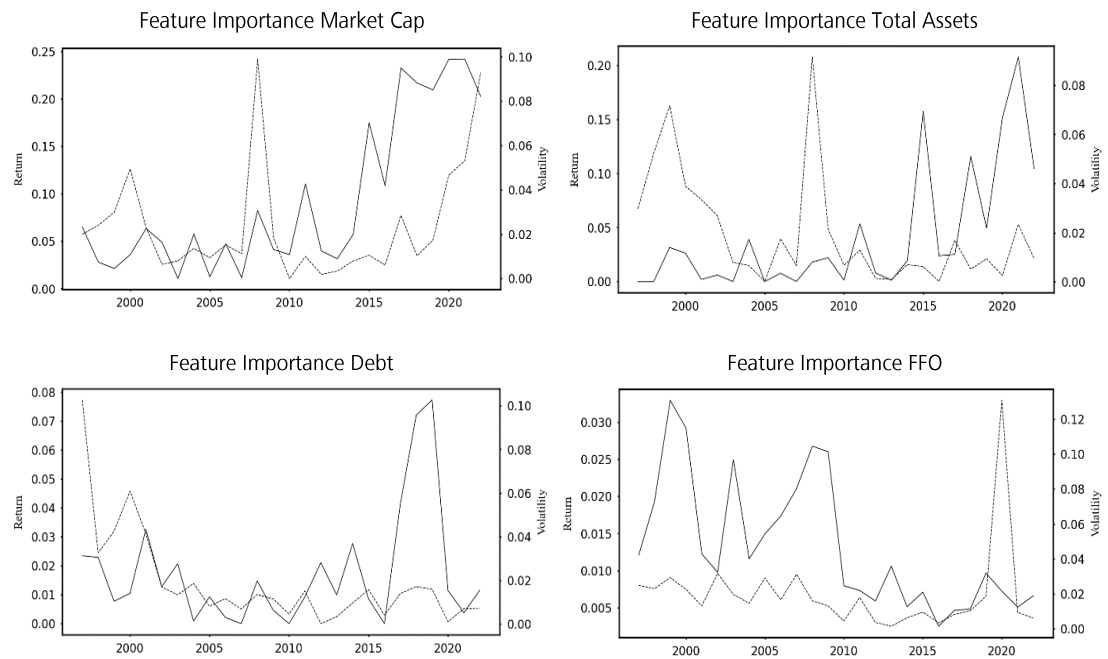
A company's outstanding debt was a significant return driver in 2018 and 2019 but not in any other year. The surge in the impact of outstanding debt on REIT returns during 2018 and 2019 can be attributed to the interest rate hike following an extended period of near-zero interest rates. This led to increased costs of debt, prompting REITs to make strategic decisions such as debt restructuring or deleveraging initiatives. Giacomini et al. (2017) find that highly-levered REITs perform worse on average than low-levered REITs. A REIT's funds from operations usually do not impact the underlying volatility excessively, except for the COVID-19 period. As a result of the disruptions caused by the pandemic, REITs faced challenges such as rent deferrals, lease negotiations, and reduced occupancy rates. As a result, investors reevaluated their investment strategies and risk tolerance levels and placed

greater emphasis on FFO as a critical indicator, amplifying its impact on REIT volatility. An overview of all remaining REIT-specific variables can be found in Appendix IV.

For REIT-specific variables, the deviation of impacts on return and volatility is generally more severe, especially during crisis periods, compared to the economic impacts. In other words, on the REIT financial side, the drivers for return and risk differ more significantly compared to the economic features. The average Spearman coefficient is 0.07, which is 68% lower than the coefficient for the economic features. The significantly lower correlation coefficient may be due to the broad and pervasive impact of economic conditions on the overall market sentiment. Macroeconomic factors such as inflation, interest rates, and market indices directly influence investor behavior, risk perceptions, and capital flows, leading to a stronger and more immediate alignment between returns and volatility. In contrast, REIT-specific financials, such as debt levels, asset valuations, and operational metrics, affect individual REITs differently based on their specific financial health and management strategies. This leads to more heterogeneous impacts on returns and volatility, resulting in a lower overall correlation coefficient for REIT-related financial variables. Hence, investors care less about individual REITs compared to the overall economy when making investment decisions and assessing risk.

Since the time-varying effects of individual drivers on returns and risk are mostly volatile and not trending, predictive models used by practitioners should be refitted regularly to avoid prediction errors. This dynamic nature of factors underscores the need of continuous monitoring and updating of predictive models to capture the latest market conditions and investor behaviors. Additionally, investment strategies should be flexible and adaptive, allowing for adjustments based on the most current data to effectively manage risk and optimize returns.

Figure 2.7 Time-Varying Feature Importance of REIT-Specific Variables



Notes: The solid line denotes the return feature importance and the dashed line the volatility feature importance.

2.6.3 Local Directional Feature Effects

We have seen in the SHAP subperiod analysis that the influence of most variables has changed significantly over time. To gain a deeper understanding of the actual relationship between predictors and returns and volatility, we use ALE plots to visualize feature effects. Figure 2.8 displays the four overall most influential variables (i.e., small-caps, market cap, dividend yield, and VIX) on REIT returns across all subperiods as identified through the feature importance. The values on the y-axis show how much the return prediction deviates from the average prediction when the variable takes on certain values indicated on the x-axis.

The overall impact of the small-caps return is generally positive except for subperiod two, indicating that an increase in the small-caps return is related to an increase in REIT returns. Small-cap stocks are considered to be more closely linked to economic growth compared to large caps, as outlined by Serrano and Hoesli (2010). In the case of market expansion, smaller companies benefit from economic activity, which can also drive demand for real estate, ultimately yielding higher REIT returns. In the second subperiod, a turning point can be observed at around 0% small-cap return. Negative small-cap returns up to 0% correspond to a decline in REIT returns, whereas positive small-cap returns beyond this point are associated with an increase in REIT returns. The economic justification centers around the dynamics of risk sentiment, investor preferences, and macroeconomic

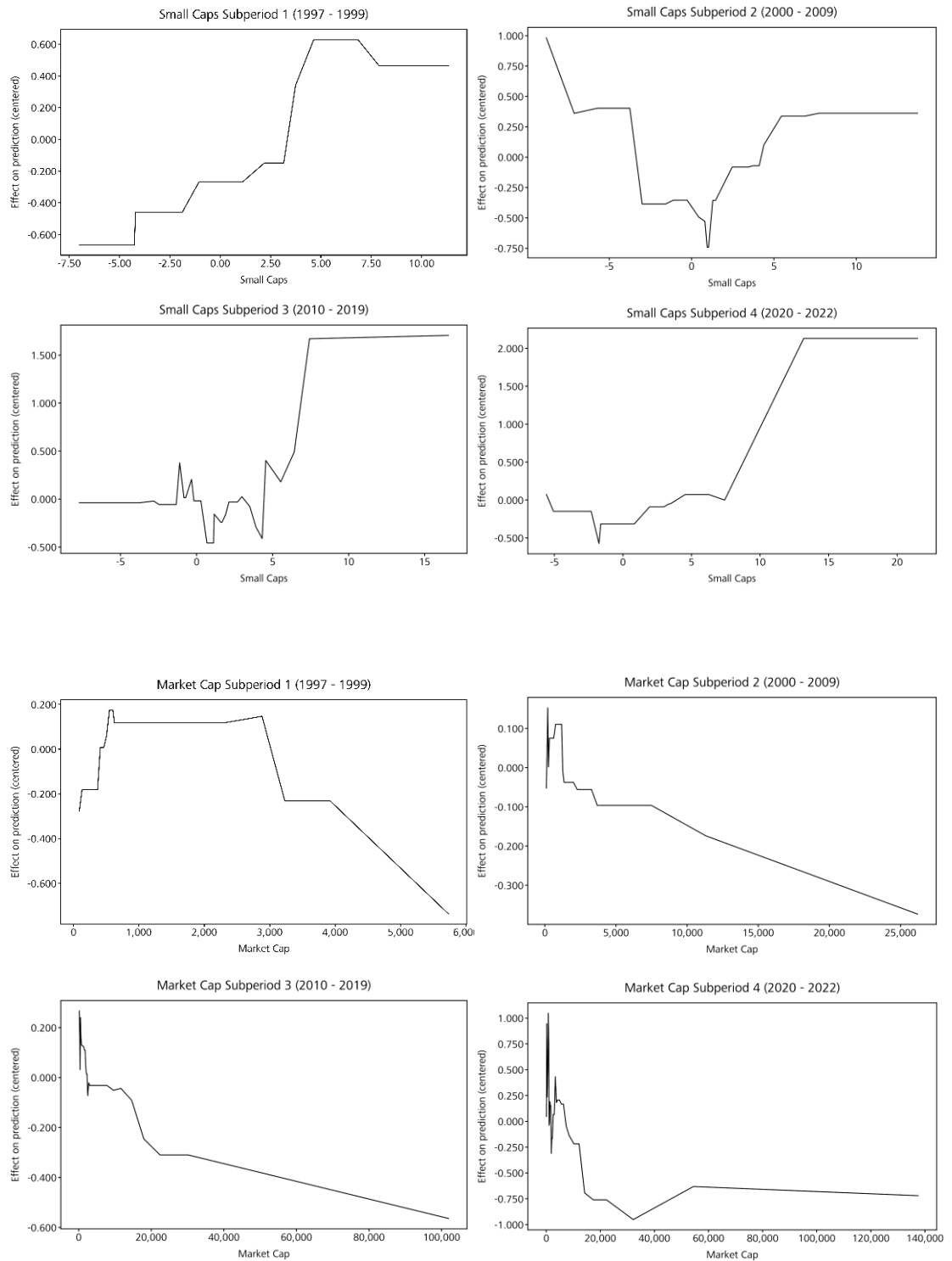
conditions. In periods of negative small-cap returns, REITs likely acted as a safer alternative. However, as economic conditions improved and small-cap returns turned positive, both asset classes benefitted from the broader recovery, leading to a positive relationship. The turning point around 0% reflects this shift in investor behavior and market fundamentals. We therefore add a more nuanced perspective in contrast to previous research on the established co-movement relationship between small-caps and REITs (see, e.g., Liu and Mei, 1992; Clayton and MacKinnon, 2003).

The relationship of market cap and REIT returns reveals a negative impact overall, meaning that larger REITs tend to have lower returns and vice versa. Based on the risk-return tradeoff, smaller REITs exhibit higher volatility than large REITs but also compensate investors with higher returns. Another reason could be that growing REITs face increasing operational inefficiencies, which is supported by Beracha et al. (2019), who find that higher efficiency REITs outperform lower efficiency REITs in terms of risk-adjusted returns. Our finding contrasts with Goebel et al. (2013), who see that large REITs outperform small REITs in a restrictive environment. Moreover, we can detect that for very small REITs, there is a massive variation in the impact on company size on returns especially in subperiods of economic uncertainty (i.e., subperiods 2 and 4).

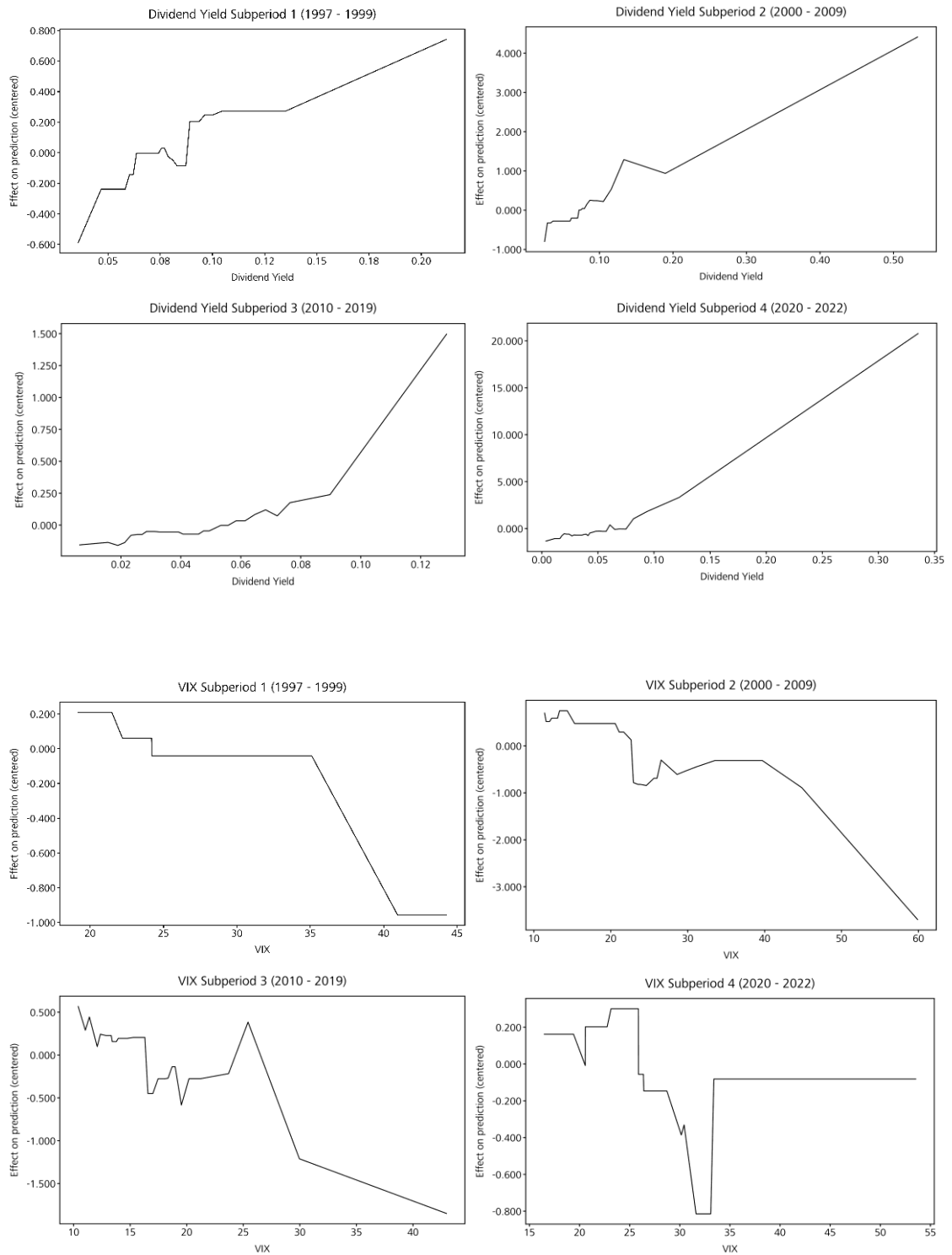
Returns increase significantly when high dividend yields are paid to the investors. Since REITs are required by law to distribute a large amount of their income to shareholders, high dividend-paying REITs indicate strong income generation from their real estate portfolio, driving up prices and returns for investors. Liu and Mei (1992), Brooks and Tsolacos (1999) and Chiang (2015) also show a positive impact of dividend yields on REIT returns. However, inferential assumptions are missing to determine the causal relationship between these two variables.

The ALE plots for the VIX also show a negative slope across all subperiods, meaning that higher market volatility is reflected in lower REIT returns, which generally is in line with Anoruo and Murthy (2017). However, the visualization of the effect creates a more in-depth picture of the actual relationship for different intervals. The shapes of the first three subperiod graphs reveal a rather concave shape, meaning that the negative effect appears softly until a certain level of volatility, from which the effect is intensified significantly. In other words, moderate increases in overall market volatility result in only slight declines in returns. Though, once the index value reaches approximately 25 to 40, further increases in volatility have a significantly greater impact on REIT returns.

Figure 2.8 ALE Plots for Return Drivers over the Subperiods



Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach



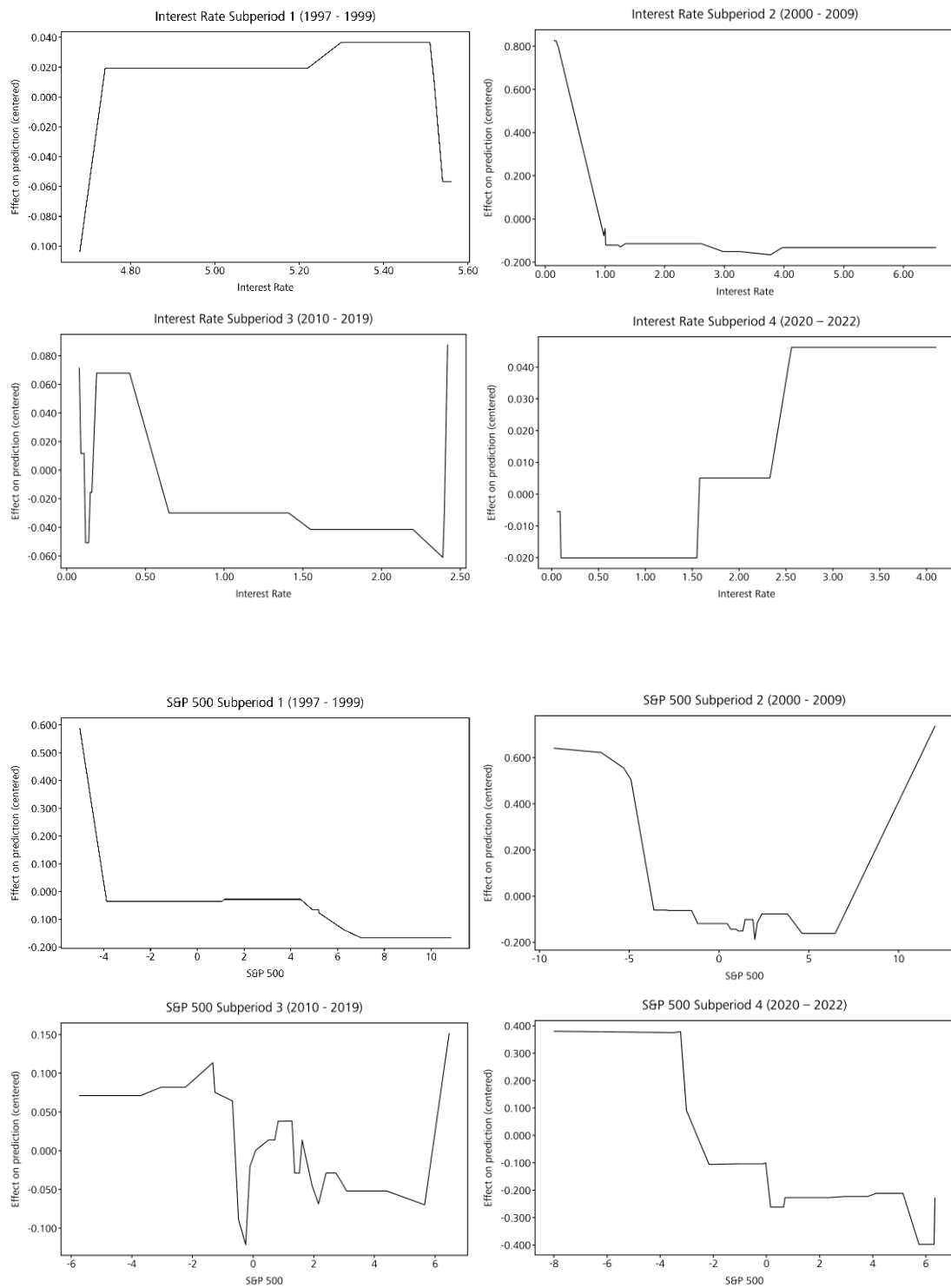
While the overall impact of return drivers is mostly similar across all subperiods, the volatility plots in Figure 2.9 show more significant time-variations. Regarding interest rates, especially the values on the outer bounds seem to impact REIT volatility quite significantly. Examining the lower bound, this aligns with our SHAP findings: when interest rates are low, investors become much more sensitive to any changes. This heightened sensitivity explains why return volatility spikes in both directions.

While the first and last subperiods of the S&P 500 reveal a negative slope, the second and third show some sort of u-shape. If the general stock market is doing well, this also reflects on the REIT market by decreasing volatility. During the 2000 to 2009 period, this relationship shifts at the higher bounds, where the graph spikes again. This could be attributed to stock overvaluation, as seen during the dot-com bubble and the GFC (both periods characterized by overvalued markets) which may signal impending economic distress.

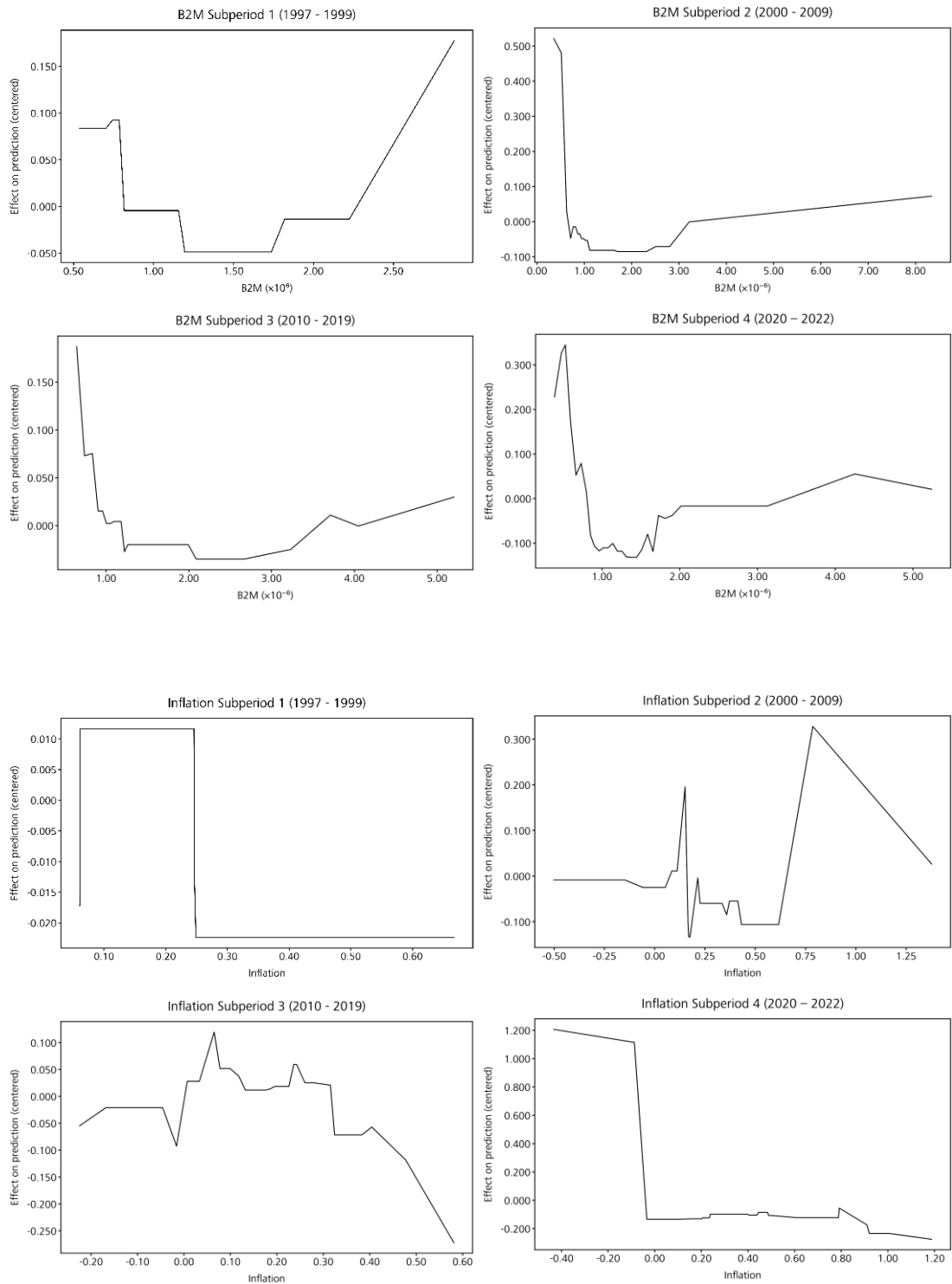
Zhang (2013) shows for the general U.S. stock market that stocks with low book-to-market ratios have significantly more volatility in their return distributions compared to high book-to-market stocks. Our analysis reveals a u-shape across all subperiods, which presents a more nuanced interpretation. REITs with low book-to-market values are growth stocks. At very low book-to-market levels, volatility may be relatively high due to speculative investments, but as the ratio increases, valuations may stabilize, reflecting investor confidence in the firm's fundamentals and growth prospects, leading to lower volatility. High book-to-market values, on the contrary, can indicate potential investor concerns about poor growth prospects, operational inefficiencies, or financial distress, which can amplify return volatility.

Except for subperiod two, higher inflation levels are generally related to lower volatility. REITs are often seen as natural inflation hedges since property values and rental income tend to rise with inflation. This may reduce volatility compared to other asset classes that may suffer in inflationary environments. This contrasts Li (2012), who finds that higher unexpected inflation is related to higher volatility. During the second subperiod, the relationship lacks a clear trend. Speculative housing markets and monetary policy uncertainty amplified investor sensitivity to inflation shocks and created inconsistent effects on return volatility.

Figure 2.9 ALE Plots for Volatility Drivers over the Subperiods



Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach



2.7 Conclusion

Global financial markets have recently been experiencing enormous turmoil amidst unprecedented economic uncertainty and geopolitical tensions. The real estate market has been massively impacted by factors such as fluctuating interest rates and high inflation, which is reflected in the REIT market by the significant decline in the NAREIT index from 2021 to 2024. To navigate the challenges posed by periods with heightened economic uncertainty, it is crucial for investors and policymakers to not only understand the main determinants of REIT returns but also, from a risk management perspective, to concurrently analyze characteristics that influence their associated volatility.

To address this issue, we analyze the time-varying constitution of REIT returns and REIT volatility simultaneously, using a diverse set of REIT-specific financials and overall economic indicators. Our results show that no variables exert the strongest influence on the estimates over the entire period from 1996 to 2022 but that the impact of the variables changes significantly over time in many cases. Therefore, stock prediction models should be refitted regularly to incorporate the latest information on influential drivers in a certain period. The predominant features in explaining REITs vary between market phases, while a few overarching observations can be made. Macroeconomic factors were found to be predominant drivers of REIT returns in the earlier years. At the same time, REIT-specific variables have gained importance over time and make up almost two-thirds of the financial returns for investors nowadays, highlighting that investors place greater emphasis on REIT-specific factors such as property type or operational performance when evaluating investment opportunities. Looking at the volatility, the weights of economic indicators and REIT-specific data are rather equal, and constant compared to returns. However, the macroeconomy was a stronger driver during the 2010s up until the COVID-19 crisis. This effect is due to the increased sensitivity of investor sentiment coupled with investors' focus on the impact of regulatory changes after the GFC.

Crises cause substantial deviations of REIT drivers from otherwise long-run patterns for both return and risk. However, our results reveal that the correlation between risk and return drivers increases during periods of uncertainty. The heightened correlation indicates that economic uncertainty strongly aligns risk and return drivers, underscoring the importance of considering both risk and return in a unified framework, especially during times of economic instability.

By isolating economic and REIT-related financials from each other, our results provide an answer to the question of whether the same drivers can be used to explain return and risk. On the economic side, individual risk and return drivers show stronger co-movements

compared to the REIT-specific features, demonstrating that the same macroeconomic indicators can be used more effectively than REIT indicators to simultaneously explain both returns and risk. This is due to the broader and more uniform impact of macroeconomic conditions on market sentiment and investor behavior, whereas REIT-specific factors influence returns and volatility in a more heterogeneous manner based on individual financial health and management strategies.

Lastly, we reveal substantial non-linearities in the way the features interact with REIT returns and volatility. Traditional OLS models usually do not identify these relationships. We apply two state-of-the-art XAI technique, SHAP and ALE, to conduct a comprehensive analysis of the model's estimates, and thus open the black box. This approach enables a more profound understanding of real estate market dynamics.

Our results increase transparency and help investors, risk managers, and policymakers better understand REIT market dynamics and make more informed decisions. This can lead to a more efficient capital allocation and ultimately contributes to the overall stability and growth of the market. Risk managers can incorporate novel insights about the dynamic relationships between different risk and return factors to help create more robust portfolios. Lastly, policymakers can use our findings to support the real estate sector by implementing effective financial policies based on the market cycle analysis and tracking the impact of existing policies.

2.8 Appendix

2.8.1 Appendix I – Chosen Variables and Related Literature

Table 2.5 Chosen Variables and Related Literature

Variable	Related Literature
Real Estate Variables	
REIT Type	Anderson et al. (2015)
Accounting Variables	
Market Capitalization	Goebel et al. (2013)
Total Assets	Beracha et al. (2019)
Return on Asset	Bond and Xue (2017)
Funds From Operations	Graham and Knight (2000)
Normalized EBITDA	Beracha et al. (2019)
Total Debt Outstanding	Chung et al. (2016)
Cash and short-term Investments	Hung and Glasscock (2010)
Interest Expense	Beracha et al. (2019)
Total Revenue	Beracha et al. (2019)
Capex	Suzuki et al. (2023)
Dividends	Akinsomi et al. (2016)
Dividend Yield	Brooks and Tscolacos (2003)
Retained Earnings	Giacomini et al. (2017)
Operating Income	Bond and Xue (2017)
Book-to-market ratio	Goebel et al. (2013)
Macroeconomic Variables	
S&P 500	Asteriou and Begiazi (2013)
Fed Funds Rate	Chen et al. (2014)
GS10	Liu and Mei (1992)
CPI	Li and Wang (1995)
Sentiment Index	Akinsomi et al. (2016)
VIX	Anoruo and Murthy (2017)
Small Cap Return	Clayton and Mackinnon (2003)

Notes: This table shows related literature that examines the factors used in this paper that were found to drive REIT performance (i.e., returns or volatility). For each variable, only one related paper is presented in this table for simplicity reasons. The variable “Property, Plant & Equipment” have not been investigated so far.

2.8.2 Appendix II – Model Performance Comparison

The evaluation metrics in Table 2.6 show the average annual performance of the XGBoost compared to the LASSO regression. XGBoost consistently outperforms LASSO regression across all samples and metrics for return estimation, demonstrating its capability to detect non-linearities and interactions between predictors and the target variable. The relatively low R^2 values are common for stock return prediction due to the noise in the market, see Gu et al. (2020).

Table 2.6 Evaluation of XGBoost and LASSO

Models	R ²	MSE	MAE
Return			
XGBoost _{in sample}	0.15	11.26	2.26
XGBoost _{out of sample}	0.09	12.07	2.31
LASSO _{in sample}	0.01	13.35	2.39
LASSO _{out of sample}	0.01	13.34	2.39
Volatility			
XGBoost _{in sample}	-1.79	3.16	1.43
XGBoost _{out of sample}	-1.81	3.19	1.44

Figure 2.10 (a) displays the evolution of the in-sample and out-of-sample R^2 over time. Following the GFC, a notable structural break in model performance is evident, particularly with a substantial increase observed in out-of-sample variation post-GFC. Possible reasons for this increase are the post-GFC market stability, which led to more stable relationships between predictors and REIT returns. Moreover, financial policy interventions like the Dodd-Frank Act, the FSOC, and SEC regulations helped stabilize the REIT market, making it easier for models to explain return variations. The remaining performance metrics plots can be found in Figure 2.11 and Figure 2.12.

Figure 2.10 R^2 Evolution over Time



Figure 2.11 MSE Evolution over Time

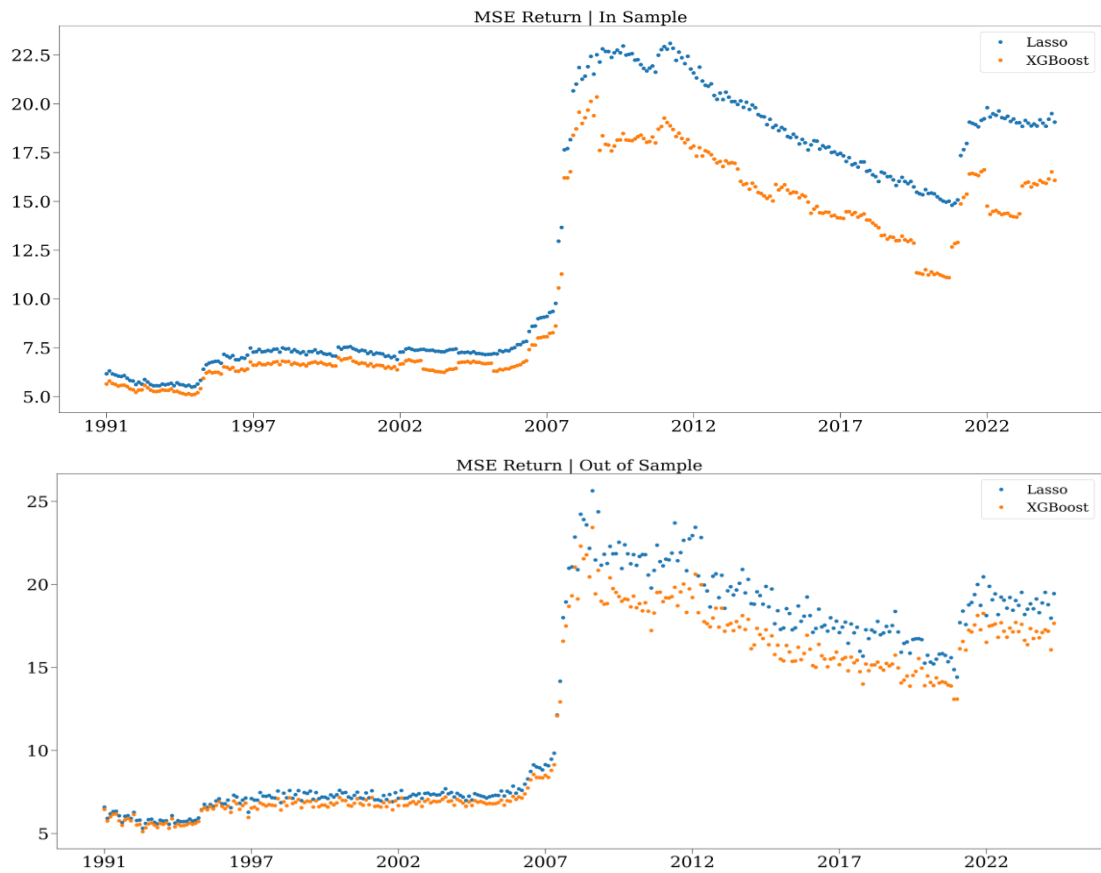
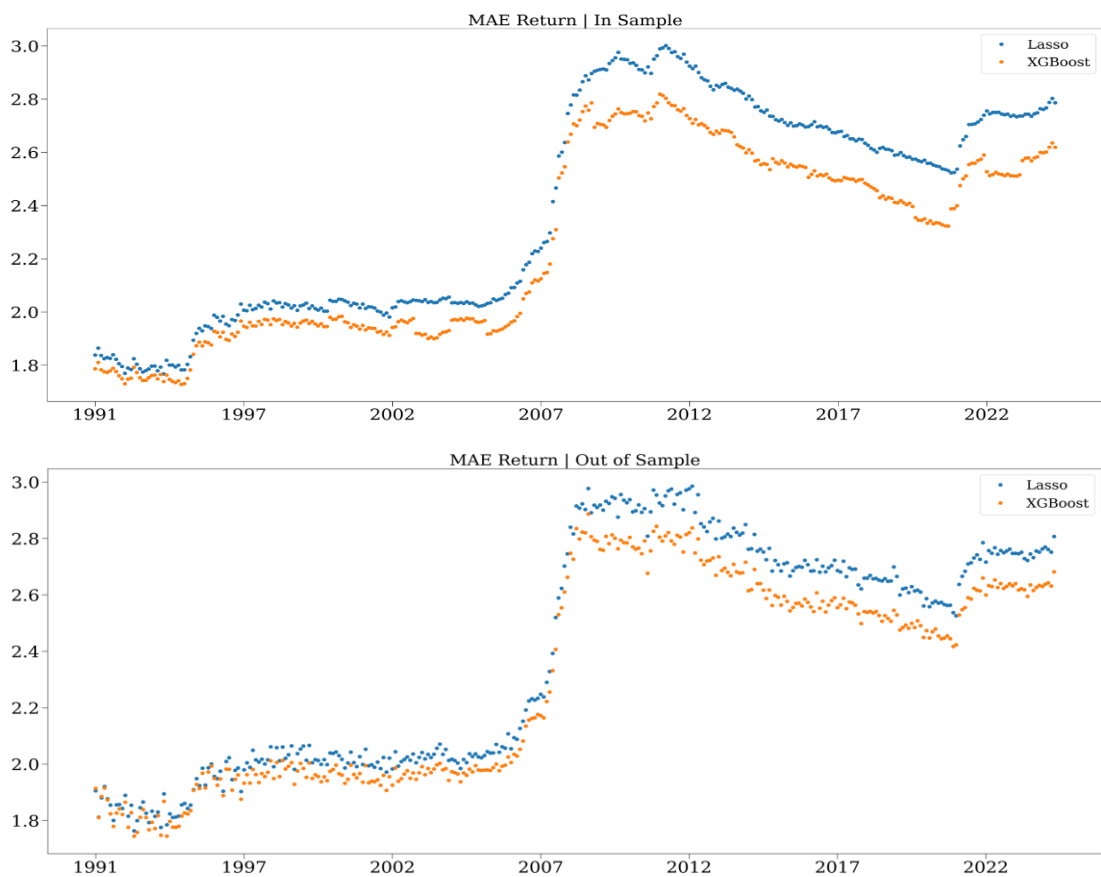
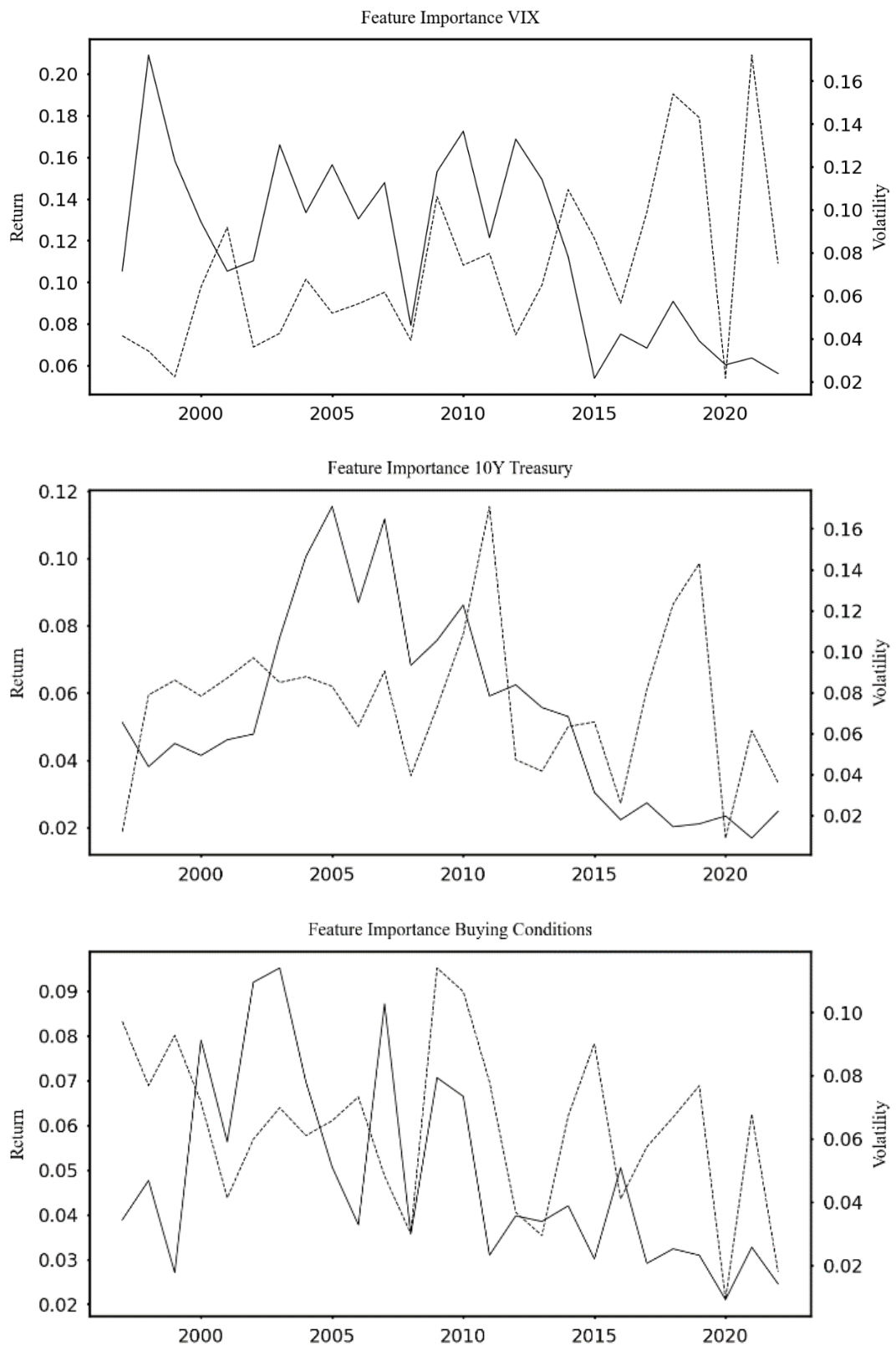


Figure 2.12 MAE Evolution over Time



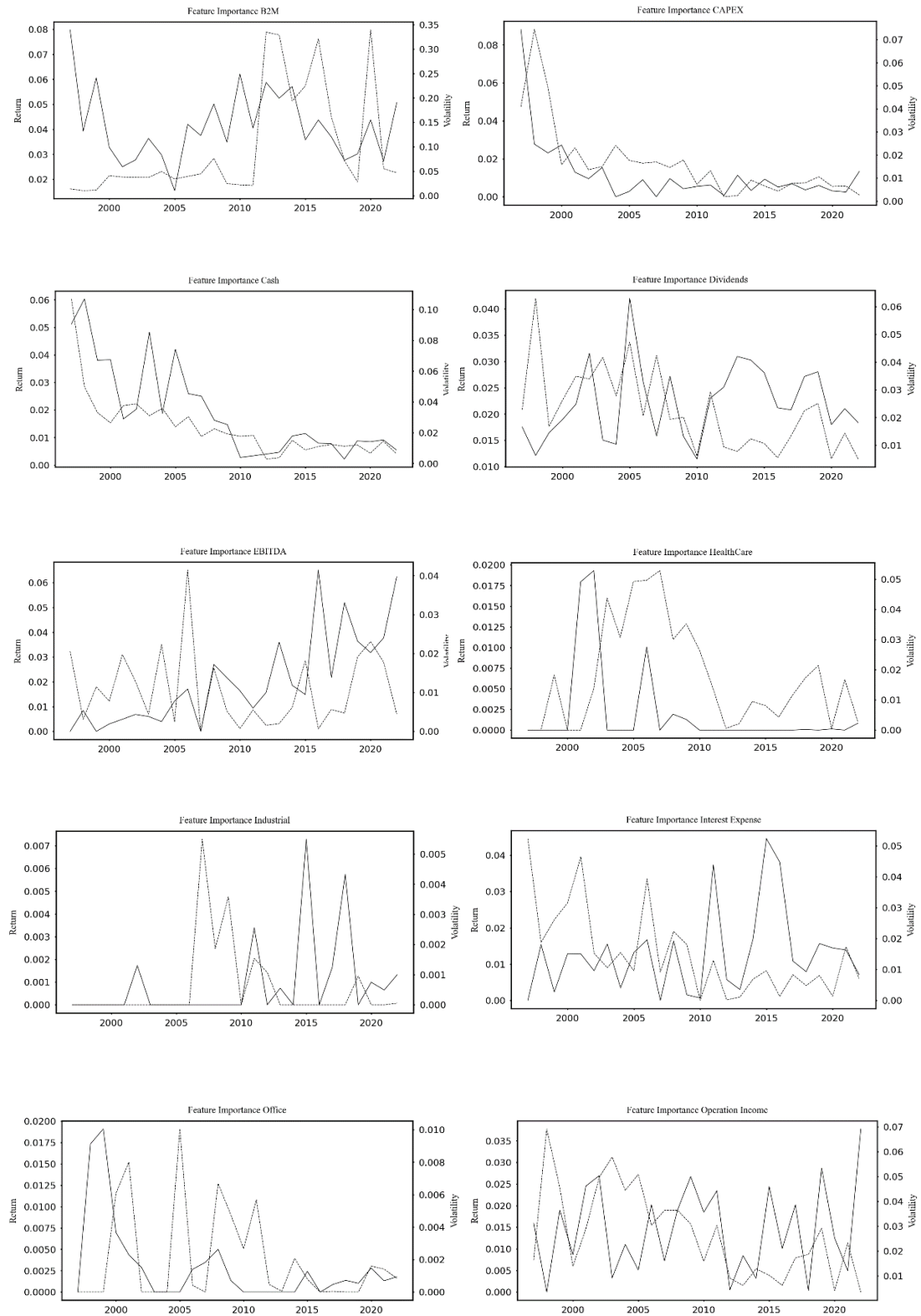
2.8.3 Appendix III – Feature Importance of Macroeconomic Variables

Figure 2.13 Time-Varying Feature Importance of Macroeconomic Variables

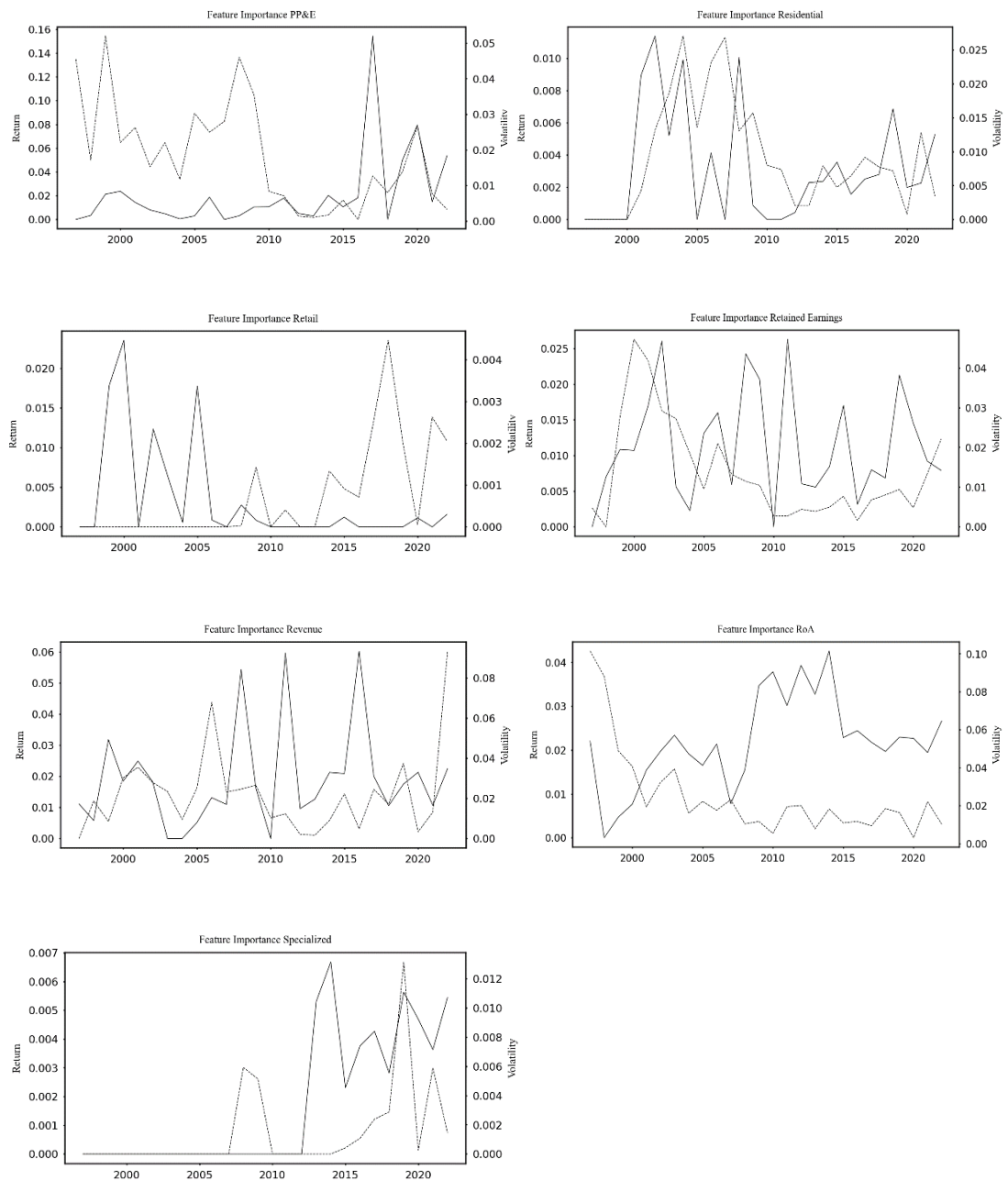


2.8.4 Appendix IV – Feature Importance of REIT-Specific Variables

Figure 2.14 Time-Varying Feature Importance of REIT-Specific Variables



Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach



2.9 References

- Akinsomi, O., Aye, G. C., Babalos, V., Economou, F. and Gupta, R. (2016).** Real estate returns predictability revisited: novel evidence from the US REITs market. *Empirical Economics*, 51(3), 116–1190.
- Anderson, J., Anderson, R., Guirguis, H. S., Proppe, S. and Seiler, M. J. (2021).** Time-Varying Correlations of REITs and Implications for Portfolio Management. *Journal of Real Estate Research*, 43(3), 317–334.
- Anderson, R., Benefield, J. D. and Hurst, M. E. (2015).** Property-type diversification and REIT performance: an analysis of operating performance and abnormal returns. *Journal of Economics and Finance*, 39(1), 48–74.
- Anoruo, E. and Murthy, V. N. R. (2017).** An examination of the REIT return–implied volatility relation: a frequency domain approach. *Journal of Economics and Finance*, 41(3), 581–594.
- Antipov, E. A. and Pokryshevskaya, E. B. (2012).** Mass appraisal of residential apartments: an application of random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, 39(2), 1772–1778.
- Apley, D. W. and Zhu, J. (2020).** Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society*, 82(4), 1059–1086.
- Asteriou, D. and Begiazi, K. (2013).** Modeling of daily REIT returns and volatility. *Journal of Property Investment & Finance*, 31(6), 589–601.
- Babaei, G., Giudici, P. and Raffinetti, E. (2022).** Explainable artificial intelligence for crypto asset allocation. *Finance Research Letters*, 47(1), 102941.
- Below, S. D., Stansell, S. R. and Coffin, M. (2000).** The Determinants of REIT Institutional Ownership: Tests of the CAPM. *Journal of Real Estate Finance and Economics*, 21(3), 263–278.
- Beracha, E., Feng, Z. and Hardin, W. G. (2019).** REIT Operational Efficiency: Performance, Risk, and Return. *Journal of Real Estate Finance and Economics*, 58(3), 408–437.
- Bianchi, D., Buechner, M. and Tamoni, A. (2021).** Bond risk premiums with machine learning. *The Review of Financial Studies*, 34(2), 1046–1089.
- Bonato, M., Cepni, O., Gupta, R. and Pierdzioch, C. (2021).** Forecasting realized volatility of international REITs: The role of realized skewness and realized kurtosis. *Journal of Forecasting*, 41(2), 303–315.
- Bond, S. and Xue, C. (2017).** The Cross Section of Expected Real Estate Returns: Insights from Investment-Based Asset Pricing. *Journal of Real Estate Finance and Economics*, 54(3), 403–428.

- Brooks, C. and Tsolacos, S. (2003).** International evidence on the predictability of returns to securitized real estate assets: econometric models versus neural networks. *Journal of Property Research*, 20(2), 133–155.
- Bryzgalova, S., Pelger, M. and Zhu, J. (2020).** Forest through the trees: Building cross-sections of stock returns. Working Paper. Available at: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3493458.
- Bussmann, N., Giudici, P., Marinelli, D. and Papenbrock, J. (2021).** Explainable Machine Learning in Credit Risk Management. *Computational Economics*, 57(1), 203–216.
- Chatrath, A. and Liang, Y. (1998).** REITs and Inflation: A Long-Run Perspective. *Journal of Real Estate Research*, 16(3), 311–326.
- Chen, J., Chang, T., Ho, C. and Diaz, J. F. (2014).** Grey Relational Analysis and Neural Network Forecasting of REIT returns. *Quantitative Finance*, 14(11), 2033–2044.
- Chen, T. and Guestrin, C. (2016).** XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August, San Francisco.
- Chen, L., Pelger, M. and Zhu, J. (2024).** Deep Learning in Asset Pricing. *Management Science*, 70(2), 714–750.
- Chiang, K. C. H. (2015).** What Drives REIT Prices? The Time-Varying Informational Content of Dividend Yields. *Journal of Real Estate Research*, 37(2), 173–190.
- Chui, A. C. W., Titman, S. and Wei, K. C. J. (2003).** The Cross Section of Expected REIT Returns. *Real Estate Economics*, 31(3), 451–479.
- Chung, R., Fung, S., Shilling, J. D. and Simmons-Mosley, T. X. (2016).** REIT Stock Market Volatility and Expected Returns. *Real Estate Economics*, 44(4), 968–995.
- Clayton, J. and MacKinnon, G. (2003).** The Relative Importance of Stock, Bond and Real Estate Factors in Explaining REIT Returns. *Journal of Real Estate Finance and Economics*, 27(1), 39–60.
- Cotter, J. and Stevenson, S. (2006).** Multivariate Modeling of Daily REIT Volatility. *Journal of Real Estate Finance and Economics*, 32(2), 305–325.
- Cotter, J. and Stevenson, S. (2007).** Uncovering Volatility Dynamics in Daily REIT Returns. *Journal of Real Estate Portfolio Management*, 13(2), 119–128.
- Cotter, J. and Stevenson, S. (2008).** Modeling Long Memory in REITs. *Real Estate Economics*, 36(3), 533–554.
- Daniel, K. and Titman, S. (1997).** Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance*, 52(1), 1–33.
- DeLisle, R. J., Price, S. M. and Sirmans, C. F. (2013).** Pricing of Volatility Risk in REITs. *Journal of Real Estate Research*, 35(2), 223–248.

- Feng, K. and Li, Q. (2016).** Using Stepwise Regression and Support Vector Regression to Comprise REITs' Portfolio. Working Paper. Available at: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7065026>.
- Feng, G., He, J., Polson, N. G. and Xu, J. (2023).** Deep Learning in Characteristics-Sorted Factor Models. Working Paper. Available at: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3243683.
- Franke, M. and van de Minne, A. (2024).** Combining machine learning and econometrics: Application to commercial real estate prices. *Real Estate Economics*, 52(1), 1308–1339.
- Freyberger, J., Neuhierl, A. and Weber, M. (2020).** Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5), 2326–2377.
- Friedman, J. H. (2001).** Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Giacomini, E., Ling, D. C. and Naranjo, A. (2017).** REIT Leverage and Return Performance: Keep Your Eye on the Target. *Real Estate Economics*, 45(4), 930–978.
- Goebel, P. R., Harrisn, D. M., Mercer, J. M. and Whitby, R. J. (2013).** REIT Momentum and Characteristic-Related REIT Returns. *Journal of Real Estate Finance and Economics*, 47(3), 564–581.
- Graham, C. and Knight, J. (2000).** Cash Flow vs. Earnings in the Valuation of Equity REITs. *Journal of Real Estate Portfolio Management*, 6(1), 17–25.
- Gu, S., Kelly, B. and Xiu, D. (2020).** Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273.
- Habbab, F. Z., Kampouridis, M. and Papastylianou, T. (2023).** Improving REITs Time Series Prediction Using ML and Technical Analysis Indicators. 2023 International Joint Conference on Neural Networks (IJCNN), June, Gold Coast.
- Habbab, F. Z. and Kampouridis, M. (2024).** An in-depth investigation of five machine learning algorithms for optimizing mixed-asset portfolios including REITs. *Expert Systems With Applications*, 235(1). 121102.
- Hastie, T., Friedman, J. and Tibshirani, R. (2009).** The elements of statistical learning. Springer New York, New York.
- Hung, S. K. and Glascock, J. L. (2010).** Volatilities and Momentum Returns in Real Estate Investment Trusts. *Journal of Real Estate Finance and Economics*, 41(2), 126–149.
- Kok, N., Koponen, E. and Martínez-Barbosa, C. A. (2017).** Big Data in Real Estate? From Manual Appraisal to Automated Valuation. *Journal of Portfolio Management*, Special Issue, 202–211.

- Kovvuri, V. R. R., Fu, H., Fan, X. and Seisenberger, M. (2023).** Fund performance evaluation with explainable artificial intelligence. *Finance Research Letters*, 58(1), 104419.
- Kraemer, B., Nagl, C., Stang, M. and Schaefers, W. (2023a).** Explainable AI in a Real Estate Context – Exploring the Determinants of Residential Real Estate Values. *Journal of Housing Research*, 32(2), 204–245.
- Kraemer, B., Stang, M., Doskoč, V., Schaefers, W. and Friedrich, T. (2023b).** Automated valuation models: improving model performance by choosing the optimal spatial training level. *Journal of Property Research*, 40(4), 365–390.
- Letdin, M., Sirmans, C. S., Sirmans, S. S. and Zietz, E. N. (2019).** Explaining REIT Returns. *Journal of Real Estate Literature*, 27(1), 1–25.
- Li, L. (2012).** The Determinants of REIT Volatility. Working Paper. Available at: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=00ec02407242f57824d9b74d2a1ccea8c81d3a59>.
- Li, L. and Zhu, B. (2022).** Trading and Volatility in Dual Market: Theory and Evidence from Real Estate. *Journal of Real Estate Research*, 44(2), 151–183.
- Li, Y. (2016).** Time Variation of Expected Returns on REITs: Implications for Market Integration and the Financial Crisis. *Journal of Real Estate Research*, 38(3), 321–358.
- Li, Y. and Wang, K. (1995).** The Predictability of REIT Returns and Market Segmentation. *Journal of Real Estate Research*, 10(4), 471–482.
- Lin, B. and Bai, R. (2022).** Machine learning approaches for explaining determinants of the debt financing in heavy-polluting enterprises. *Finance Research Letters*, 44(1), 102094.
- Liu, C. H. and Mei, J. (1992).** The Predictability of Returns on Equity REITs and Their Co-Movement with Other Assets. *Journal of Real Estate Finance and Economics*, 5(1), 401–408.
- Loo, W. K. (2019).** Predictability of HK-REITs returns using artificial neural network. *Journal of Property Investment & Finance*, 38(4), 291–307.
- Lorenz, F., Willwersch, J., Cajias, M. and Fuerst, F. (2023).** Interpretable machine learning for real estate market analysis. *Real Estate Economics*, 51(1), 1178–1208.
- Lundberg, S. M. and Lee, S. (2017).** A unified Approach to Interpreting Model Predictions. 31st Conference on Neural Information Processing Systems, Long Beach.
- McCue, T. and Kling, J. (1994).** Real Estate Returns and the Macroeconomy: Some Empirical Evidence from Real Estate Investment Trust Data, 1972-1991. *Journal of Real Estate Research*, 9(3), 277–287.
- McCluskey, W. J., McCord, M., Davis, P. T., Haran, M. and McIlhatton, D. (2013).** Prediction accuracy in mass appraisal: a comparison of modern approaches. *Journal of Property Research*, 30(4), 239–265.

- McCracken, M. W. and Ng, S. (2016).** FRED-QD: A Quarterly Database for Macroeconomic Research. Working Paper. Available at: <https://research.stlouisfed.org/econ/mccracken/fred-databases/>.
- Mori, M. and Ziobrowski, A. J. (2011).** Performance of Pairs Trading Strategy in the U.S. REIT Market. *Real Estate Economics*, 39(3), 409–428.
- Nagl, M. (2024).** Intricacy of cryptocurrency returns. *Economics Letters* 239(1), 111746.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.
- Rossi, A. G. (2018).** Predicting Stock Market Returns with Machine Learning. Working Paper. Available at: <https://mendoza.nd.edu/wp-content/uploads/2019/07/2018-AlbertoRossi-Fall-Seminar-Paper-1-Stock-Market-Returns.pdf>.
- Serrano, C. and Hoesli, M. (2007).** Forecasting EREIT Returns. *Journal of Real Estate Portfolio Management*, 13(4), 293–310.
- Serrano, C. and Hoesli, M. (2010).** Are Securitized Real Estate Returns more Predictable than Stock Returns? *Journal of Real Estate Finance and Economics*, 41(1), 170–192.
- Stang, M., Kraemer, B., Nagl, C. and Schaefers, W. (2023).** From human business to machine learning—methods for automating real estate appraisals and their practical implications. *Zeitschrift fuer Immobilienoekonomie*, 9(2), 81–108.
- Suzuki, M., Ong, S. E., Asami, Y. and Shimizu, C. (2023).** Long-Run Renewal of REIT Property Portfolio Through Strategic Divestment. *Journal of Real Estate Finance and Economics*, 66(1), 1–40.
- Tibshirani, R. (1996).** Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1), 267–288.
- Wang, X., Xiao, X. and Xiao, Z. (2016).** S-reits' performance forecast using a small sample model associating support vector machine with vector auto-regression model. *International Journal of Innovative*, 12(1), 15–40.
- Wu, W., Chen, J., Yang, Z. and Tindall, M. L. (2021).** A cross-sectional machine learning approach for hedge fund return prediction and selection. *Management Science*, 67(7), 4577–4601.
- Yilmazer, S. and Kocaman, S. (2020).** A mass appraisal assessment study using machine learning based on multiple regression and random forest. *Land Use Policy*, 99(1), 104889.
- Yoo, S., Im, J. and Wagner, J. E. (2012).** Variable selection for hedonic model using machine learning approaches: a case study in Onondaga County, NY. *Landscape and Urban Planning*, 107(3), 293–306.
- Zhang, X. (2013).** Book-to-market ratio and skewness of stock returns. *The Accounting Review*, 88(6), 2213–2240.

Zhang, W., Li, B. and Roca, E. (2023a). Moments and momentum in the returns of securitized real estate: A cross-country study of risk factors driving real estate investment trusts before and during COVID-19. *Heliyon*, 9(8), e18476.

Zhang, W., Li, B., Liew, A. W., Roca, E. and Sing, T. (2023b). Predicting the returns of the US real estate investment trust market: evidence from the group method of data handling neural network. *Financial Innovation*, 9(1), 1–33.

Zhou, J. (2020a). A comparison of realized measures for daily REIT volatility. *Journal of Property Research*, 37(1), 1–24.

Zhou, J. (2020b). Combining realized measures to forecast REIT volatility. *Journal of European Real Estate Research*, 14(1), 19–39.

Zhou, J. and Kang, Z. (2011). A Comparison of Alternative Forecast Models of REIT Volatility. *Journal of Real Estate Finance and Economics*, 42(3), 275–294.

3 Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt

3.1 Abstract

Purpose: This study examines the impact of location-based risk factors, specifically, local real estate market risk and geographic diversification, on the bond risk premia of U.S. Real Estate Investment Trusts (REITs). It addresses a gap in the literature on how locational factors influence REIT debt pricing.

Design/ methodology/ approach: We quantify local real estate market risk via a REIT-specific local beta, capturing sensitivity of local real estate markets to national real estate shocks. Geographic diversification is measured using the Herfindahl-Hirschman-Index (HHI). Using over 30,000 monthly REIT bond yield spreads from 2010 to 2023, we combine machine learning methods (artificial neural network and accumulated local effect plots) and ordinary least squares regression with polynomial terms.

Findings: We find a local real estate market risk premium for REIT bonds, indicating that investors seek compensation for the additional risk. Furthermore, we uncover a non-linear relationship between geographic diversification and bond risk premia. While moderate diversification lowers risk premia, strong diversification increases them. The optimal point of diversification is found at an HHI of 0.25, where REITs face the lowest public debt cost.

Practical implications: REIT managers can reduce borrowing costs by optimizing geographic diversification and avoiding overexposure to volatile markets. Bond investors can better price credit risk using REITs' local beta and diversification metrics.

Originality/value: To the best of the authors' knowledge, this study is the first to show that REIT bondholders price location risk, expanding prior equity-based findings. It uncovers a non-linear relationship between diversification and the cost of public debt. It further highlights how combining machine learning methods with traditional regression methods can enhance model interpretability and performance.

Keywords: REIT Bond, Local Beta, Real Estate Market Risk, Geographic Diversification, Machine Learning

3.2 Introduction

Location is one of the most important value drivers for real estate. A growing body of literature examines the impact of geographic characteristics on the risk and return profiles of both private and public real estate equity investments (Fisher et al., 2022; Zhu and Lizieri, 2022). However, there is limited research on how location risks influence REIT debt. This is a notable gap, given that U.S. REITs rely heavily on bond financing to support their operations, with bond issuances remaining robust in recent years despite rising interest rates.¹ In addition, bondholders may price risk differently than equity holders because they have a higher claim priority and face primarily downside risk, as their upside potential is limited to fixed coupon payments and principal repayment.² Our study aims to fill this gap by investigating how location-based risk characteristics influence the pricing of REIT bonds.

Our analysis is motivated by the intuition underlying the structural credit risk model of Merton (1974), which implies that a firm's asset volatility increases its probability of default, thereby raising the required risk premium on debt. For REITs, whose assets largely consist of property holdings, local real estate market volatility affects asset value volatility, thereby raising their required risk premium on debt.

On the other hand, geographic diversification may reduce asset volatility by offsetting local shocks across different real estate markets. While prior studies modeled the effect of geographic diversification linearly, we argue (drawing on modern portfolio theory) that the benefits of diversification diminish as exposure becomes more widespread. Accordingly, we assume a non-linear relationship between geographic diversification and debt risk premia.

Building on this theoretical premise, we examine the relationship between these two location-based characteristics (local real estate market risk and geographic diversification) with REIT bond risk premia. Based on Zhu and Lizieri (2022), we construct a local beta factor to measure an individual REIT's sensitivity to local real estate market risk. This beta reflects the sensitivity of each local real estate market's total return in which a REIT holds assets to shocks in the aggregated national real estate market total return, thereby capturing the REIT's exposure to higher-risk markets. Our local beta is calculated at the granular level of U.S. Core-Based Statistical Areas (CBSA) to account for the heterogeneity

¹ In 2020, a record-setting year with historically low interest rates, REITs issued \$81.55 billion in bonds. However, with the rise in interest rates, this figure stabilised to \$48.1 billion by 2024, slightly above the ten-year average of approximately \$40 billion prior to 2020.

² Indeed, recent studies on empirical asset pricing (i.e., Gospodinov and Robotti, 2021) suggest that robust evidence for common factors in pricing across equities and bonds remains elusive.

of real estate markets. Additionally, to capture geographic diversification of a REIT's property portfolio, we calculate the Herfindahl-Hirschman-Index (HHI).

We employ a dataset of 30,186 monthly bond yield spreads from 2010 to 2023, along with 27 explanatory variables. The data includes REIT bond information, property location data, REIT-specific accounting measures, equity market data and macroeconomic indicators. In our empirical analysis, we combine machine learning and traditional econometric methods, leveraging the former's ability to capture complex, non-linear patterns while retaining the interpretability offered by the latter, as demonstrated in prior studies (see, e.g., Kraemer et al., 2023b; Lorenz et al., 2023). Specifically, we employ an artificial neural network (ANN) model to predict REIT bond yield spreads. The data patterns are then visualized using the eXplainable Artificial Intelligence (XAI) method, Accumulated Local Effect (ALE) plots. These visualizations guide the selection of model specifications for the final step, where we apply a traditional ordinary least squares (OLS) regression with polynomial transformations based on the inflections identified in the ALE plots, to provide clear measures of statistical significance.

The results show a local real estate market risk premium for REIT bonds. Investors demand higher compensation for holding bonds from REITs with property portfolios in locations that are more sensitive to real estate market shocks. This effect is consistently observed across all REITs, regardless of whether they operate in more volatile or more stable markets. Consistent with our hypothesis, the effect is less pronounced than the findings for REIT equities by Zhu and Lizieri (2022), likely due to the differing priority of claims, as bondholders have higher priority than equity holders and are therefore less sensitive to geographic risks.

Additionally, the results reveal a more nuanced, non-linear relationship between geographic diversification and REIT bond risk premia than previously reported in the literature. For REITs with more concentrated portfolios, increasing geographic diversification leads to a decrease in bond risk premia, reflecting that the benefits of diversification, such as reduced cash flow volatility and lower exposure to local economic shocks, outweigh the potential increased costs associated with managing a dispersed property portfolio. However, for REITs that are already well-diversified, further diversification results in higher bond risk premia, suggesting that the costs of additional diversification exceed its benefits. This pattern aligns with modern portfolio theory, which posits that diversification initially reduces risk, but its marginal benefits diminish as portfolio diversification increases (Markowitz, 1952). The turning point occurs at an HHI value of 0.25, where we identify a structural break in the data that marks a shift in how investors

price the risks of further diversification. This point can be seen as the optimal portfolio diversification for minimizing the cost of public debt from a REIT perspective. When interacting local beta with HHI, we find that for REITs operating in more volatile markets, the effect of geographic concentration on bond risk premia is amplified, resulting in a disproportionately larger increase in risk premia as concentration rises.

This study makes several contributions to the literature on REIT bond pricing and geographic diversification. Firstly, we add to Zhu and Lizieri (2022) as well as Fisher et al. (2022) by showing that local real estate market risk affects REIT public debt, which constitutes a major funding source for REITs. Secondly, we refine the notion of the previously documented geographic diversification discount, contrasting with prior works that model geographic diversification effects linearly (see, e.g., Feng et al., 2021; Hartzell et al., 2014; Ibrahim and Falkenbach, 2024; Ro and Ziobrowski, 2011). Thirdly, we show that the cost of concentration rises non-linearly in high-risk markets, amplifying bond risk premia beyond what either factor would suggest in isolation. Taken together, our findings enhance the understanding of how REIT bondholders price risk and underscore the importance of considering both real estate market conditions and portfolio structure when evaluating debt costs. Lastly, we demonstrate how combining machine learning and traditional models can enhance economic interpretations and model accuracy.

Overall, this study offers practical insights for REIT managers, investors, and bond portfolio managers. REIT managers can use these findings to optimize their real estate market allocations and reduce their cost of public debt. For bond portfolio managers and investors, our analysis of local beta provides a valuable tool for assessing REITs' sensitivity to local market shocks when making investment and hedging decisions.

The remainder of this paper is organized as follows. The next section provides an overview of the related literature and derives the research hypotheses. The "Data and Variable Construction" section outlines the dataset, data processing steps and the construction of the key variables. The "Methodology" section details the applied methods, followed by the empirical findings presented in the "Results" section. Finally, the paper concludes with a summary of insights and implications.

3.3 Literature Review and Hypotheses Development

This study touches on two strands of literature: (1) the impact of location-specific risks on firm performance and asset pricing and (2) the role of diversification, particularly geographic diversification, in shaping REIT value, performance, and debt costs

3.3.1 Location and Firm Risk

A growing body of finance literature emphasizes the role of firm location in determining risk exposure and financial outcomes. Several studies develop local beta factors to capture sensitivity to geographic-specific shocks. Tuzel and Zhang (2017), for example, measure the sensitivity of local industrial production to national GDP shocks, showing that high-beta metropolitan areas experience greater cyclicalities in both wages and real estate prices. Fisher et al. (2022) apply a similar approach in the REIT context, linking location density and systematic stock risk exposure to rental growth and cap rates.

Zhu and Lizieri (2022) introduce a real estate-specific local beta, measuring the sensitivity of local property market returns to national real estate market returns. Their findings indicate that REITs with high local beta, particularly those with geographically concentrated portfolios, exhibit higher equity returns, reflecting compensation for greater location-based risk.

3.3.2 Diversification and REIT Equity Performance and Debt Costs

A second body of literature analyzes the effect of property-type and geographic diversification on REIT equity performance, valuations, and debt costs. While Ro and Ziobrowski (2011) find no significant difference between specialized and diversified REITs in terms of equity excess returns, Benefield et al. (2009) document outperformance for diversified REITs during favorable market conditions. In contrast, European real estate companies, Haran et al. (2020) show that property-type diversification is associated with lower equity returns, underscoring regional differences in outcomes. Taken together, the mixed findings highlight the challenges of isolating diversification effects and their influence on equity performance.

In terms of valuations, several U.S. focused studies report a “diversification discount” for geographically diversified REITs (Feng et al., 2021; Hartzell et al., 2014; Huerta and Mothorpe, 2024). However, this effect is moderated by institutional ownership and transparency. Huerta and Mothorpe (2024) further argue that moderate geographic clustering can enhance operational efficiency.

Evidence on debt costs is more limited but similarly nuanced. Capozza and Seguin (1999) find that REITs with diversified property types face higher bank borrowing costs, potentially due to increased information asymmetries. This is echoed by Cheok et al. (2011) for the Asian property markets, by showing that regionally diversified REITs face higher interest expenses. Demirci et al. (2020) refine this notion by showing that property-type diversification is associated with higher spreads on bank loans and mortgages, whereas

geographic diversification leads to lower spreads. Their findings suggest that, overall, diversification can reduce the cost of debt by enhancing creditworthiness through risk dispersion. On the public debt side, Ibrahim and Falkenbach (2024) provide evidence that geographic diversification is associated with higher REIT bond spreads, though they do not consider non-linearities or interactions with location-specific risk. However, from a modern portfolio theory perspective, the relationship between diversification and risk is inherently non-linear (Markowitz, 1952). Early diversification reduces idiosyncratic risk, while the benefits taper off at higher levels due to increasing correlations among assets. In the context of our study, higher correlations among regions and rising operational complexity may induce the diminishing effect of geographic diversification.

Taken together, these strands reveal that diversification effects differ depending on type (property vs. geography), outcome (equity returns, firm valuations or private vs. public debt costs) and geography (U.S., Europe and Asia). Yet, the role of geographic diversification in REIT bond pricing, particularly its potential non-linearities and interaction with local market risk, remains underexplored. Our study addresses this gap by adopting a granular, asset-level approach and a machine learning framework to detect non-linear patterns.

3.3.3 Hypotheses Development

Our research is conceptually grounded in the structural credit risk model of Merton (1974), which posits that a firm defaults when the value of its assets falls below the face value of its liabilities at maturity. A key implication of the model is that asset volatility increases default probability, as it raises the likelihood that asset values will breach the default threshold. In the case of REITs, whose asset base largely consists of commercial real estate, this volatility is closely tied to fluctuations in the value of their property portfolios. We capture this through a local beta factor, which measures a REIT's exposure to volatility in local real estate markets and thus serves as a proxy for location-based asset risk. A higher local beta implies greater exposure to volatile markets, increasing asset value uncertainty and, by extension, the likelihood of default. Accordingly, bondholders should demand a higher risk premium to compensate for this increased credit risk. At the same time, such volatility may offer upside potential for equity holders through higher returns, highlighting the asymmetry between debt and equity financing. This reasoning supports our hypothesis that REITs with greater exposure to local real estate market risk will face higher bond yield spreads. Therefore, we formulate the following research question:

- I. Is there a local real estate market risk premium in REIT bonds, i.e., do investors demand higher compensation for REITs exposed to more volatile property markets?*

In addition, we argue that geographic diversification of a REIT's property portfolio may lower asset volatility and default risk by mitigating location-specific exposures. However, in line with modern portfolio theory, we expect this effect to be non-linear, with diminishing or even adverse impacts at higher diversification levels. Based on this, we state the second research question:

II. Does the relationship between geographic diversification and REIT bond risk premia exhibit non-linearities?

Finally, if both local market risk and portfolio concentration affect bond risk premia, their interaction could be particularly relevant. For example, REITs operating in volatile markets may benefit more from diversification, while concentration in such markets could amplify default risk. This motivates our third research question:

III. What is the interaction effect between local real estate market risk and geographic diversification on REIT bond risk premia?

3.4 Data and Variable Construction

The sample used in this study consists of 30,186 monthly bond yield spreads of 60 U.S. equity REITs that are constituents of the FTSE NAREIT Index from January 2010 to December 2023.³ The observation period starts in 2010 due to very limited REIT bond transaction data prior to this year, as noted by Freybote (2016). Therefore, the study excludes 2002 to 2009 and focuses on the post-crisis years. The dataset is based on different data types, namely (1) REIT bond data, (2) REIT property data, (3) REIT-specific accounting/market data and (4) macroeconomic data.

3.4.1 REIT Bond Risk Premium

Data on REIT bond yields come from the Trade Reporting and Compliance Engine (TRACE) database by FINRA, which has required broker-dealers to self-report their over-the-counter (OTC) bond transactions since 2002.⁴ We calculate monthly yield spreads by averaging the volume-weighted yields of all transactions in a month of a REIT bond and then subtracting the yield on U.S. Treasuries with corresponding maturities. U.S. Treasury data come from the Liu and Wu (2021) dataset. Since TRACE does not offer bond characteristics such as

³ To prevent survivorship bias, our analysis includes all REITs that issued bonds during the study period, regardless of whether they were present for the entire duration of the sample.

⁴ As bonds are trades over-the-counter (OTC), data collection on bond transactions is limited compared to equity trading data. TRACE is the most relevant database for researching bond transactions in the U.S.. However, its nature of self-reporting can lead to erroneous entries. By winsorising our data, we try to ensure that such errors do not distort the data.

the bond maturity date, call features, or coupon rate, this data was manually collected from the FINRA website, which is based on the TRACE database. Bond transactions with missing data are excluded. Also, bonds with maturities under one month and above 360 months are excluded since no Treasury matching was feasible. Following Bessembinder et al. (2009), TRACE data entries marked as canceled are removed from the dataset, while corrected entries are kept. Furthermore, bond and firm liquidity metrics are included in the sample.

3.4.2 Local Real Estate Market Risk

Our primary research objective is to document the relationship between the systematic risk of local property markets and the risk premium of REIT bonds. Therefore, we use NCREIF NPI total returns of 136 CBSA divisions as well as REIT property portfolio data from S&P Capital IQ to construct our local market risk factor, called local beta. Firstly, we calculate β_c , which denotes the sensitivity of local commercial real estate prices in each CBSA to systematic real estate shocks and is calculated as follows:

$$r_{c,t}^{NPI} - r_{f,t} = \alpha_c + \beta_c(MKT_t^{NPI} - r_{f,t}) + \varepsilon_{c,t}, \quad (3.1)$$

where $r_{c,t}^{NPI}$ is the NCREIF real estate total return in CBSA c in month t and $r_{f,t}$ the risk-free rate, measured by the yield on the 1-month Treasury bill. MKT_t^{NPI} is the mean NPI across all CBSAs in month t and represents the national real estate market total return. We compute the monthly total return of each CBSA and across all CBSAs, in excess of the risk-free rate.⁵ We then regress the CBSA-specific excess return on the overall market excess return to estimate β_c , which captures the sensitivity of the monthly excess return on a given CBSA portfolio to fluctuations in the overall market's excess return. Even though the NCREIF return data is available for different property types, we opt to use the aggregate total return to circumvent issues with sparse data, which could potentially lead to unreliable results.

Next, we calculate local beta for REIT i at time t , representing the mean systematic risk across all CBSAs where the REITs' properties are located, as follows:

$$\beta_{i,t}^{REIT} = \sum_{c=1}^C w_{i,c,t} \times \beta_c, \quad (3.2)$$

where β_c is the CBSA beta, and $w_{i,c,t}$ is the weight assigned to each CBSA, measured as the exposure of REIT i in CBSA c in month t . We calculate this exposure as the number of

⁵ While the NCREIF total return index is only published on a quarterly basis, the risk-free rate is at monthly frequency and hence we can derive a varying monthly difference between these two.

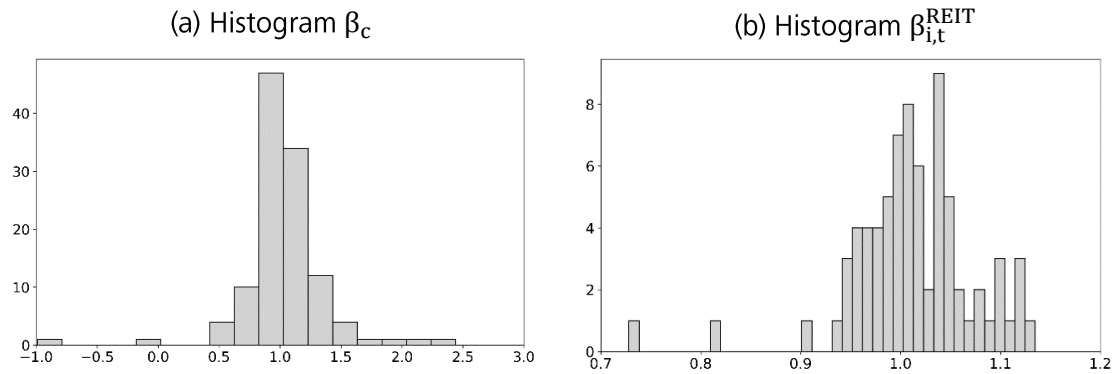
properties held by REIT i in CBSA c in month t divided by the number of total properties held by REIT i at time t across all C CBSAs.

Although β_c is constant over time, $\beta_{i,t}^{REIT}$ adjusts whenever there are changes in a REIT's property portfolio. The variable $\beta_{i,t}^{REIT}$ represents a REIT's exposure to the cyclical behavior of the local real estate markets across the various CBSAs where it holds commercial property investments. For robustness reasons, we additionally calculate a rolling β_c based on a 12-month window. The construction and OLS results can be found in Appendix I. Zhu and Lizieri (2022) prove that the local real estate market risk, as defined in this study, is non-diversifiable. Hence, it should be priced in the bond yield spread even for investors with diversified REIT bond portfolios.

The distribution of β_c and $\beta_{i,t}^{REIT}$ are displayed in Figure 3.1. Table 3.1 reports the five highest and lowest β_c 's as well as the five highest and lowest $\beta_{i,t}^{REIT}$'s. Looking at the histograms, we can observe that most β_c 's lie between 0.5 and 1.5, with only a few observations on the outer bounds. On the contrary, the relatively narrow range of REIT-specific betas, $\beta_{i,t}^{REIT}$, suggests that REITs typically maintain geographically diversified portfolios, which helps smooth out exposure to high- or low-beta CBSAs and manage overall risk more effectively. Peoria, IL and Albany-Schenectady-Troy, NY, are the only CBSAs with a negative beta, indicating that the commercial property markets in these areas move inversely to the national market. As a result, they can provide a hedge against national market trends. On the other side of the spectrum, California-Lexington Park, MD, Wilmington, NC and Cleveland-Elyria-Mentor, OH, have betas exceeding two. This suggests that these areas are highly sensitive to aggregate real estate shocks. For example, the β_c of Wilmington can be interpreted as follows: if the overall market total return increases by 1 percentage point, the return in the specific CBSA increases by 2.362 percentage points. Hence, the real estate market in Wilmington is overly reactive to changes in the aggregate real estate market.

On the $\beta_{i,t}^{REIT}$ side, the three REITs with the lowest beta are Diversified REITs. For example, the $\beta_{i,t}^{REIT}$ of Alexander & Baldwin, Inc. can be interpreted as follows: if the overall market total return increases by 1 percentage point, the property-weighted total return of all CBSAs where this REIT holds assets increases by 0.728 percentage points. Therefore, this REIT is less sensitive to aggregate real estate market shocks.

Figure 3.1 Distribution of Betas



Notes: The figure plots the distribution of the CBSA beta β_c and the average REIT-specific local beta $\beta_{i,t}^{REIT}$, which is based on the REITs property portfolio. The mean CBSA beta is 1.045 and the mean REIT-specific local beta is 1.012. The number of observations for β_c corresponds to the total number of CBSAs and the number of observations for $\beta_{i,t}^{REIT}$ corresponds to the total number of REITs in the study.

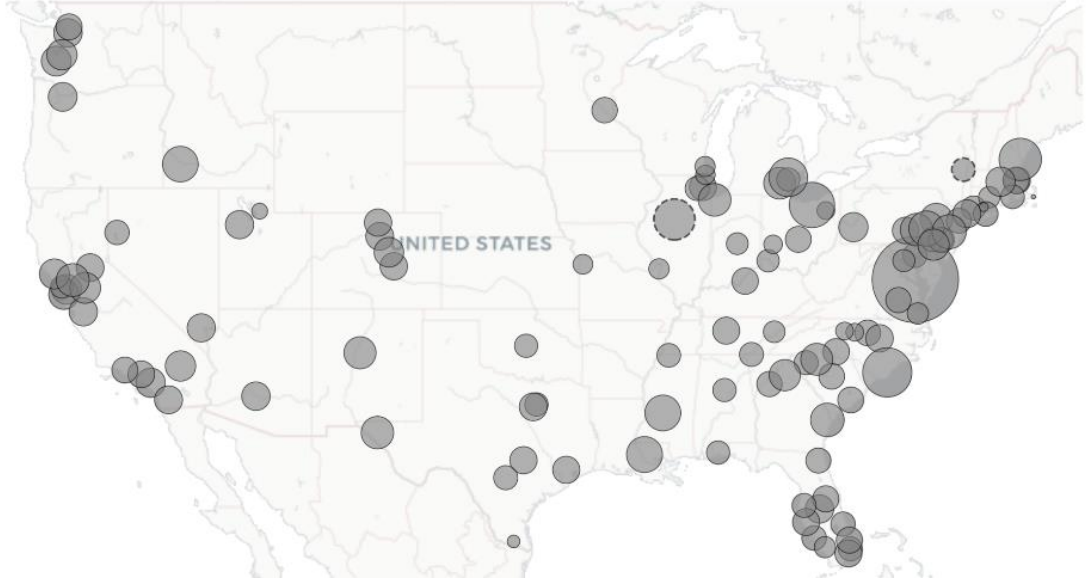
Table 3.1 Highest and Lowest Beta CBSAs and REITs

CBSA Name	β_c
Lowest β_c CBSAs	
Peoria, IL	-1.796
Albany-Schenectady-Troy, NY	-0.899
Palm Bay-Melbourne-Titusville, FL	0.015
Barnstable Town, MA	0.490
New Haven-Milford, CT	0.539
Highest β_c CBSAs	
California-Lexington Park, MD	6.276
Wilmington, NC	2.362
Cleveland-Elyria-Mentor, OH	2.077
Rockingham County-Strafford County, NH	1.873
Warren-Troy-Farmington Hills, MI	1.641
REIT Name	$\beta_{i,t}^{REIT}$
Lowest $\beta_{i,t}^{REIT}$ REITs	
Alexander & Baldwin, Inc.	0.728
Essential Properties Realty Trust, Inc.	0.811
Broadstone Net Lease, Inc.	0.910
Innovative Industrial Properties, Inc.	0.936
Office Properties Income Trust	0.944
Highest $\beta_{i,t}^{REIT}$ REITs	
Sun Communities, Inc.	1.135
Sotherly Hotels Inc.	1.121
Kimco Realty Corporation	1.121
Retail Opportunity Investments Corp.	1.116
Rayonier Inc.	1.110

Figure 3.2 illustrates the β_c distribution across the U.S., with each circle denoting one CBSA. The size of each circle is proportional to the absolute value of the respective β_c (i.e., the higher β_c , the larger the circle on the map), while the two negative values have dashed frames. Coastal CBSAs, especially those on the East Coast, are more sensitive to aggregate real estate shocks than inland CBSAs. The average beta of all coastal CBSAs is 1.12, while

the average beta for all inland CBSAs is 0.97 (coastal is defined as <100 km from the coast). Zhu and Lizieri (2022) show that there is a significant positive relationship between local beta and measures of land regulatory strictness, while the most restricted markets are located along the northeast coast (from Boston through Washington D.C.) and the west coast (Seattle, San Francisco and Los Angeles).

Figure 3.2 Geographic Distribution of CBSA Beta



Notes: The figure plots the geographic distribution of CBSA betas. The size of the circles proportionally coincides with the absolute value of the betas. Negative betas have a dashed frame. Obviously, the circles only cover a small area of the U.S. However, these are the most populated areas and, therefore, contain a large share of the commercial real estate. Only those CBSAs for which a NCREIF NPI total return value exists for at least two years are included in the analysis.

3.4.3 Geographic Diversification

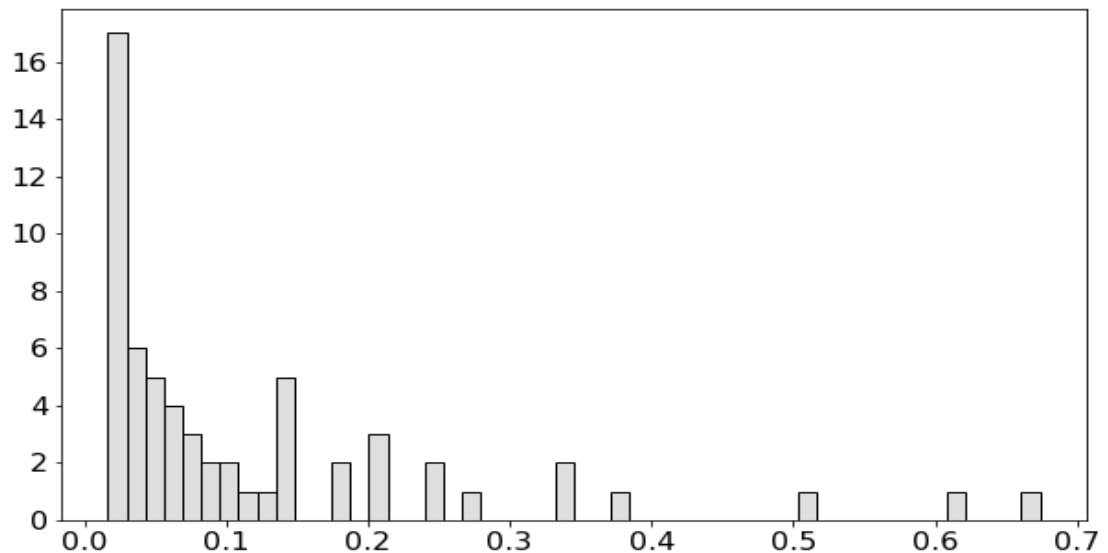
In addition, following Ibrahim and Falkenbach (2024), who established a negative, linear relationship between regional concentration and REIT bond risk premia, we also account for market concentration using the HHI. It measures the geographic concentration of a REITs' property portfolio at month t and is calculated as:

$$HHI_{i,t} = \sum_{c=1}^C \left(\frac{P_{i,c,t}}{N_{i,t}} \right)^2, \quad (3.3)$$

where $P_{i,c,t}$ is the number of properties of REIT i in CBSA c at time t , and $N_{i,t}$ is the total number of properties of REIT i at time t . HHI values can range from zero to one, with higher values indicating greater portfolio concentration. A value of one signifies that all properties are situated within a single CBSA. Figure 3.3 shows the histogram for HHI. Generally, REIT portfolios tend to be well diversified across the U.S., with a mean value of 0.0822 and most observations below 0.2. The correlation coefficient between HHI and

$\beta_{i,t}^{REIT}$ is 0.038, indicating a slightly positive relationship, suggesting that REITs with more concentrated portfolios may be more exposed to market risks due to missing diversification.

Figure 3.3 Distribution of HHI



Notes: The figure plots the distribution of the independent variable HHI. For each REIT, the mean HHI value is reported. Therefore, the total number of observations corresponds to the total number of REIT in this study. The histogram reveals a right-skewed distribution, as most values are close to zero, indicating property diversification.

3.4.4 REIT-Specific and Macroeconomic Control Variables

Monthly REIT market and quarterly REIT accounting data come from Refinitiv Eikon and comprise various relevant firm characteristics. Since not all REITs in the NAREIT index have debt offerings, we only draw REIT-specific data for those REITs that have been issuing bonds over the sample period, i.e., matching the bond data. Moreover, a REIT property sector dummy is also included, with “Diversified” as the reference category. Of the 174 REITs that were NAREIT constituents at any time between 2010 and 2023, 89 REITS have been issuing bonds. Due to missing property data reports of some REITs, we end up with a sample of 60 REITs.

Macroeconomic data are from the Federal Reserve Economic (FRED) monthly database, as presented by McCracken and Ng (2016). These measures include overall economic indicators such as GDP and unemployment rate, as well as equity and debt market metrics. Debt market metrics include the term spread and default risk spread, which serve as bond market risk exposure indicators. On the equity side, the S&P 500 index and CBOE volatility index are utilized. A comprehensive overview of all variables used in this study can be found in Table 3.2.

Table 3.2 Variable Description

Variable	Description
Dependent Variable	
Bond Risk Premium	Difference between the volume-weighted yield to maturity of the REIT bond and the maturity-matched Treasury yield to maturity
Bond Variables	
Coupon Rate	Coupon rate on bond in percent
Bond Price	Bond price inclusive of any commission, mark-ups, and/or mark-downs
Transaction volume	Volume reported on the trade
Bond Yield	Effective rate of return to maturity weighted by volume
Time to Maturity	Difference in month between maturity date and transaction date
Credit Rating Dummy	1 if investment grade bond, 0 if high yield bond
Call Status Dummy	1 if callable, 0 if not
Liquidity Variables	
Bond Zero Trading Days	Percentage of days in the month that the bond was not traded
Firm Zero Trading Days	Percentage of days in a month on which no outstanding bond of a REIT was traded
Bid-Ask Spread	Monthly mean of the difference in the average daily bid and ask prices of each bond
REIT-Specific Variables	
Local Beta	Local real estate market risk of a REIT based on its property portfolio
HHI	Geographic concentration of REIT properties across the CBSAs
Property Type Dummy	8 REIT sectors based on property portfolio, defined by NAREIT, reference category is Diversified
Market Capitalization	Measure of company size, expressed as total dollar market value of a REIT's outstanding shares of stock (in million USD)
Return	Expressed as the current adjusted price of the REIT stock minus the original adjusted price of the stock (month t-1) divided by the original adjusted price
Return on Asset	Net Income divided by total assets
Funds From Operations	Cash flow from operations
Leverage	Total debt outstanding divided by total assets
Dividends	Distribution of a REIT's earnings to its shareholders
Market-to-Book	Market capitalization divided by total assets minus total outstanding debt
Equity Market Variables	
S&P 500	Return of the S&P 500 composite index
VIX	CBOE volatility index
Macroeconomic Variables	
Term Spread	10-Year Treasury rate minus 3-Month Treasury Bill
Default Risk Spread	Moody's Baa corporate bond yield minus Moody's Aaa corporate bond yield
Interest Rate Volatility	Standard deviation of monthly 3-month constant maturity Treasury rate over a 12-month period prior to bond transaction date
Unemployment Rate	Unemployment rate in percent
GDP	Real gross domestic product (in billions of Chained 2017 USD)

Various steps are undertaken to preprocess the combined dataset. The $\beta_{i,t}^{REIT}$ and HHI are lagged as we want to use the current information to predict the following month. To address potential reverse causality, we perform Granger-causality tests, which show that local beta and HHI affect bond risk premia (at lags 3 to 5 and 4 to 5, respectively), while bond risk premia have no significant impact on them. We handle missing values by imputing data for up to one year, meaning that an existent value is being perpetuated until a new value occurs or a year has passed. The bond data in TRACE include manually reported information from broker-dealers, which can be prone to input errors, potentially leading to inaccurate data and outliers. Therefore, the yield spreads and REIT-level accounting data are winsorized at the 5% and 95% percentile to prevent outliers from distorting the results. This process is conducted monthly to ensure that future data points or REITs not included in the sample at a specific time are not incorporated. Finally, we standardize the features on a scale from -1 to 1 before feeding them into the model for training and testing. This is to ensure that importance weights are not biased by variable magnitudes. An overview of the summary statistics of the utilized variables is provided in Table 3.3.

Table 3.3 Descriptive Statistics

Variable	Unit	Mean	SD	Min	Max
Dependent Variable					
Bond Risk Premium	Percentage	0.0149	0.0092	-0.143	0.0637
Bond Variables					
Coupon Rate	Percentage	0.039	0.0114	0.0075	0.0975
Bond Price	U.S.-Dollar	100.21	10.8518	10.1375	176.504
Transaction volume	U.S.-Dollar	477,288	523,365	1,000	5,000,000
Bond Yield	Percentage	0.038	0.0271	-1.8324	0.8897
Time to Maturity	Months	98	81	2	359
Credit Rating Dummy	Percentage	0.6077	0.4883	0	1
Call Status Dummy	Percentage	0.9608	0.194	0	1
Liquidity Variables					
Bond Zero Trading Days	Percentage	0.5952	0.2013	0.2581	0.9677
Firm Zero Trading Days	Percentage	0.3531	0.0902	0.2258	0.9
Bid-Ask Spread	Percentage	0.0091	0.0092	0.0006	0.1305
REIT-Specific Variables					
Local Beta	Integer	1.0172	0.0422	0.7357	1.296
HHI	Integer	0.0822	0.0887	0.0155	0.7921
Market Capitalization	Mil. \$ (log)	23.2652	0.9565	20.2933	25.53
Return	Percentage	0.003	0.0768	-0.6003	0.4559
Return on Asset	Percentage	0.0081	0.0073	-0.0163	0.1681
Funds From Operations	Integer	5.1158	3.601	0.436	16.542
Leverage	Percentage	0.4843	0.1164	0.2716	0.8075
Dividends	U.S.-D	3.263	2.2666	0.04	11
Market-to-Book	Integer	2.232	1.3562	0.3581	8.2112
Health Care REIT	Percentage	0.1169	0.1169	0	1
Hotel REIT	Percentage	0.0253	0.1571	0	1
Industrial REIT	Percentage	0.0385	0.0385	0	1
Office REIT	Percentage	0.1612	0.3677	0	1
Residential REIT	Percentage	0.2288	0.4201	0	1
Retail REIT	Percentage	0.2792	0.4486	0	1
Specialized REIT	Percentage	0.131	0.3374	0	1
Diversified REIT	Percentage	0.019	0.0467	0	1
Equity Market Variables					
S&P 500	Integer	3,295	925.593	1,079	4,685
VIX	Integer	19.1419	6.8775	10.0785	58.0813
Macroeconomic Variables					
Term Spread	Percentage	-0.0082	0.0275	-0.0662	0.0356
Default Risk Spread	Percentage	0.0094	0.0021	0.0055	0.017
Interest Rate Volatility	Percentage	0.0041	0.0043	0.0001	0.0157
Unemployment Rate	Percentage	0.0485	0.0196	0.034	0.148
GDP	Integer	20,661	1,342	16,582	22,679

3.5 Methodology

The methodological approach of this study is based on a combination of machine learning methods and traditional econometric models, motivated by the need to uncover complex, non-linear relationships in the data while retaining interpretability and statistical rigor. Prior studies have shown that such hybrid approaches, particularly when combined with XAI methods like ALE plots, can enhance both predictive accuracy and theoretical insight by

linking flexible model structures with transparent, hypothesis-driven inference (see, e.g., Kraemer et al., 2023b; Lorenz et al., 2023).

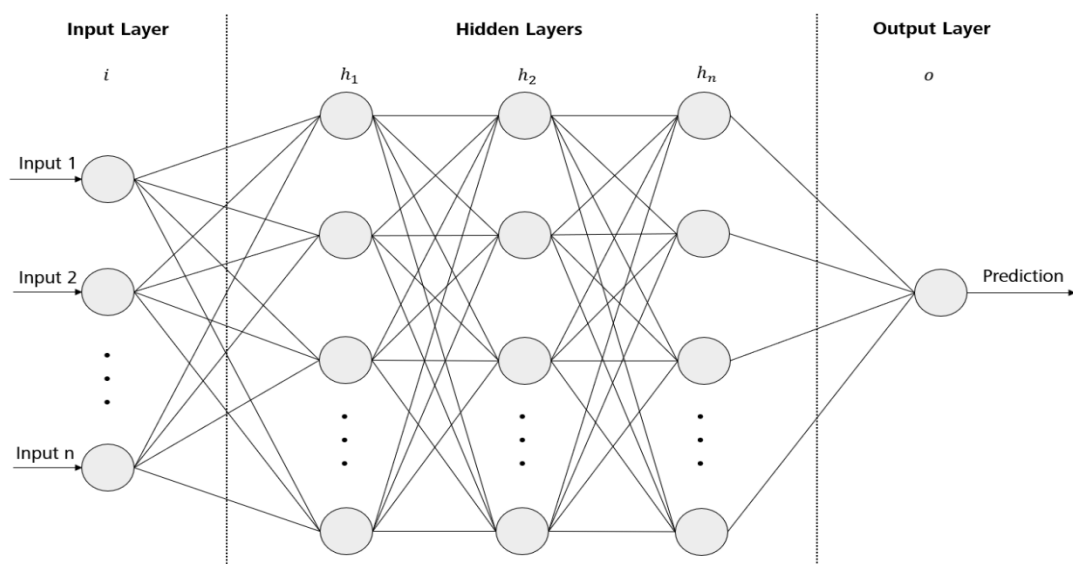
Therefore, we first employ an ANN with an expanding window scheme to derive predictions based on the preprocessed data. Subsequently, we make use of ALE plots to allow for an interpretation of the potential non-linear impacts of the variables on the average REIT bond risk premium prediction. We then specify an OLS regression with polynomial transformations based on the insights of the ALE plots.

3.5.1 Artificial Neural Network

Prior studies on REITs have shown the superior performance of ANNs over traditional regression methods due to their ability to address the non-linearity and non-stationary data properties (see, e.g., Loo, 2020; Serrano and Hoesli, 2007). We chose an ANN over simpler non-linear models, such as Generalized Additive Models, because its structure is even more flexible and allows us to capture complex interactions and inflection points without imposing restrictive functional assumptions. Prior studies suggest that such machine learning approaches yield more meaningful and accurate predictions compared to both linear and semi-parametric alternatives (Kraemer et al., 2023b; Lorenz et al., 2023).

The structure of an ANN consists of three types of layers: an input layer, a variable number of computational layers called hidden layers, and an output layer. In the input layer, each node represents one independent variable. In the hidden layers, the variables are transformed non-linearly and passed onto the next layer. Eventually, the output layer contains the predicted bond risk premium. Figure 3.4 illustrates a simplified ANN used in the study.

Figure 3.4 Exemplary ANN Architecture



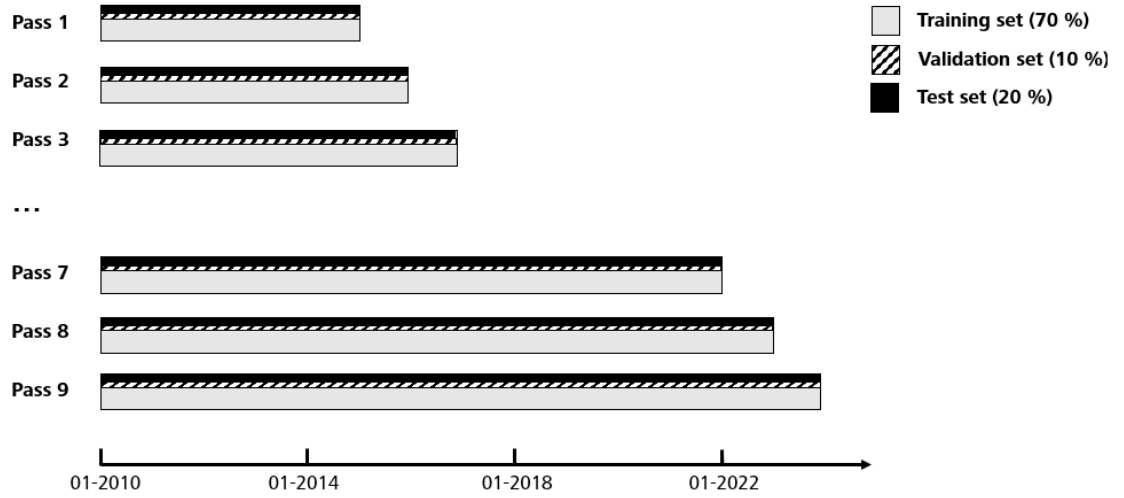
Additionally, a number of hyperparameters are defined and tuned by a hyperparameter search procedure. In this study, we use a randomized hyperparameter search method where the search space spans a distribution, allowing us to identify the most optimal hyperparameter combination. Appendix II shows the hyperparameter search spaces in detail.

To prevent overfitting of the ANN models to the training data this study employs four regularization techniques. Firstly, a dropout layer is added after every hidden layer. Srivastava et al. (2014) prove that randomly dropping neurons along with their connections during training significantly reduces the overfitting of ANNs. Secondly, an L2 kernel regularization is used in the hidden layers, allowing weight penalties during the optimization process. Thirdly, an early stopping parameter is added to the training process, terminating the optimization when the validation loss fails to show improvement over 50 consecutive iterations. Fourthly, for each month in the prediction process, 20 ANNs are fitted, and the 10 instances exhibiting the lowest validation loss are chosen for the out-of-sample predictions. The models' predictions are averaged and reported for this study. Using the mentioned regularization tools simultaneously can improve model performance, as shown by Gu et al. (2020).

3.5.2 Model Setup

An expanding window approach based on Gu et al. (2020) is used to deal with the time-series nature of the data. The expanding window process is illustrated in Figure 3.5. During the first iteration, 70% of the data from the first five years is utilized for training, 10% for validation, and the remaining 20% is reserved for testing. The model's hyperparameter search is done at the beginning of each year using the training and validation data. With every iteration, the dataset expands by one year to account for the time variation of the input features. Fixed time effects are included to control for heterogeneity over time. The advantage of this approach is that the algorithm tests on data it has not seen before and can provide robust, unbiased results.

Figure 3.5 Cross-Sectional Expanding Window Approach



3.5.3 Accumulated Local Effect (ALE) Plots

ANNs are often referred to as "black boxes" because it is not readily apparent how the model learns the relationships between input and output data. The branch of research that focuses on making ANNs and other machine learning methods interpretable is called XAI. The global model-agnostic method of ALE plots provides insights into how an independent variable influences the average prediction of the dependent variable across all its values in the dataset. A drawback of other methods, such as Shapley Additive Explanations (SHAP) by Lundberg and Lee (2017) and Partial Dependency Plots (PDP) by Friedman (2001), is that they assume that variables are uncorrelated, which is uncommon in finance and real estate applications. ALE plots address this issue by allowing for correlated features (Apley and Zhu, 2020). As a result, they offer a more reliable approach to understanding the effects of independent features on predictions. ALE plots are gaining increasing attention in the real estate literature, as evidenced by studies such as Lorenz et al. (2023) or Kraemer et al. (2023a, 2023b).

The total range of observed values for each independent feature $x_r \in \mathbb{R}^{N \times 1}$ is split into K bins. We define $Z_{r,k}$ as the $\frac{k}{K}$ quantile of its empirical distribution with $Z_{r,0}$ being the minimum and $Z_{r,K}$ the maximum of Z_r . Moreover, assume that $S_{r,k}$ defines the array of values within the left open interval from $Z_{r,k-1}$ to $Z_{r,k}$ with $n_{r,k}$ being the number of observations within the interval $S_{r,k}$. We define $k(x_r)$ as an index returning the bin for a value of x_r . Consequently, the uncentred accumulated local effect can be stated as:

$$g_{ALE}(x_r) = \sum_{k=1}^{k(x_r)} n_{r,k}^{-1} \sum_{i \in S_{r,k}} [f(Z_{r,k}, X_{\setminus r,i}) - f(Z_{r,k-1}, X_{\setminus r,i})], \quad (3.4)$$

where $x_{\setminus r} \in \mathbb{R}^{N \times P-1}$ represents the set of all features except for feature r among the total P variables, and $f(\cdot)$ denotes the model's prediction function. The expression within the squared brackets captures the prediction of $f(\cdot)$ when the observation i is replaced with $Z_{r,k}$ in the minuend and the prediction with $Z_{r,k-1}$ instead of the true value i in the subtrahend. These differences are then summed over all observations contained in $S_{r,k}$. Since this procedure is repeated for every bin k , the term $g_{ALE}(x_r)$ represents the total sum of these inner differences, weighted by the number of observations in each bin (Apley and Zhu, 2020). The centered ALE main effect is normalized to have a zero mean with respect to the marginal distribution of x_r , and can be expressed as:

$$\Theta_{ALE}(x_r) = g_{ALE}(x_r) - N^{-1} \sum_{i=1}^N g_{ALE}(x_{r,i}). \quad (3.5)$$

To ensure that the final ALE plots remain interpretable, they are being centered. In this way, the plot consistently illustrates the impact of a feature as a function of the average prediction of the ANN across the entire feature space. The visual analysis of ALE plots allows us to detect any non-linear relationships between the independent variables and the predicted bond risk premium.

3.5.4 Ordinary Least Squares Regression

A crucial limitation of machine learning methods such as ANNs is their lack of interpretability regarding statistical significance, which motivates the complementary use of an OLS regression. In this regression, all variables for which non-linear relationships were found in the ALE plots are modeled as the n^{th} degree polynomial based on the inflection points observed in the data. The OLS is set up on a cross-sectional basis with an expanding window and time-fixed and property-fixed effects. We estimate the following linear regression model using monthly REIT bond yield spreads:

$$\begin{aligned} r_{i,t}^{REIT} - r_{i,t}^{Treasury} = & c_0 + \beta_1 \beta_{i,t}^{REIT} + \beta_2 HHI + \sum_{i=1}^7 \chi_i \text{Bond Controls} \\ & + \sum_{i=1}^3 \chi_i \text{Liquidity Controls} + \sum_{i=1}^8 \chi_i \text{REIT Controls} \\ & + \sum_{i=1}^2 \chi_i \text{Equity Market Controls} + \sum_{i=1}^5 \chi_i \text{Macroeconomic Controls} \\ & + \gamma_t^{Time} + \gamma_t^{Property Type} + \epsilon_{i,t}, \end{aligned} \quad (3.6)$$

where $r_{i,t}^{REIT} - r_{i,t}^{Treasury}$ is the monthly REIT bond yield spread as the difference between the monthly REIT bond yield of REIT i at time t and the U.S. Treasury yield with matched maturity i at time t . $\beta_{i,t}^{REIT}$ is the local real estate market beta of REIT i at time t . γ_t^{Time} is

the year-fixed effects, $\gamma_t^{Property\ Type}$ are property type-fixed effects and $\epsilon_{i,t}$ is the residual. Table 3.4 shows the different performance metrics used to evaluate the ANN and OLS.

Table 3.4 Employed Performance Metrics

Error Metric	Formula
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\bar{y})^2}$
Mean Squared Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $

3.6 Results

The following section provides the results of our empirical analysis. Firstly, we analyze the ALE plots (specifically for local beta and HHI) to assess non-linear relationships between these variables and the REIT bond risk premium. Secondly, we specify an optimized OLS based on the uncovered non-linear data patterns in the ALE plots including all control variables. We then assess the drivers of REIT bond risk premia, focusing on local real estate market risk, geographic concentration and their interplay.

3.6.1 Results ALE Plots

To better understand the influence of a single feature on the ANN model's predictions, we turn to the ALE plots, visualizing the data patterns uncovered by the ANN. The ALE plots show how a feature interacts with the target variable REIT bond risk premia. The values on the y-axis show how much the risk premia prediction deviates from the average prediction when the variable takes on certain values indicated on the x-axis. The results of the ALE plots can be used to improve the performance of other parametric or semi-parametric models by specifying the functional form for each feature. To assess the presence of non-linearities in the data, we display a univariate OLS fitted line in red alongside the effect direction of the ANN as shown in the ALE plot. We can then conclude whether the OLS can capture the functional form of the variables well or whether it should be specified manually based on the uncovered patterns of the ALE plots.

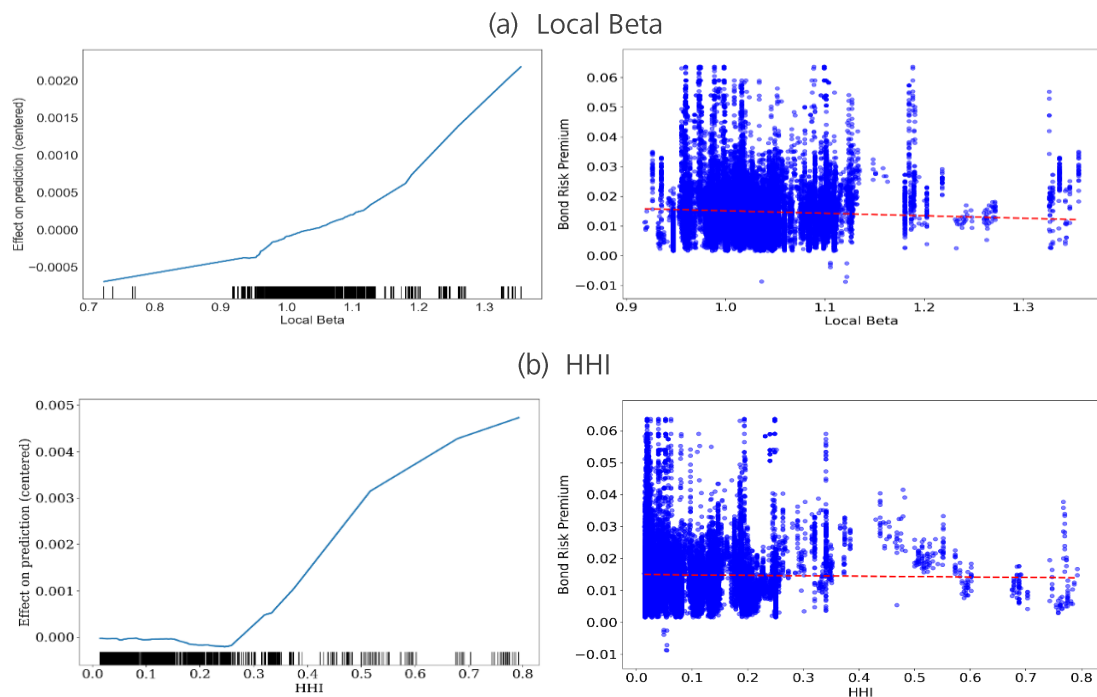
In general, for most variables, the ANN is able to pick up non-linear relationships in the data. Only two variables, namely volume and bid-ask-spread, have a linear relationship with REIT bond risk premia. In addition, with the ALE plots, we can assess whether there is a positive or negative relationship across all variable values or whether the relationships

are changing. It is evident that an ANN, together with ALE plots, can uncover non-linearities that help to specify a stronger OLS in the next step. Appendix III and Appendix IV show the ALE plots of all control variables and the corresponding univariate OLS.

Figure 3.6 (a) shows the effect of the local beta factor, i.e., the risk of the local real estate market, on the risk premium prediction of REIT bonds. The relationship between the local beta and the REIT bond risk premia indicates that higher local real estate market risk is reflected in higher risk premia for the REIT bonds. We assume that this is because investors want to be compensated for the additional risk they take on when holding REIT bonds exposed to more volatile real estate markets, which could result in less predictable cash flows and asset values.

Figure 3.6 (b) shows the effect of geographic concentration measured as HHI on the risk premia prediction of REIT bonds. In line with modern portfolio theory, which posits diminishing benefits from diversification, the relationship exhibits a clear non-linear pattern. Starting from high concentration levels (i.e., high HHI), increased geographic diversification (moving from right to left along the graph) reduces predicted bond risk premia up to an HHI of around 0.25. Beyond this point, however, further diversification is associated with rising risk premia, suggesting that the costs of managing highly dispersed portfolios, such as increased complexity or reduced transparency, may outweigh the risk-reducing benefits.

Figure 3.6 ALE Plot and Univariate Regression of Local Beta and HHI



3.6.2 Optimized Regression Results

Based on the observed patterns, the baseline variables are transformed to higher degree terms. These polynomial transformations of all control variables in the OLS based on the inflections of the ALE plots can be seen in Appendix V. Table 3.5 shows the regression summary for all model specifications. The coefficients are mostly in line with the shown relationship in the ALE plots in Figure 3.6 and Appendix III and Appendix IV. Also, the control variables are in line with previous literature (e.g., default risk captured by market capitalization, see Lui (2021) and liquidity risk see Longstaff et al. (2005)).

The results of Model 1 (the optimized OLS based on inflections in the ALE plots) show that local real estate market risk, i.e., the local beta, is statistically significant in all model specifications. It becomes evident that with an increase in real estate market risk, there is an increase in REIT bond risk premia. Thus, we can confirm that investors require higher bond risk premia to be adequately compensated for higher exposure to riskier local real estate markets. However, the magnitude of the local beta coefficient is smaller than that observed in the study of Zhu and Lizieri (2022) concerning REIT equity returns. This aligns with our theoretical argument that bondholders, due to their higher priority of claims are less sensitive to geographic risks compared to equity holders.

For robustness reasons, Model 2 shows that this positive relation between local real estate market risk and REIT bond risk premia is consistent across all REITs, irrespective of their exposure to riskier local real estate markets (high beta REITs, i.e., local beta above median) or more stable local real estate markets (low beta REITs, i.e., local beta below median). In monetary terms, for a typical \$500 million bond issuance for REITs in the high local beta group, an increase in local beta from the 25th to the 75th percentile translates into approximately \$295,711 to \$583,999 in additional annual interest costs, illustrating the tangible financial impact of local real estate market risk on REIT debt financing.

The analysis of Models 1 and 2 reveals a statistically significant negative relationship between the HHI and REIT bond risk premia, suggesting that greater geographic concentration (or lower diversification) corresponds with reduced risk premia across the entire dataset. However, the ALE plot in Figure 3.6 (b) shows a structural break in the data around the value of an HHI of 0.25. To account for this observed change in the data, we specify Model 3, where the sample is divided into REITs with high geographic concentration (HHI above 0.25), and those with low geographic concentration (HHI below 0.25). The results show that increased geographic concentration for REITs with higher HHI relates to higher bond risk premia. Conversely, for REITs with lower HHI, increased geographic concentration is associated with reduced bond risk premia. This suggests that

for more concentrated REITs, the advantages of geographic diversification, such as reduced earnings volatility, outweigh the associated information and monitoring costs, leading investors to require lower risk premia. In contrast, for more diversified REITs, the additional costs of further geographic diversification surpass its benefits, leading to higher required bond risk premia. This adds a public debt perspective to the discussion on the coined “diversification discount” found in previous literature for REIT value and other operating metrics (see, e.g., Feng et al., 2021; Hartzell et al., 2014; Huerta and Mothorpe, 2024; Ibrahim and Falkenbach, 2024). Ibrahim and Falkenbach (2024) find a REIT bond risk premium for geographic diversification, regardless of the REIT’s level of diversification. Our finding that the effect of diversification depends on the given diversification level of a REIT is consistent with modern portfolio theory, which suggests that the benefits of diversification diminish as diversification increases (Markowitz, 1952). This non-linear pattern may have been overlooked in prior studies on geographic diversification. To further validate the structural break in HHI, we perform a Chow test, which confirms a significant break at the HHI level of 0.25.

However, it is important to note that the effect size (0.48%) of geographic concentration on risk premia for REITs with high HHI is less pronounced than the effect size for REITs with lower HHI (1.71%). It seems that investors require a higher compensation for the costs of additional geographic diversification for REITs that are already well diversified (i.e., REITs with an HHI <0.25) than they are willing to offer a REIT bond risk premia discount for geographic diversification for concentrated REITs (i.e., REITs with an HHI >0.25). This new perspective reinforces the benefits of combining machine learning and traditional methods.

Since the effect of geographic diversification on risk premia appears to depend on the level of concentration, it implies that for REITs exposed to more volatile markets, diversification might influence bond costs differently than for REITs with a lower local beta. To explore this relationship more formally, Model 4 includes an interaction term between local beta and HHI. The interaction term yields a coefficient of 0.0191, indicating that the positive impact of local beta on bond risk premia is amplified in more concentrated markets. Contrary, this suggests that increased geographic diversification for REITs in high-risk markets may lead to significantly lower bond risk premia, effectively mitigating volatility-related concerns. This supports the hypothesis that REITs exposed to volatile markets benefit more from diversification and concentrating in a few areas becomes costly in terms of bond risk premia beyond a certain point. Conversely, in more stable markets, the same level of diversification might incur additional management costs and operational complexities, which could explain the inflection point for HHI at 0.25. In other words, while

concentration alone might reduce risk premia initially, for REITs operating in more volatile markets, concentration beyond a certain point is associated with increased risk premia. This finding emphasizes the importance of maintaining an optimal balance in geographic diversification and risk market strategies.

Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT
Public Debt

Table 3.5 Regression Results

	(1) Polynomial	(2) Beta	(3) HHI	(4) Interaction
$\beta_{i,t-1}^{REIT}$	0.0021*** (0.001)		0.0022*** (0.001)	0.0004* (0.001)
High $\beta_{i,t-1}^{REIT}$		0.0323*** (0.001)		
Low $\beta_{i,t-1}^{REIT}$		0.0321*** (0.001)		
HHI _{t-1}	-0.0015*** (0.001)	-0.0015*** (0.001)		-0.0205** (0.008)
High HHI _{t-1}			0.0048*** (0.001)	
Low HHI _{t-1}			-0.0171*** (0.001)	
$\beta_{i,t-1}^{REIT} * HHI_{t-1}$				0.0191** (0.008)
Volume	1.477e-10** (6.64e-11)	1.469e-10** (6.64e-11)	1.461e-10** (6.64e-11)	-1.421e-10** (6.64e-11)
Bond Price	-0.0003*** (5.84e-06)	-0.0003*** (5.83e-06)	-0.0003*** (5.82e-06)	-0.0003*** (5.84e-10)
Coupon Rate	0.2421*** (0.005)	0.2435*** (0.005)	0.2401*** (0.005)	0.2428*** (0.005)
Volume Weighted Yield	0.0485*** (0.001)	0.0485*** (0.001)	0.0486*** (0.001)	0.0485*** (0.001)
Time to Maturity	5.849e-06*** (5.26e-07)	5.731e-06*** (5.25e-07)	5.929e-06*** (5.26e-07)	5.875e-06*** (5.26e-07)
Credit Rating	0.0013*** (9.05e-05)	0.0013*** (9.03e-05)	0.0013*** (9.04e-05)	0.0013*** (9.06e-05)
Call Status	-0.0023*** (0.000)	-0.0023*** (0.000)	-0.0024*** (0.000)	-0.0023*** (0.000)
Market Cap	-0.0036*** (6.03e-05)	-0.0037*** (6.07e-05)	-0.0037*** (6.03e-05)	-0.0036*** (6.03e-05)
Equity Return	0.0025*** (0.000)	0.0025*** (0.000)	0.0025*** (0.000)	0.0025*** (0.000)
Leverage	0.0073*** (0.001)	0.0074*** (0.000)	0.0073*** (0.001)	0.0075*** (0.001)
Return on Assets	-0.0682*** (0.005)	-0.0694*** (0.005)	-0.0689*** (0.005)	-0.0679*** (0.005)
Market-to-Book	8.17e-05 (5.05e-05)	9.378e-05* (5.06e-05)	8.762e-05* (5.04e-05)	9.024e-05* (5.07e-05)
Default Risk Spread	1.0788*** (0.026)	1.0781*** (0.026)	1.0803*** (0.026)	1.0789*** (0.026)
IVOL	-0.1114*** (0.013)	-0.1112*** (0.013)	-0.1105*** (0.013)	-0.1114*** (0.013)
Bond Zero Trading Days	-0.0049*** (0.000)	-0.0049*** (0.000)	-0.0049*** (0.000)	-0.0049*** (0.000)
Firm Zero Trading Days	-0.0041*** (0.000)	-0.0042*** (0.000)	-0.0042*** (0.000)	-0.0041*** (0.000)
Bid-Ask Spread	0.0869*** (0.004)	0.0868*** (0.004)	0.0870*** (0.004)	0.0864*** (0.004)
Health Care	0.0019*** (0.000)	0.0021*** (0.000)	0.0019*** (0.000)	0.0019*** (0.000)
Hotel & Resort	0.0063*** (0.000)	0.0063*** (0.000)	0.0063*** (0.000)	0.0061*** (0.000)
Industrial	0.0021*** (0.000)	0.0021*** (0.000)	0.0020*** (0.000)	0.0021*** (0.000)
Office	0.0003 (0.000)	0.0003 (0.000)	0.0004 (0.000)	0.0001 (0.000)
Residential	-0.0003 (0.000)	-0.0002 (0.000)	1.791e-05 (0.000)	-0.0003 (0.000)
Retail	-0.0010*** (0.000)	-0.0009*** (0.000)	-0.0010*** (0.000)	-0.0010*** (0.000)
Specialized	0.0021*** (0.000)	0.0022*** (0.000)	0.0022*** (0.000)	0.0021*** (0.000)
R ²	0.729	0.727	0.728	0.729

Notes: Dependent Variable: REIT bond yield spread in decimals (e.g., 0.012 = 1.2%). Standard errors are reported in parenthesis. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively. The variables Dividends, FFO, VIX, S&P 500, Unemployment Rate, Term Spread, and GDP were excluded from the OLS for multicollinearity reasons (VIF values exceeded 10). To further ensure that multicollinearity is not affecting the standard errors, we also estimate the regressions without control variables while keeping the fixed effects. The results remain robust. All model specifications are based on the polynomial regression with variable transformations reported in Appendix V.

3.6.3 Performance Evaluation of Applied Models

Table 3.6 illustrates the performance indicators of the ANN and the OLS over the entire sample period. $OLS_{Baseline}$ represents the regression without any variable transformations (which is not displayed in Table 3.5 and used here solely for performance comparison), while $OLS_{Polynomial}$ corresponds to Model 1 in Table 3.5 and includes polynomial transformations guided by the ALE plots. The results show that incorporating insights from the ALE plots improves the OLS model's fit, with R^2 increasing from 67.1% in the basic model to 72.9% in the polynomial specification. The ANN displays the highest explanatory power, with an R^2 value of 82.5%, reinforcing the value of using machine learning models to inform and enhance traditional econometric models such as an OLS.

Table 3.6 Performance Comparison of All-Year Models

Metrics	$OLS_{Baseline}$	$OLS_{Polynomial}$	$\Delta OLS_{Polynomial} / OLS_{Baseline}$	ANN	$\Delta ANN / OLS_{Baseline}$
MSE	0.000026	0.000022	-15,4%	0.000015	-42,3%
MAE	0.0034	0.0032	-5,9%	0.0028	-17,6%
R^2	0.671	0.729	+8,6%	0.825	+23,0%

3.7 Conclusion

To the best of our knowledge, this study is the first to examine the relationship between local real estate market risk and REIT bond risk premia. The motivation behind our analysis is that REITs exposed to more volatile local real estate markets (i.e., REITs with a local beta above 1) should have more volatile asset values, thereby increasing default risk. As a result, we hypothesize that bondholders will demand higher risk premia from REITs with greater exposure to local market volatility, and that this effect may differ from what is observed for equity holders, given bondholders' limited upside, priority of claims, and heightened sensitivity to downside risk. We investigate whether these local asset-level risks, as well as geographic diversification, are reflected in the pricing of REIT bonds.

Our dataset includes 30,186 monthly observations of REIT bond risk premia from 2010 to 2023. To capture local market risk, we construct a local beta factor using REIT-level property data, which measures the sensitivity of local real estate markets to aggregate national real estate market shocks. Additionally, we quantify geographic concentration of REIT portfolios by calculating HHI values for each REIT.

This study makes several contributions to the literature on geographic characteristics as determinants of commercial real estate investment risk and return. Regarding our first

research question, we find a local real estate market risk premium in REIT bonds. Investors demand higher compensation for REITs that are exposed to more volatile local markets. This positive relationship between local-market volatility and bond risk premia holds across the entire sample, both for REITs concentrated in high-volatility areas and for those operating mainly in more stable regions. This aligns with the notion of structural models of credit risk, such as Merton (1974), where increased asset volatility elevates the default risk perceived and priced by debt investors. As expected, the effect is less pronounced for REIT bonds compared to REIT equity as shown by Zhu and Lizieri (2022) which can be explained by the differences in risk-return profiles and claim order in the capital structure between bondholders and shareholders.

Turning to our second research question, we reveal a more nuanced, non-linear impact of geographic diversification on REIT bond risk premia, consistent with modern portfolio theory's assumption that diversification benefits only up to a certain threshold. For REITs with more concentrated portfolios, increased diversification is associated with lower bond risk premia, suggesting that the benefits of diversification, such as more stable cash flows, outweigh the potential costs related to information and monitoring. However, for REITs that are already well-diversified, further diversification leads to higher bond risk premia, indicating that the additional costs of diversification outweigh its benefits. This finding refines the concept of a geographic diversification discount in REIT bonds, as documented in prior research. We identify a turning point at an HHI concentration value of around 0.25, which marks the threshold where further diversification becomes less beneficial from a REITs perspective.

Finally, addressing our third research question, we reveal an important interaction effect: in more volatile markets (high local beta), higher geographic concentration (high HHI) leads to a disproportionately larger increase in bond risk premia. This interaction suggests that the combined effect of local market volatility and concentration amplifies risk, especially in high-risk environments.

Methodologically, we demonstrate the value of integrating non-parametric machine learning approaches, such as ANNs and ALE plots, to enhance traditional parametric models like OLS. By using ALE plots to visualize non-linear relationships, we optimize the OLS model and achieve improved predictive performance, offering a methodological contribution that extends beyond the REIT bond literature and can be applied to other asset classes or contexts.

Understanding the link between fundamental location characteristics of commercial real estate and REIT bond risk premia provides valuable insights for REIT managers, investors,

and bond portfolio managers. For REIT managers, these findings offer a framework for optimizing geographic market allocations over the long term, with the aim of lowering the cost of future public debt issuances. While short-term portfolio restructuring in response to specific issuance conditions is generally not feasible due to asset illiquidity and transaction costs, strategic adjustments in acquisition, development, and divestment decisions over time can shape geographic exposure in ways that benefit future financing costs. For investors, the results show that local real estate market risk is a material driver of REIT bond risk premia in the secondary market, providing an additional lens for credit risk assessment beyond firm-level financial metrics. Bond portfolio managers can leverage these insights to refine pricing models, improve investment decisions, and enhance risk-adjusted returns.

3.8 Appendix

3.8.1 Appendix I – Rolling Local Beta Construction and OLS Results

For robustness reasons, we additionally calculate a rolling local beta, denoted as $\beta_{c,y}^{REIT}$, which captures the local real estate market risk for CBSA c in year y . This is achieved by applying the regression specified Equation 3.1 annually, yielding time-varying local betas for each CBSA:

$$r_{c,(y-1,y)}^{NPI} - r_{f,(y-1,y)} = \alpha_{c,y} + \beta_{c,y} (MKT_{(y-1,y)}^{NPI} - r_{f,(y-1,y)}) + \varepsilon_{c,y}. \quad (3.7)$$

Figure 3.7 displays the average rolling $\beta_{c,y}^{REIT}$ for all CBSAs and for the ten top and bottom 10 CBSAs. While the average $\beta_{c,y}^{REIT}$ of all CBSAs is around 1 and therefore in line with the average of β_c , the top, and bottom 10 CBSAs exhibit more extreme spreads ($\beta_{c,y}^{REIT}$ values > 4 and < -3) compared to the histogram in Figure 3.1. This is likely because the fixed beta (β_c) smoothes out short-term property market volatility, leading to more stable coefficients. Moreover, the top 10 and bottom 10 CBSAs show more fluctuations compared to the overall average. The regression results are displayed in Table 3.7 below. The results remain robust.

Figure 3.7 Rolling Betas over Time

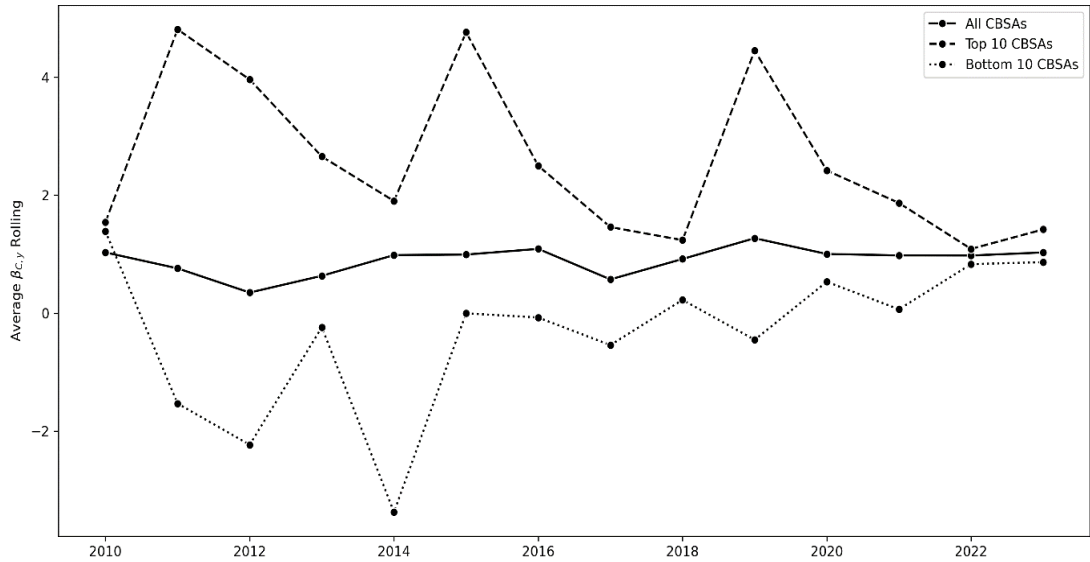


Table 3.7 Rolling Beta Regression Results

Variables	Rolling Beta Coefficients
$\beta_{i,t-1}^{REIT}$	0.0021*** (0.001)
HHI _{t-1}	-0.0015*** (0.001)
Volume	1.477e-10** (6.64e-11)
Bond Price	-0.0003*** (5.84e-06)
Coupon Rate	0.2421*** (0.005)
Volume Weighted Yield	0.0485*** (0.001)
Time to Maturity	5.849e-06*** (5.26e-07)
Credit Rating	0.0013*** (9.05e-05)
Call Status	-0.0023*** (0.000)
Market Cap	-0.0036*** (6.03e-05)
Equity Return	0.0025*** (0.000)
Leverage	0.0073*** (0.001)
Return on Assets	-0.0682*** (0.005)
Market-to-Book	8.17e-05 (5.05e-05)
Default Risk Spread	1.0788*** (0.026)
IVOL	-0.1114*** (0.013)
Bond Zero Trading Days	-0.0049*** (0.000)
Firm Zero Trading Days	-0.0041*** (0.000)
Bid-Ask Spread	0.0869*** (0.004)
Health Care	0.0019*** (0.000)
Hotel & Resort	0.0063*** (0.000)
Industrial	0.0021*** (0.000)
Office	0.0003 (0.000)
Residential	-0.0003 (0.000)
Retail	-0.0010*** (0.000)
Specialized	0.0021*** (0.000)
R ²	0.728

Notes: Dependent Variable: REIT bond yield spread in decimals (e.g., 0.012 = 1.2%). Standard errors are reported in parenthesis. ***, ** and * denote significance at the 1%, 5%, and 10% level, respectively. The variables Dividends, FFO, VIX, S&P 500, Unemployment Rate, Term Spread, and GDP were excluded from the OLS for multicollinearity reasons (VIF values exceeded 10). To further ensure that multicollinearity is not affecting the standard errors, we also estimate the regressions without control variables while keeping the fixed effects.

3.8.2 Appendix II – ANN Hyperparameter Distribution

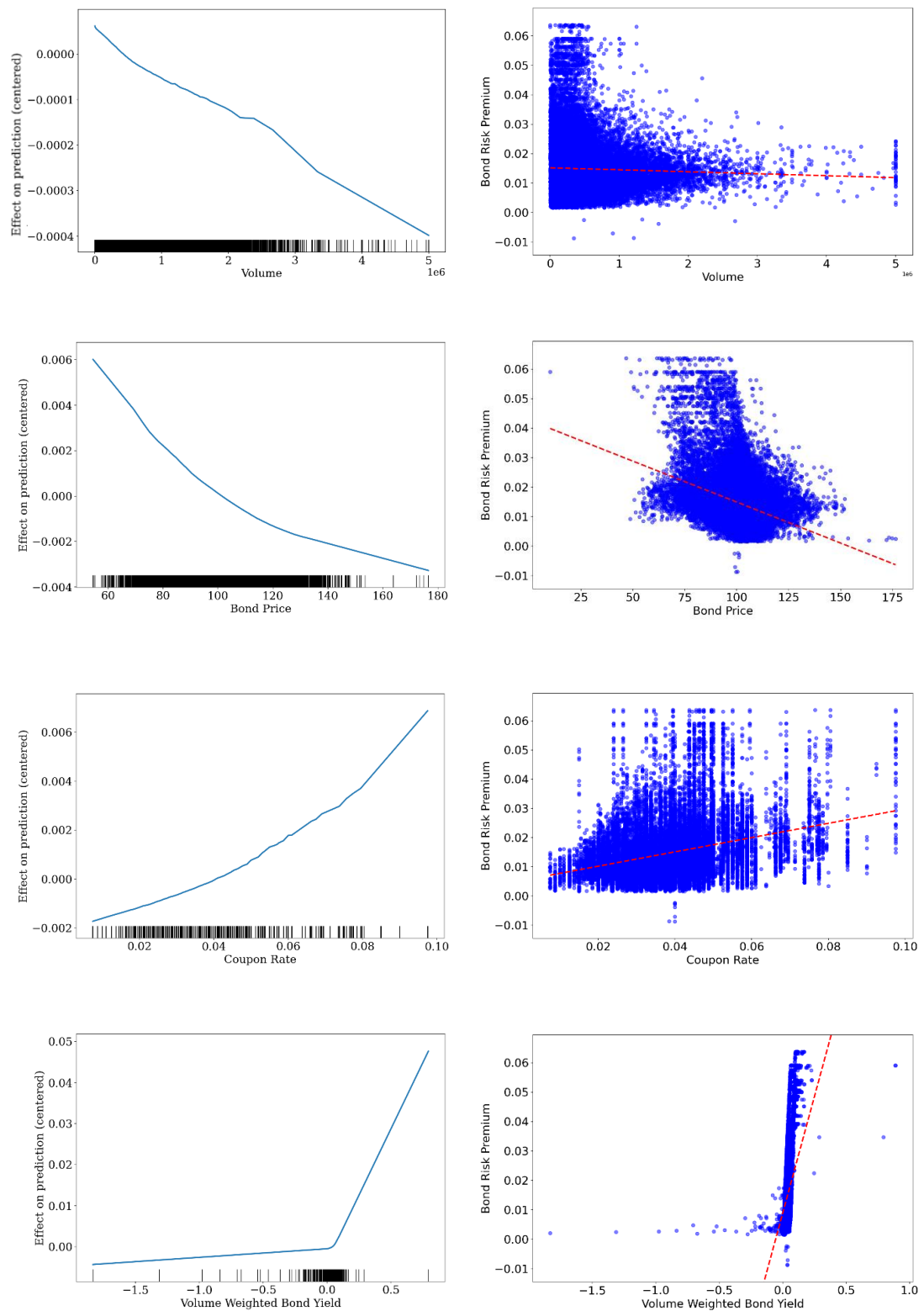
Turning the hyperparameters is a crucial aspect in setting up the right network architecture, which are defined ex-ante and refer to the parameters regulating the model design. Hyperparameters being tuned in this study include the learning rate, the dropout rate, the number of hidden layers, lambda L2, and the multiple. Table 3.8 displays the distribution of each hyperparameter. Every constellation of hyperparameters is fitted annually on the training data and subsequently applied to the validation set. We chose not to include the activation function in the hyperparameter search but use the ReLu function for each hidden layer. This is because the ReLu function is less likely to create a vanishing gradient problem and can be computed efficiently, which are major benefits compared to other activation functions (Wang et al., 2020).

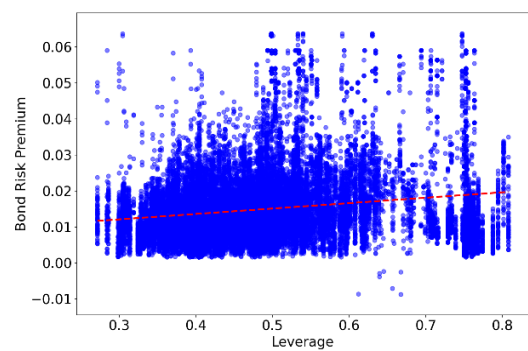
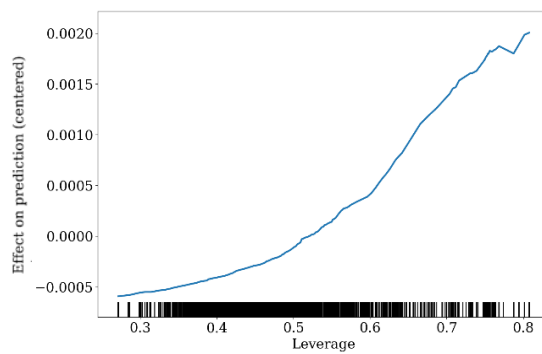
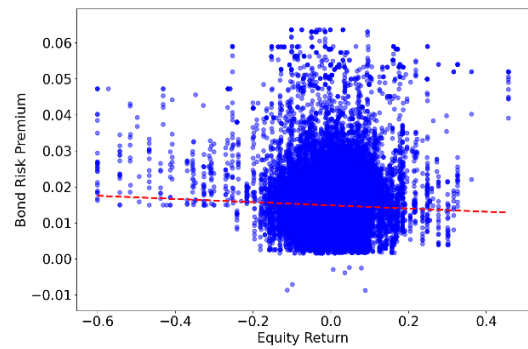
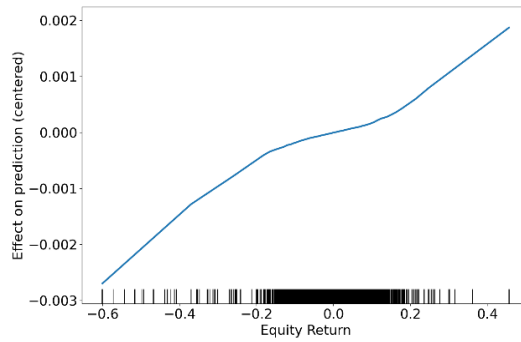
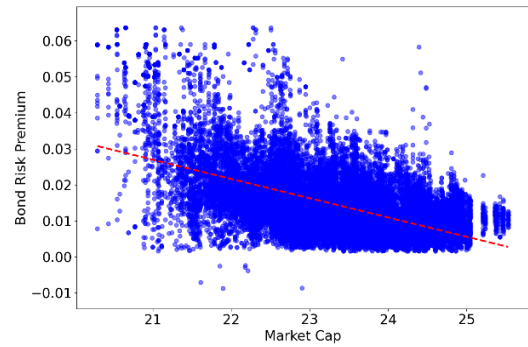
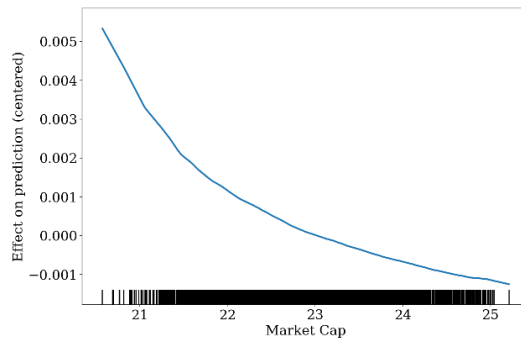
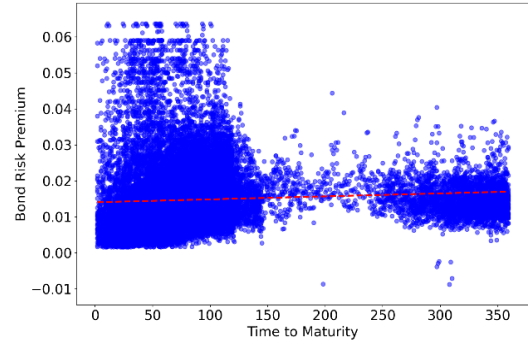
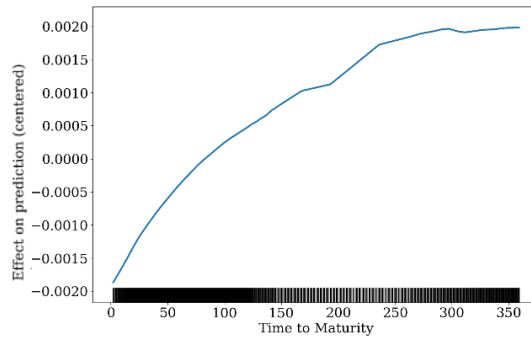
Table 3.8 Hyperparameter Distribution

Hyperparameter	Distribution
Hidden Layer	$U^d \sim [1, 4]$
Neuron Multiple	$U^d \sim [1, 6]$
Lambda L2	$U^c \sim [0.000001, 0.01]$
Learning Rate	$U^c \sim [0.0001, 0.01]$
Dropout Rate	$U^c \sim [0.10, 0.40]$
Hidden Layer	$U^d \sim [1, 4]$
Neuron Multiple	$U^d \sim [1, 6]$

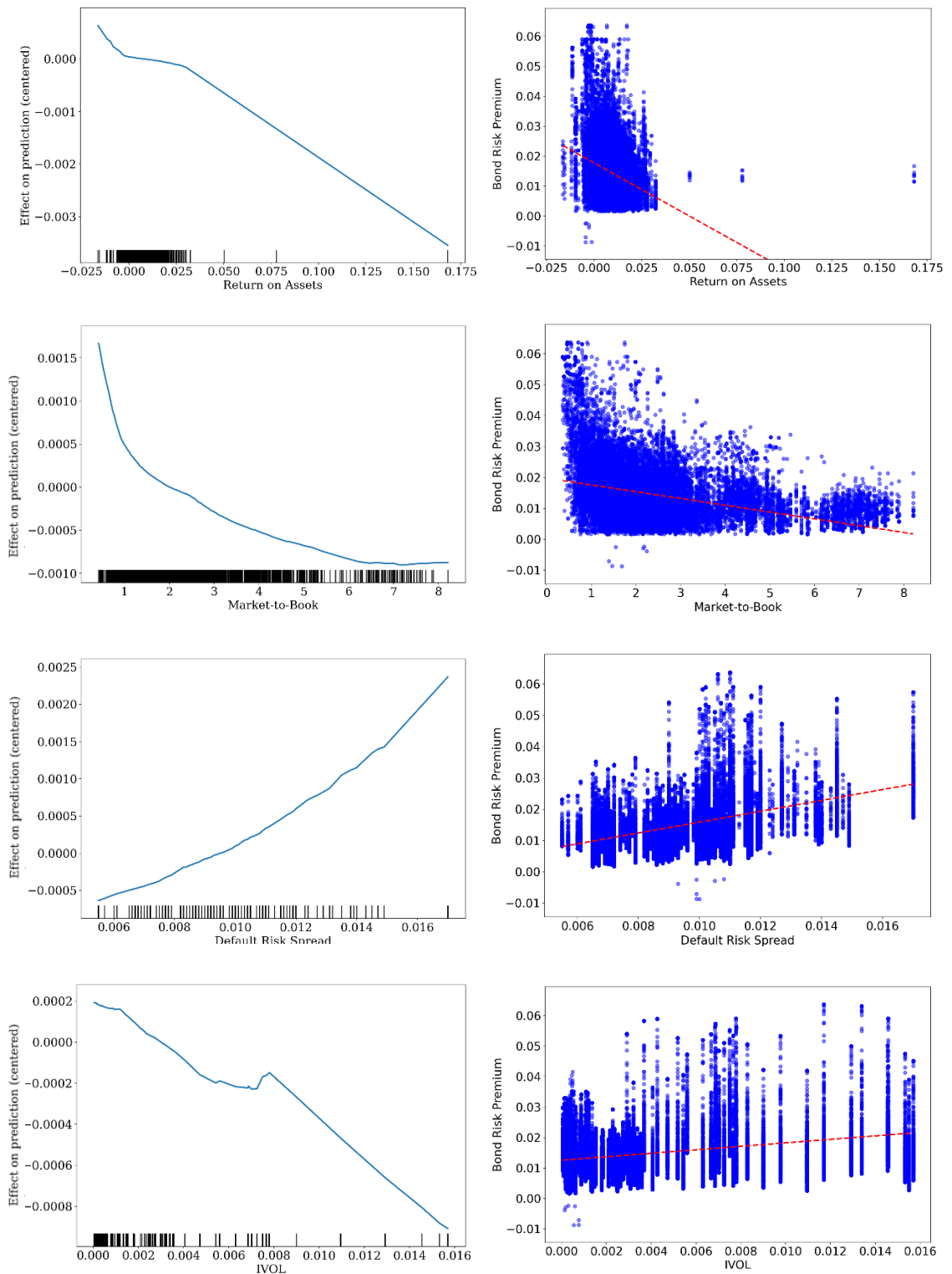
3.8.3 Appendix III – ALE Plots and Regressions of Continuous Variables

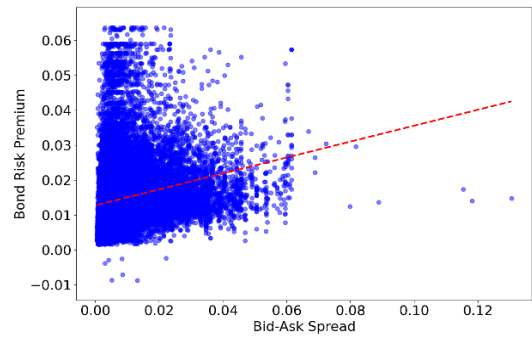
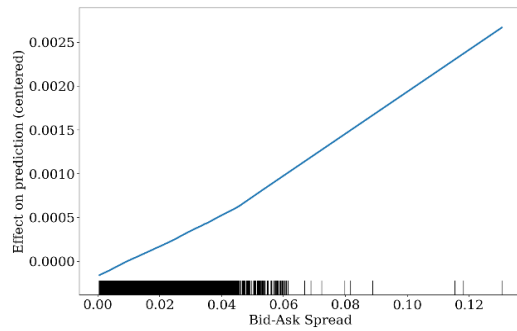
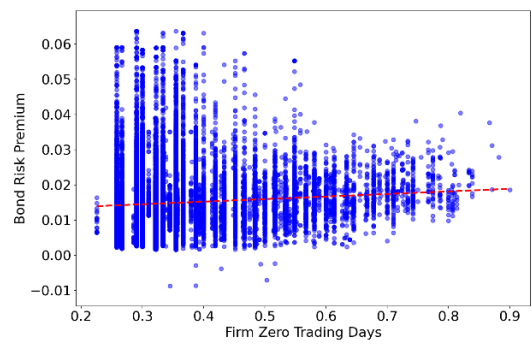
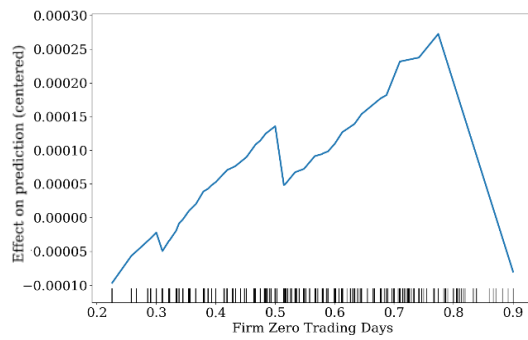
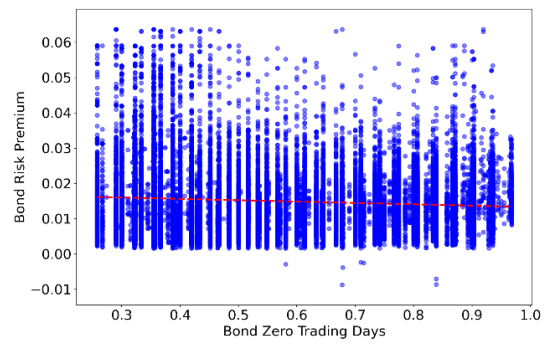
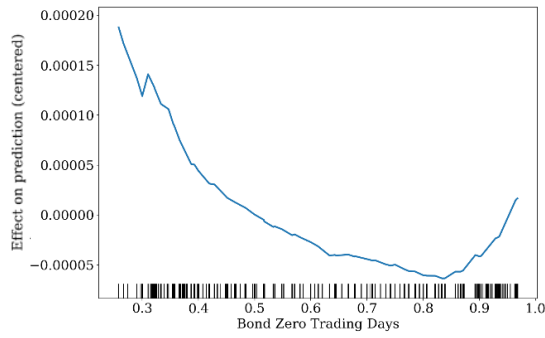
Figure 3.8 ALE Plots and Regressions of Continuous Variables





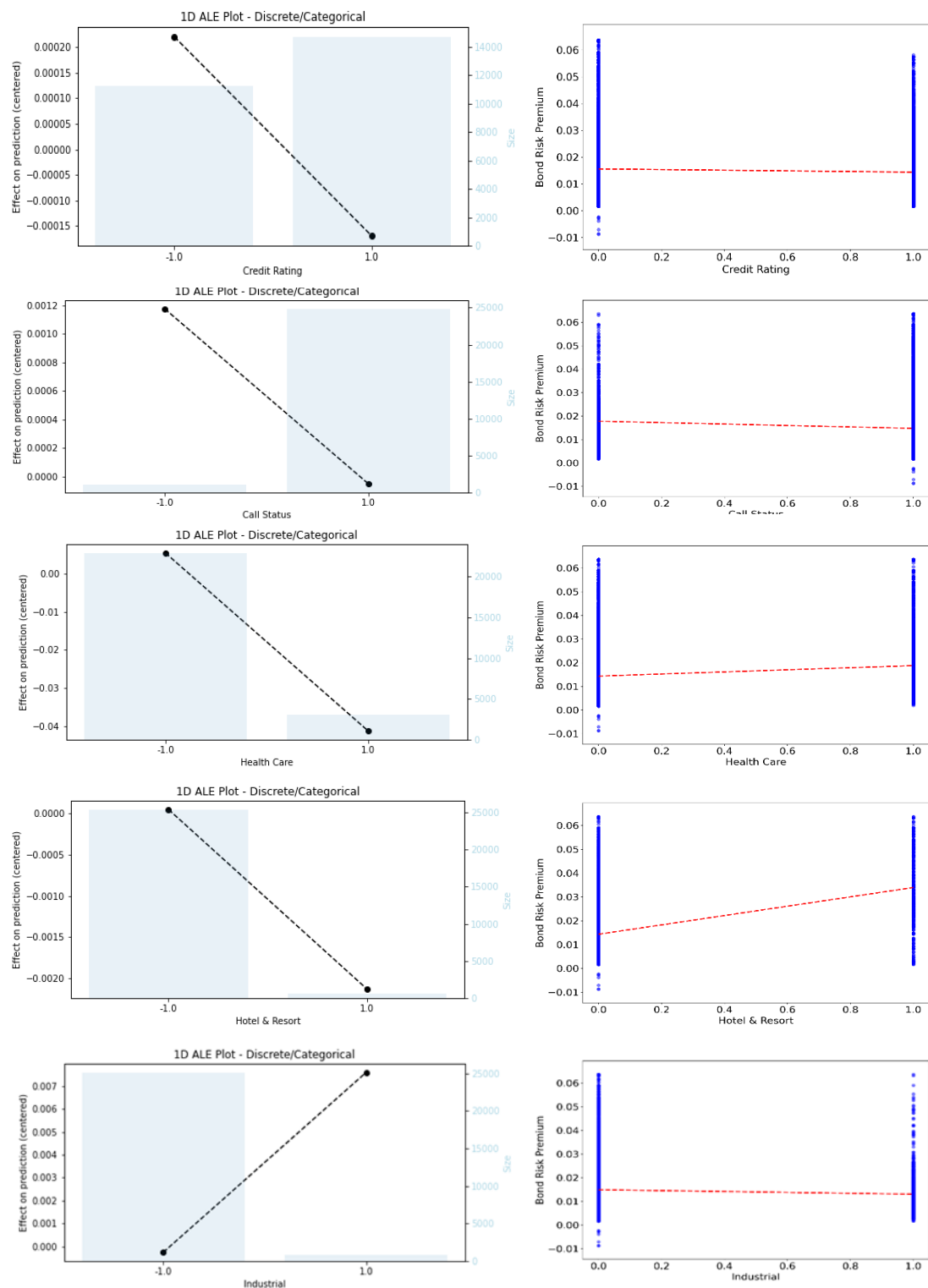
Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt



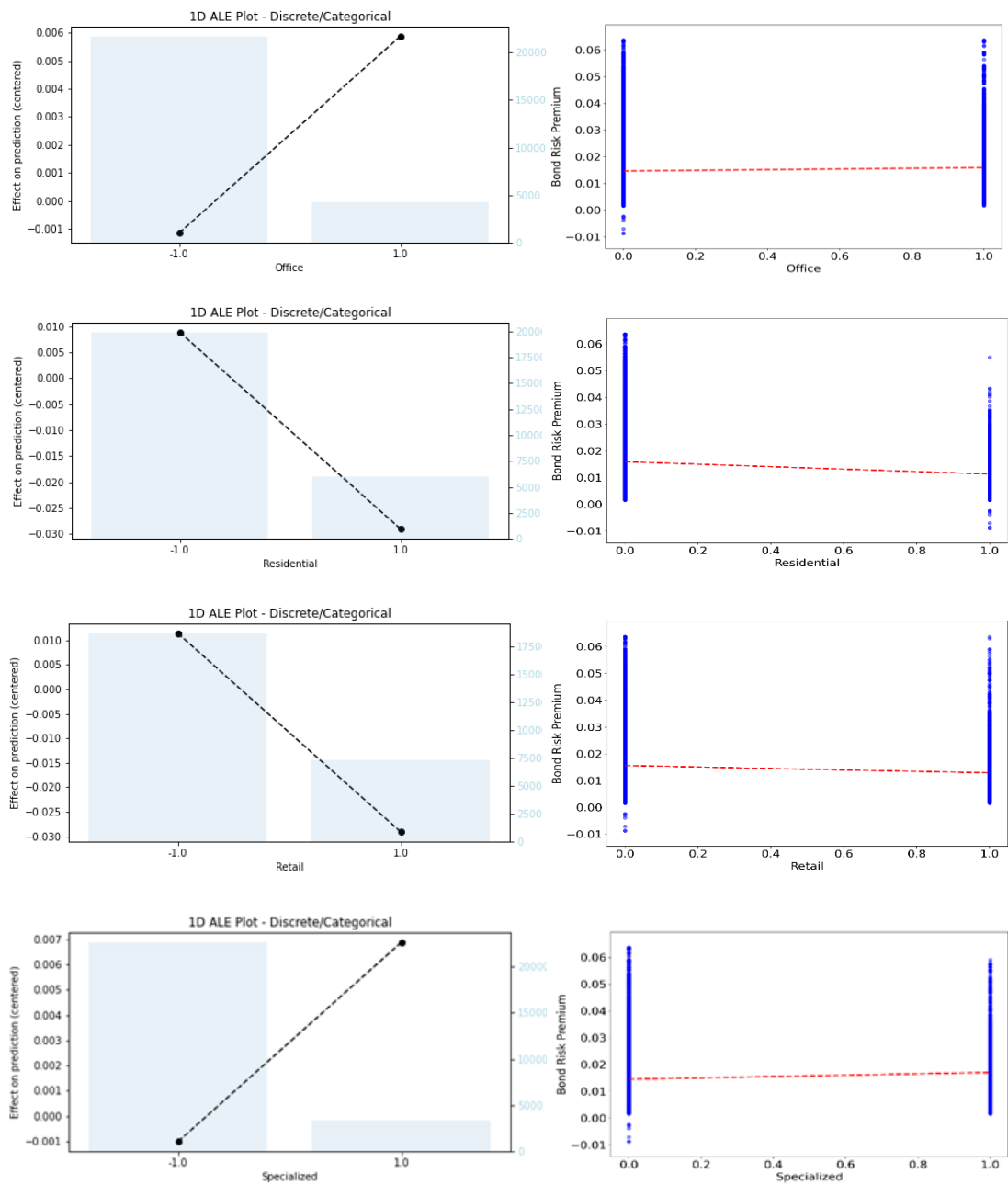


3.8.4 Appendix IV – ALE Plots and Regressions of Categorical Variables

Figure 3.9 ALE Plots and Regressions of Categorical Variables



Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt



3.8.5 Appendix V – Variable Transformation to Optimize Polynomial OLS

Table 3.9 Variable Transformations

Variable	Transformation
$\beta_{i,t}^{REIT}$	Squared ⁽²⁾
Volume	–
Bond Price	Squared ⁽²⁾
Coupon Rate	Squared ⁽²⁾
Volume Weighted Bond Yield	Squared ⁽²⁾
Time to Maturity	Squared ⁽²⁾
HHI	Cubed ⁽³⁾
Market Cap	Squared ⁽²⁾
Equity Return	Cubed ⁽³⁾
Leverage	Squared ⁽²⁾
Return on Assets	Cubed ⁽³⁾
Market-to-Book	Cubed ⁽³⁾
Default Risk Spread	Squared ⁽²⁾
IVOL	Cubed ⁽³⁾
Bond Zero Trading Days	Squared ⁽²⁾
Firm Zero Trading Days	Power of ⁽⁴⁾
Bid-Ask Spread	–

Notes: This table illustrates the transformations of each variable of the OLS regression based on the shape of the respective ALE plot. Dummy variables are not transformed due to their linear nature. The variables Dividends, FFO, VIX, S&P 500, Unemployment Rate, Term Spread, and GDP were excluded from the OLS for multicollinearity reasons (VIF values exceeded 10), therefore, were also not transformed.

3.9 References

- Apley, D. W. and Zhu, J. (2020).** Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 82(4), 1059–1086.
- Benefield, J. D., Anderson, R. I. and Zumpano, L. V. (2009).** Performance differences in property-type diversified versus specialized real estate investment trusts (REITs). *Review of Financial Economics*, 18(2), 70–79.
- Bessembinder, H., Kahle, K. M., Maxwell, W. F. and Xu, D. (2009).** Measuring abnormal bond performance. *The Review of Financial Studies*, 22(10), 4219–4258.
- Capozza, D. R. and Seguin, P. J. (1999).** Focus, Transparency and Value: The REIT Evidence. *Real Estate Economics*, 27(4), 587–619.
- Cheok, S. M. C., Sing, T. F. and Tsai, I.-C. (2011).** Diversification as a Value-Adding Strategy for Asian REITs: A Myth or Reality? *International Real Estate Review*, 14(2), 184–207.
- Demirci, I., Eichholtz, P. and Yoender, E. (2020).** Corporate Diversification and the Cost of Debt. *The Journal of Real Estate Finance and Economics*, 61(3), 316–368.
- Feng, Z., Pattanapanchai, M., Price, S. M. and Sirmans, C. F. (2021).** Geographic diversification in real estate investment trusts. *Real Estate Economics*, 49(1), 267–286.
- Fisher, G., Steiner, E., Titman, S. and Viswanathan, A. (2022).** Location density, systematic risk, and cap rates: Evidence from REITs. *Real Estate Economics*, 50(2), 366–400.
- Freybote, J. (2016).** Real estate sentiment as information for REIT bond pricing. *Journal of Property Research*, 33(1), 18–36.
- Friedman, J. H. (2001).** Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Gospodinov, N. and Robotti, C. (2021).** Common pricing across asset classes: Empirical evidence revisited. *Journal of Financial Economics*, 140(1), 292–324.
- Gu, S., Kelly, B. and Xiu, D. (2020).** Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273.
- Haran, M. E., Lo, D., McCord, M., Davis, P. and Lim, L. C. (2020).** Impact of firm-level attributes on listed real estate company performance. *Journal of Property Investment & Finance*, 39(4), 323–348.
- Hartzell, J. C., Sun, L. and Titman, S. (2014).** Institutional investors as monitors of corporate diversification decisions: Evidence from real estate investment trusts. *Journal of Corporate Finance*, 25, 61–72.

- Huerta, D. and Mothorpe, C. (2024).** The Impact of Property Clustering on REIT Operational Efficiency and Firm Value. *The Journal of Real Estate Finance and Economics*, 71(1), 476–509.
- Ibrahim, I. and Falkenbach, H. (2024).** Diversification and Cost of Public Debt for REITs: Evidence from the US. *Journal of Real Estate Portfolio Management*, 30(1), 36–53.
- Kraemer, B., Nagl, C., Stang, M. and Schaefers, W. (2023a).** Explainable AI in a Real Estate Context – Exploring the Determinants of Residential Real Estate Values. *Journal of Housing Research*, 32(2), 204–245.
- Kraemer, B., Stang, M., Doskoč, V., Schaefers, W. and Friedrich, T. (2023b).** Automated valuation models: Improving model performance by choosing the optimal spatial training level. *Journal of Property Research*, 40(4), 365–390.
- Liu, Y. and Wu, J. C. (2021).** Reconstructing the yield curve. *Journal of Financial Economics*, 142(3), 1395–1425.
- Longstaff, F. A., Mithal, S. and Neis, E. (2005).** Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance*, 60(5), 2213–2253.
- Loo, W. K. (2020).** Predictability of HK-REITs returns using artificial neural network. *Journal of Property Investment & Finance*, 38(4), 291–307.
- Lorenz, F., Willwersch, J., Cajias, M. and Fuerst, F. (2023).** Interpretable machine learning for real estate market analysis. *Real Estate Economics*, 51(5), 1178–1208.
- Lui, A. (2021).** Does size matter? Incorporating company size in credit risk assessment. S&P Global Market Intelligence. <https://www.spglobal.com/marketintelligence/en/news-insights/blog/does-size-matter-incorporating-company-size-in-credit-risk-assessment>.
- Lundberg, S. and Lee, S.-I. (2017).** A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*.
- Markowitz, H. (1952).** Portfolio Selection. *The Journal of Finance*, 7(1), 77–91.
- McCracken, M. W. and Ng, S. (2016).** FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574–589.
- Merton, R. C. (1974).** On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2), 449–470.
- Ro, S. and Ziobrowski, A. J. (2011).** Does focus really matter? Specialized vs. diversified REITs. *The Journal of Real Estate Finance and Economics*, 42(1), 68–83.
- Serrano, C. and Hoesli, M. (2007).** Forecasting EREIT Returns. *Journal of Real Estate Portfolio Management*, 13(4), 293–310.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56), 1929–1958.

Tuzel, S. and Zhang, M. B. (2017). Local Risk, Local Factors, and Asset Prices. *The Journal of Finance*, 72(1), 325–370.

Wang, Y., Li, Y., Song, Y. and Rong, X. (2020). The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition. *Applied Sciences*, 10(5), 1897.

Zhu, B. and Lizieri, C. (2022). Local Beta: Has Local Real Estate Market Risk Been Priced in REIT Returns? *The Journal of Real Estate Finance and Economics* 69(1), 682–718.

4 The Non-Linear Dynamics of Disclosure Tone and REIT Bond Yields

4.1 Abstract

The tone in corporate disclosures represents a source of information for bond investors that complements quantitative information, allows to predict a firm's default risk and subsequently bond yields. However, previous studies investigating the predictive value of disclosure tone for bond yields assume a linear relationship. We argue that linearity cannot be assumed and hypothesize two dimensions of non-linearity in the tone-yield relationship. In our empirical investigation, we employ a multi-method approach combining machine learning methodologies and conventional regression analysis. Using a sample of 68 bond-issuing equity REITs over the period of 2011 to 2024, we provide evidence that the relationship of disclosure tone and bond yield spread is non-linear. Net positive tone, which jointly accounts for positive and negative disclosure tone in REIT financial statements, has a quadratic relationship with future REIT bond yield spread. The more positive or negative disclosure tone is, the higher are future bond yield spreads. Additionally, net positive tone explained by REIT fundamentals negatively predicts bond yield spread in the next month while non-fundamental tone positively predicts it. Our findings suggest that the disclosure tone-bond yield relationship is more complex than assumed by previous studies.

Keywords: Corporate Bonds, REITs, Corporate Disclosures, Textual Tone, Machine Learning

4.2 Introduction

When assessing the credit risk of a firm, bond market investors are faced with informational asymmetries between themselves and firms issuing bonds, which can explain bond yields (Han and Zhou, 2014; Duffie and Lando, 2001). One issue is that accounting information provided by firms to bond investors and credit analysts to assess default risk is incomplete and infrequent (Duffie and Lando, 2001). As a consequence, an emerging literature investigates the informative value of soft or qualitative information in corporate disclosures, as a complement to quantitative information, for bond market participants.

Previous research has shown that not only what information a firm's management discloses (Gupta et al., 2025; Lou, 2019), but also how information is disclosed matters for bond market participants. In particular, the tone in corporate disclosures has been found to predict the default risk of firms (Jin et al., 2024) and bond investor behavior (Cao et al., 2025). This suggests that the tone in corporate disclosures signals risk-related information to bond investors and represents a source of information improving the information environment around a firm. Previous studies assume a linear relationship between disclosure tone and bond yields (Schmeling and Wagner, 2025; Hirsch et al., 2024; Wang, 2020). However, the question arises as to whether the relationship is indeed linear for the following reasons:

Corporate management has been found to prefer a positive tone over a neutral or negative tone in disclosures (Brockman et al., 2015; Price et al., 2012; Schleicher and Walker, 2010; Rutherford, 2005; Hyland, 1998; Hildebrandt and Snyder, 1981) and manage disclosure tone (Davis and Tama-Sweet, 2012). The appeal of this positive framing is that it may yield a positive stock market reaction (Davis et al., 2012; Henry, 2008), influence investor perception (Huang et al., 2014) and offset the effects of poor earnings (Price et al., 2012). An overly positive tone in disclosures, however, has been found to negatively predict financial performance (Jiang et al., 2019; Huang et al., 2014) and positively predicts litigation risk (Rogers et al., 2011). Furthermore, institutional stock market investors have been found to penalize firms using an overly optimistic disclosure tone (Baganski et al., 2018; Blau et al., 2015; Tan et al., 2014).

Considering their focus on assessing a firm's default risk and the limited upside potential of fixed income securities, bond investors, who are predominantly institutional investors, have been found to have a greater demand for negative information (DeFranco et al., 2009) and are more sensitive to the negative or pessimistic tone in corporate disclosures (Jin et al., 2024). Consequently, bond investors may appreciate a mildly negative tone in financial statements as it indicates a managerial dedication to transparency and higher

disclosure quality, which in turn is expected to lead to lower bond yields. This expectation is in line with previous findings for the stock market, where a change to a more pessimistic disclosure language has been found to predict lower cost of equity (Iatridis, 2016). However, as negative tone in financial statements increases, bond investors may consider it as a signal of an increased default risk, which is expected to result in higher bond yields. Similarly, an overly optimistic tone may signal risk-related information to bond investors with regards to a firm's litigation risk, future performance, and managerial disclosure practices.

Furthermore, previous studies use aggregated tone measures, which ignores that tone can be either based on firm-fundamentals or not. An optimistic or pessimistic tone that is explained by firm fundamentals may signal information to bond investors about a firm's dedication to transparency and disclosure quality and reduce their required yield. An overly negative or positive disclosure tone not explained by fundamentals may signal information about a firm to bond investors that increases their risk perception and required yield. Furthermore, the informative value of disclosure tone based on, amongst others, a firm's leverage, operating performance, or growth opportunities to bond investors may be limited considering the availability of quantitative information capturing this information. Tone that is not based on firm-fundamentals, however, may capture new information for bond investors and thus have more informative value.

The purpose of this study is to investigate these two dimensions of non-linearity in the relationship of disclosure tone and bond yield spread. We focus on net positive tone, which jointly accounts for positive and negative disclosure tone. With regards to the first dimension of non-linearity, we hypothesize a quadratic tone-bond yield relationship. In particular, we expect a mildly negative or neutral tone to reduce the bond yield spread, and an overly positive or negative tone to increase it. With regards to the second dimension of non-linearity, we expect the fundamentals-based net positive tone to reduce bond yields, and the non-fundamental tone to increase it. Additionally, we hypothesize non-fundamental tone to have more predictive value for the bond yield than fundamentals-based tone.

For our empirical analysis, we follow Kozak et al. (2025) and use publicly traded equity real estate investment trusts (REITs) that issue bonds as a laboratory for the following reasons: First, equity REITs represent companies that focus on the investment in, development, and management of income-producing real estate. The property portfolios owned by REITs generate more stable cash flows compared to firms focused on products and services as well as allow bond investors to assess the credit risk of REITs more easily. Thus, the

homogeneity of equity REITs in terms of business model allows eliminating the effects of, for example, differing levels of cash flow uncertainty, fixed assets or business focus among firms in an industry on bond yields. Secondly, while covenants are uncommon in the investment grade bond market for non-REIT firms, covenants relating to, amongst others, secured and overall debt ratios, unencumbered assets, and debt service coverage ratios are standard for all REIT bonds (see, e.g., Riddiough and Steiner, 2018; Deng et al., 2016). Riddiough and Steiner (2018) provide evidence for the value of bond covenants to alleviate agency conflicts between REIT management and shareholders. The presence of standardized bond covenants across the REIT industry allows us to mitigate any effects on bond yields that stem from variations in bond covenants or their absence across firms as well as agency conflicts. Thirdly, REITs are limited in their ability to retain earnings due to their high dividend payout requirement, which in turn results in a higher reliance on external financing than non-REIT firms (Ott et al., 2005), a higher transparency in terms of disclosure (Danielsen et al., 2014) and higher monitoring by external sources. Additionally, equity REITs have a comparatively higher institutional ownership (see, e.g., Devos et al., 2012; Ciochetti et al., 2002), which enhances corporate governance (Hartzell et al., 2014; Chung et al. 2012; Feng et al., 2010; Hartzell et al., 2006) and the informativeness of disclosure tone (Jain et al., 2023). Thus, the institutional characteristics of equity REITs ensure a certain industry-level baseline of disclosure quality that is less likely present in other industries.

We use a multi-method approach combining machine learning and conventional regression approaches in our analysis. First, we employ a large language model approach (RoBERTa) to extract tone measures from the financial statements (10-Q and 10-K) for 68 equity REITs over the period of 2011 to 2024. In particular, we extract quarterly positive (optimistic) and negative (pessimistic) measures to calculate a net positive tone measure for each REIT in our sample. Next, we use a two-way fixed effects regression to analyze the first dimension of non-linearity. To investigate the presence of a quadratic term in the tone-yield relationship, we employ an eXtreme Gradient Boosting (XGBoost) algorithm in combination with Accumulated Local Effects (ALE) plots. To assess the second dimension of non-linearity, we disaggregate our tone measures by regressing them on firm fundamentals and retaining fitted values (fundamentals-based tone) and residuals (non-fundamental tone). We then use disaggregated tone measures in our two-way fixed effects regression.

We find the tone-yield relationship to be non-linear across the two dimensions investigated: First, we provide evidence that net positive tone has a quadratic relationship with the REIT bond yield spread in the next month. Second, fundamental and non-

fundamental net positive tone predict future REIT bond yield spread in different directions. While fundamentals-based net positive tone negatively predicts the REIT bond yield, the non-fundamental net positive tone positively predicts it. Furthermore, we show that the non-fundamental component of disclosure tone has a higher informative value to bond investors than the fundamentals-based component.

Our study contributes to the literature in several ways: We contribute to a larger finance and accounting literature that investigates the informative value of tone in corporate disclosures for corporate and sovereign bond markets (Cao et al., 2025; Schmeling and Wagner, 2025; Hirsch et al., 2024; Jin et al., 2023; Liang et al., 2023; Wang, 2020) by showing that the tone-bond yield relationship is not linear. To our knowledge this is the first study to investigate the assumption of linearity made by previous studies by 1) separating tone into a fundamental and non-fundamental component and 2) accounting for a quadratic relationship.

Additionally, we contribute to the REIT bond literature, which compared to REIT stocks, is relatively small and has focused on factors such as commercial real estate market sentiment (Freybote, 2016), environmental performance (Eichholtz et al., 2019), ownership structure (Gilstrap et al., 2021), debt covenants (Deng et al., 2016), debt capital structure decisions (Allen and Letdin, 2020), and real estate-specific factors (Kozak et al., 2025) in the context of REIT bonds.

Last, our study contributes to an emerging literature on textual analysis in the context of commercial real estate (see, e.g., Carstens and Freybote, 2021; Beracha et al., 2019; Hausler et al., 2018), REITs (see, e.g., Paulus et al., 2023; Kim et al., 2021; Hausler et al., 2018) and housing markets (see, e.g., Aroul et al., 2025; Shen and Ross, 2021; Nowak et al., 2019; Nowak and Smith, 2017). To our knowledge, this is the first study that employs textual analysis in the context of REIT bond markets.

Our findings have implications for REIT management and bond investors. Our results suggest that the tone in REIT financial statements can be used as an additional source of information by bond investors to predict yields, irrespective of whether it is based on fundamentals or not. Our findings for non-fundamental tone suggests that REIT management should pay particular attention to align their disclosure tone with firm fundamentals. Furthermore, the quadratic relationship identified for net positive tone suggests that overly positive or negative tone in disclosures should be avoided by management.

The remainder of this study is structured as follows. Next, we review relevant studies on corporate disclosures and bond markets to develop our theoretical framework and

hypotheses. It is followed by a discussion of our data, tone measure derivation and methodology. Then, we present our results and a conclusion.

4.3 Literature Review and Hypothesis Development

4.3.1 Disclosure and Bond Markets

An emerging literature investigates the informative value of qualitative or soft information in disclosures for sovereign and corporate bond markets. Some studies focus on what information is revealed in textual disclosures such as the use of derivatives (Gupta et al., 2025) or pending lawsuits (Lou, 2019). For example, Bozanic et al. (2023) use textual analysis to investigate how credit analysts use qualitative information in corporate financial statements to assess credit risk. They find evidence that credit analysts account for qualitative (soft) information in their credit ratings. In particular, they make soft adjustments when companies provide more disclosure that reflect increases in credit risk (e.g., credit risk and debt) or decreases in credit risk (profitability and managerial outlook). The authors conclude that textual sections in annual reports provide complementary information to the private information credit analysts already have. Other studies focus on how information is disclosed, i.e., readability and textual sentiment/tone. With regards to our study, this latter stream of literature is most relevant.

Several studies have investigated readability in corporate disclosures as a dimension of how firms provide bond market participants with information: Chen and Tseng (2021) analyze the readability of the narrative “Notes to Consolidated Financial Statements” and its impact on corporate bond yields. They show that a lower readability predicts greater corporate bond yield spreads, particularly for firms in certain industries (high-tech) and with a higher equity volatility. Bonsall and Miller (2017) show that corporate disclosures with a lower readability are associated with a lower credit rating, higher disagreement between bond rating agencies and cost of debt. Chen et al. (2024) further investigate the impact of disclosure readability on corporate bond yields. The authors focus on narrative pension disclosures by firms and its association with credit risk. They find supportive evidence that a lower readability predicts higher bond yield spreads. Farrell et al. (2024) focus on the municipal bond market and show that the complexity of disclosures impact cost of debt for municipalities. It is noteworthy that the authors find the positive relationship for complexity and borrowing costs to be particularly strong for municipalities with a poorer information environment due to less complementary information from rating agencies.

Another stream of literature focuses on the impact of tone in news or disclosures by firms or governmental agencies on bond markets. Consoli et al. (2021) investigate emotions in

macroeconomic news and show that they have predictive power for sovereign bond yield spreads. In particular, they find that negative emotions in the news can predict yield spreads in periods of distress. Hirsch et al. (2024) provide evidence that the tonality of television news impacts sovereign debt spreads. Hereby, positive (negative) news predict a lower (higher) spread.

Wang (2020) uses textual analysis to analyze the impact of the tone in risk disclosures in financial statements on Credit Default Swaps (CDS). The author finds a change in CDS spreads that is positively associated with the pessimistic tone in risk disclosures. Tone hereby is distinct from the amount of disclosures. However, the market reaction to disclosure tone is stronger for entities closer to default, i.e., a higher concern of downside risk, and with weaker information environment. Overall, the author concludes that the tone in MD&A risk disclosures has informative value for debt and CDS market investors. Liang et al. (2023) show that a more pessimistic managerial tone in financial statements predicts the hedging behavior of investors in the CDS market.

Jin et al. (2024) investigate the informative value of the tone, besides length and content, of risk disclosures in financial statement for a firm's default risk. They provide evidence for the predictive power of tone for a firm's default risk, particularly in the post-GFC period. Precisely, negative tone positively predicts default risk. The relationship is stronger for firms that, for example, are undergoing shifts in default risk, face financial distress or have a heightened scrutiny from external monitors. The authors conclude that the tone in risk disclosures in annual report have informative value for bond investors to assess default risk. Cao et al. (2025) investigate whether insurance companies, which are the largest investor type in investment-grade corporate bonds, use the tone of earnings calls as a source of information in their trading of BBB-rated bonds. The authors find that negatively toned earnings calls have predictive power for the selling behavior of insurance companies. Schmeling and Wagner (2025) find that the tone used by the Central Bank in its communications can predict equity and capital market outcomes. They find positive tone to predict, amongst others, lower credit risk spreads. Rich et al. (2021) investigate disclosure tone in the context of municipal bonds and find the negative tone in the MD&A section to positively predict cost of debt and disagreements between bond rating agencies.

4.3.2 Hypotheses Development

While some studies investigating disclosure tone in the context of bond markets find positive and negative tone to have predictive value (Schmeling and Wagner, 2025; Hirsch et al., 2024), previous studies on bond markets in general and corporate bonds in particular emphasizes the importance of negative or pessimistic tone for bond market participants

(Cao et al., 2025; Jin et al., 2024; Liang et al., 2023; Consoli et al., 2021; Wang, 2020). The higher sensitivity to negative tone in corporate disclosures and a greater demand for negative information (De Franco et al., 2009) results from the limited upside potential of corporate bonds and need of bond investors to assess risks to predict a firm's default risk.

The focus of bond investors on negative or pessimistic tone differs from the focus of corporate management, which has been found to prefer a positive disclosure tone over a neutral or negative (Brockman et al., 2015; Price et al., 2012; Schleicher and Walker, 2010; Rutherford, 2005; Hyland, 1998; Hildebrandt and Snyder, 1981). The managerial positivity bias results from the ability of positive framing to influence the reaction and perception of stock market investors (see, e.g., Huang et al., 2014; Davis et al., 2012; Henry, 2008).

However, institutional stock market investors have been found to penalize firms for an overly optimistic tone in disclosures (Baganski et al., 2018; Blau et al., 2015; Tan et al., 2014). These sophisticated stock market investors can identify when the tone in corporate disclosures fails to match the quantitative information about a firm's performance and firms use tone management. Another explanation is that an overly optimistic tone signals information about firm risks as it has been found to negatively predict a firm's financial performance (Jiang et al., 2019; Huang et al., 2014) as well as positively predict litigation risk (Roger et al., 2011). Therefore, Iatridis (2016) shows that a change in reporting language to a more pessimistic or unexpectedly pessimistic language predicts lower cost of equity.

Considering the findings of previous studies on institutional investors, who are the predominant type of investor in the bond market (DeFranco et al., 2019) and disclosure tone, we expect the negative tone in REIT financial statements to have a higher predictive value for bond yield spreads than the positive tone. However, previous studies investigating tone and bond yield or CDS spreads (Schmeling and Wagner, 2025; Hirsch et al., 2024; Wang, 2020) assume a linear relationship. We argue that two dimensions of non-linearity exist in the relationship of tone, defined as net positive tone, and bond yield spread:

First, considering that institutional investors are aware of the managerial positivity bias, they may consider a mildly negative or neutral tone in REIT disclosures a signal of a firm's commitment to a high level of transparency and absence of tone management. Consequently, they may require a lower yield, which is in line with the findings of Iatridis (2016) for the stock market. An increase in negative or positive disclosure tone, however, is likely to signal an increase in risks to bond investors and increases their yield

requirements. Thus, we expect a quadratic relationship between net positive tone and REIT bond yield spread.

Second, the tone in REIT financial statements has two components: The first component can be explained with firm fundamentals such as operating performance, firm size, leverage, stock market performance, or growth opportunities. The second component cannot be explained with firm fundamentals, but other factors such as managerial sentiment or private information not captured by firm fundamentals. Previous studies on disclosure tone in the context of bond markets (Cao et al., 2025; Schmeling and Wagner, 2025; Hirsch et al., 2024; Jin et al., 2024; Liang et al., 2023; Wang, 2020) fail to make this distinction in their analysis.

Fundamentals-based net positive tone is expected to reduce bond investor risk perceptions and negatively predict the bond yield spread, while a net positive tone that is not explained by firm fundamentals may signal information to bond investors about managerial disclosure practices (e.g., tone management, transparency) and increase their risk perception. This expectation is in line with previous studies on institutional investors in the stock market, who penalize firms for an overly positive tone in their disclosures (Baganski et al., 2018; Blau et al., 2015; Tan et al., 2014). Therefore, we expect a positive relationship of non-fundamental tone and REIT bond yield spread.

Furthermore, non-fundamental tone may represent new information and thus have a higher informative value for REIT bond investors as it is not captured by fundamental variables, for which quantitative data commonly exists. Therefore, we hypothesize that tone not explained by fundamental variables has a higher predictive value for REIT bond yields than the fundamentals-based negative tone.

4.4 Data

4.4.1 Derivation of Tone Measures

We obtain the quarterly financial statements (10-Q and 10-K) for publicly traded equity REITs over the period of 2011 to 2024. The beginning of our sample period is due to limited REIT bond transaction data prior to 2011 (Freybote, 2016). Considering not all equity REITs access the bond market, we only focus on equity REITs that issued publicly traded bonds over the sample period. Of the 174 REITs that were NAREIT index constituents at any time between 2011 and 2024, 89 REITs issued bonds. Due to missing financial statements for some REITs, our final sample consists of 68 REITs.

We focus on the Management and Discussion Analysis (MD&A) section of financial statements in line with previous studies (see, e.g., Wang, 2020; Koelbl, 2020) to extract our tone measures. In this section, REITs are required to discuss their financial condition, results of operations, liquidity, and known trends or uncertainties that could impact future performance. As a robustness check, we also obtain our tone measures based on all textual sections in financial reports.

We derive our tone measures using the RoBERTa (Robustly Optimized BERT Pretraining Approach) model by Liu et al. (2019), an improved version of BERT (Bidirectional Encoder Representations from Transformers) introduced by Devlin et al. (2019). These models are Large Language Models and represent a state-of-the-art approach in natural language processing (Karanikola et al., 2023). The use of the RoBERTa model minimizes the need for an adequate training dataset, as the model comes pre-trained and is only optimized for an individual case (Liu et al., 2019). In our case, the model is optimized using the SEntFIN 1.0 dataset by Sinha et al. (2022).

The preprocessing of financial statements is crucial, as the model requires clean tokens to accurately classify tone. Therefore, we employed techniques commonly used in the literature, such as segmenting the report into smaller paragraphs, converting all words to lowercase, and removing numbers and punctuation. The final pre-processing step is lemmatization, where words are transformed into their root forms to reduce the complexity and size of the vocabulary of the textual data (Paulus et al., 2024).

Our approach to using RoBERTa for tone classification results in a probabilistic output with values ranging from 0 to 1, where values close to 0 indicate a low probability of a message being positive and vice versa. To achieve a three-way classification, we deploy a decision rule in line with Paulus et al. (2024) given as:

$$P \geq 0.667 \rightarrow \text{positive tone}$$

$$0.333 < P < 0.667 \rightarrow \text{neutral tone}$$

$$P \leq 0.333 \rightarrow \text{negative tone}$$

After classifying the tone of the segments in the financial reports, we aggregate it monthly to construct our three tone measures. The first measure, the Optimism Indicator (OI), is based on Antweiler and Frank (2004) and captures the positive disclosure tone as shown in Equation 4.1:

$$OI_t = \frac{\sum_1^I \text{positive segments}_{i,t}}{\sum \text{total number of segments}_t}, \quad (4.1)$$

where i denotes the number of segments of a financial report and t denotes the quarter the report was published.

The second measure is called Pessimism Indicator (PI), captures negative disclosure tone and measures the number of negative segments in relation to the total number of segments as shown in Equation 4.2:

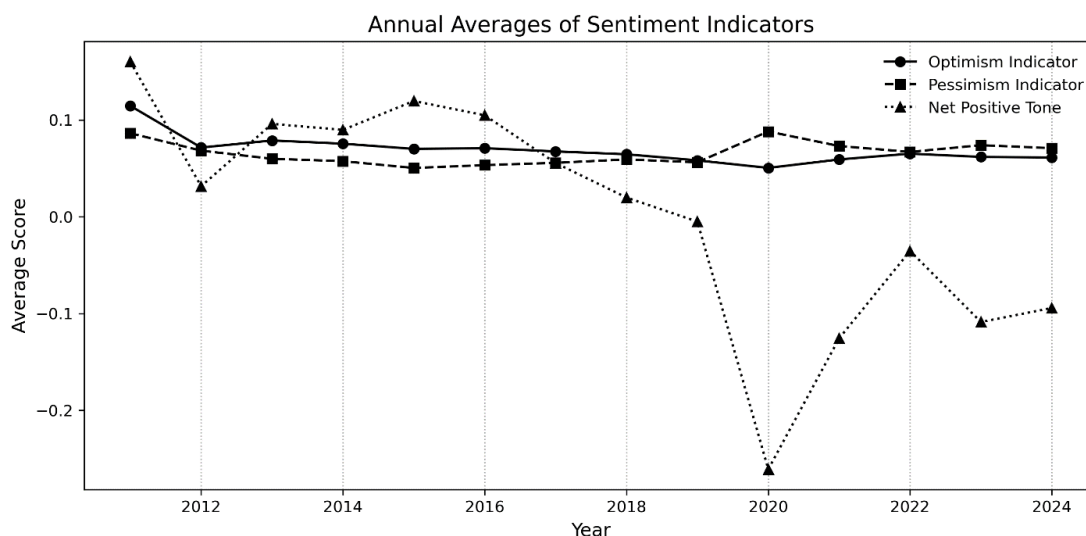
$$PI_t = \frac{\sum_1^I \text{negative segments}_{i,t}}{\sum \text{total number of segments}_t}. \quad (4.2)$$

To derive a tone measure that jointly accounts for positive and negative tone, we calculate the Net Positive Tone (NPT) in line with Doran et al. (2012) as shown in Equation 4.3. NPT ranges from -1 to 1, where values below 0 indicate a tendency toward a generally negative tone and values above 0 indicate a tendency toward a more positive overall tone.

$$NPT_t = \frac{(\sum_1^I \text{positive segments}_{i,t} - \sum_1^I \text{negative segments}_{i,t})}{(\sum_1^I \text{positive segments}_{i,t} + \sum_1^I \text{negative segments}_{i,t})}. \quad (4.3)$$

Please note, our tone measures are in levels as opposed to changes to account for the fact that bond market investors are expected to respond to the disclosure tone of the most recent financial statement as opposed to the change in disclosure tone from the previous quarter. Kravet and Muslu (2013) use changes in risk disclosure tone to account for omitted variables and reverse causality resulting from repetitions in risk disclosures by firms over multiple quarters. However, compared to risk factors discussed by firms, which are more likely to remain constant over a certain amount of time, the content of MD&A sections varies from quarter to quarter as it reflects the most recent operations, financial conditions, and expectations for future performance. Furthermore, we address endogeneity concerns due to omitted variables or reverse causality by using a fixed effects regression and lagged tone variables in our empirical analysis.

Figure 4.1 shows our tone measures, averaged per year across REITs in our sample, for the period of 2011 to 2024. The OI is above the PI for all periods until 2020, which can be explained by factors such as favorable economic conditions and/or a positivity bias in financial statements (Schleicher and Walker, 2010). Since 2020, the PI exceeded the OI due to events such as COVID-19 pandemic and shift interest rate regime. The Figure emphasizes the importance of jointly assessing positive and negative tone in form of the net positive tone (NPT) measure and motivates our focus on this tone measure. While the average positive and negative scores appear to be relatively leveled over time, the NPT shows considerable variability: Since 2015, the average annual net positive tone has been decreasing, i.e., became less positive, and in 2019 fell below 0, which means it became negative. For the remainder of our sample period, it remained in the negative range.

Figure 4.1 Positive, Negative, and Net Positive Tone Measures over Time

4.4.2 Dependent Variable and Control Variables

Our dependent variable is the spread of REIT bond yields over U.S. Treasuries of the same maturity. Data on REIT bond yields are sourced from the Trade Reporting and Compliance Engine (TRACE) database by FINRA, which has required broker-dealers to self-report their over-the-counter bond transactions since 2002.⁶ We calculate monthly yield spreads by averaging the volume-weighted yields of all transactions in a month of a given REIT bond and then subtracting the yield on U.S. Treasuries with corresponding maturities. We obtain U.S. Treasury data from the Liu and Wu (2021) dataset. Our final sample consists of 31,812 monthly bond yield spreads of U.S. equity REITs that were constituents of the FTSE NAREIT Index at some point between January 2011 and December 2024.⁷

While our tone measures are our independent variables of interest, we control for several variables that impact bond yields: First, we control for the features of REIT bond issues. Since TRACE does not offer bond characteristics such as the bond maturity date, call features, or coupon rate, this data was manually collected from the FINRA website, which is based on the TRACE database. In particular, we control for bond features such as the coupon rate, bond price, transaction volume, time to maturity, call option and credit rating of the issuing REIT. Bond transactions with missing data are excluded. Additionally, bonds with maturities of less than one month and those with maturities of more than 360 months are excluded, as no Treasury matching was feasible. Based on Bessembinder et al. (2009),

⁶ Since bonds are traded over-the-counter (OTC), collecting data on bond transactions is limited compared to equity data. TRACE is the largest database for researching bond transactions in the US. However, its nature of self-reporting can lead to false entries. By preprocessing and winsorizing our data, we try to ensure that such errors do not distort the predictions.

⁷ To prevent survivorship bias, the analysis includes all REITs that issued bonds during the entire period, regardless of whether they were present for the entire duration of the sample or not.

TRACE data entries marked as canceled are removed from the dataset, while corrected entries are kept.

Furthermore, we control for variables capturing the liquidity of REIT bonds by including the percentage of days in the month that the bond was not traded (Bond Zero Trading Days), the percentage of days in a month on which no outstanding bond a REIT was traded (Firm Zero Trading Days) and the monthly mean of the difference in the average daily bid and ask prices of each bond (Bid-Ask Spread).

In addition to bond issue and market variables, we control for REIT characteristics: We control for REIT size measured as the log of market capitalization, stock market performance measured by a REIT's stock market return, operating performance captured by Funds from Operations (FFO), leverage defined as total debt divided by total assets, dividends per share, and growth opportunities as measured by the market to book ratio. We control for the property type specialization of REITs using binary variables with diversified REITs being the reference group. Monthly REIT market and quarterly REIT accounting data are obtained from Compustat and Refinitiv Eikon.

Last, we control for stock market, bond market and macro-economic conditions that impact bond yields: We control for stock market conditions by including the return on the S&P 500 index as well as the CBOE volatility index (VIX). Bond market conditions are controlled for with the term structure defined as spread between the 10-year U.S. Treasury and 3-month Treasury Bill as well as the credit risk defined as the spread between corporate bonds rated Baa and Aaa by Moody's. Macroeconomic conditions are controlled for with the interest rate volatility defined as the standard deviation of the monthly 3-month constant maturity Treasury rate over a 12-month period prior to a bond's transaction date, the unemployment rate, and gross domestic product (GDP). Macroeconomic data are from the Federal Reserve Economic (FRED) monthly database, as presented by McCracken and Ng (2016). Table 4.1 present an overview and description for our variables.

Table 4.1 Variable Description

Variable	Description
Dependent Variable	
Bond Yield Spread	Difference between the volume-weighted yield to maturity of the REIT bond and the maturity-matched Treasury yield to maturity
Tone Variables	
Optimism Indicator	Number of positive segments in a 10-Q/K report divided by total segments
Pessimism Indicator	Number of negative segments in a 10-Q/K report divided by total segments
Net Positive Tone	Number of positive minus negative segments in a 10-Q/K report divided by the total number of positive and negative segments
Bond Variables	
Coupon Rate	Coupon rate on bond in percent
Bond Price	Bond price inclusive of any commission, mark-ups, and/or mark-downs
Transaction volume	Volume reported on the trade
Time to Maturity	Difference in months between maturity date and transaction date
Credit Rating Dummy	1 if investment grade bond, 0 if high yield bond
Call Status Dummy	1 if callable, 0 if not
Liquidity Variables	
Bond Zero Trading Days	Percentage of days in the month that the bond was not traded
Firm Zero Trading Days	Percentage of days in a month on which no outstanding bond of a REIT was traded
Bid-Ask Spread	Monthly mean of the difference in the average daily bid and ask prices of each bond
REIT-Specific Variables	
Property Type Dummy	8 REIT sectors based on property portfolio, defined by NAREIT, reference category is Diversified
Market Capitalization	Measure of company size, expressed as total dollar market value of a REIT's outstanding shares of stock (in million USD), as log
Return	Expressed as the current adjusted price of the REIT stock minus the original adjusted price of the stock (month t-1) divided by the original adjusted price
Funds From Operations	Cash flow from operations
Leverage	Total debt outstanding divided by total assets
Dividends per Share	Distribution of a REIT's earnings per share to its shareholders
Market-to-Book	Market capitalization divided by total assets minus total outstanding debt
Equity Market Variables	
S&P 500	Return of the S&P 500 composite index
VIX	CBOE volatility index
Macroeconomic Variables	
Term Spread	10-Year Treasury rate minus 3-Month Treasury Bill
Default Risk Spread	Moody's Baa corporate bond yield minus Moody's Aaa corporate bond yield
Interest Rate Volatility	Standard deviation of monthly 3-month constant maturity Treasury rate over a 12-month period prior to bond transaction date
Unemployment Rate	Unemployment rate in percent
GDP	Real gross domestic product (in billions of Chained 2017 USD)

Various steps are undertaken to preprocess the combined dataset. We handle missing values by imputing data for up to one year, meaning that an existing value is perpetuated until a new value is recorded or 52 weeks have passed. The TRACE bond data includes manually reported information from broker-dealers, which can be prone to input errors, potentially leading to inaccurate results. Therefore, the yield spreads and REIT-level accounting data are winsorized at the 5% and 95% percentiles to prevent outliers from distorting the results (Gu et al., 2020). This process is conducted on a monthly basis to ensure that future data points or REITs not included in the sample at a specific time are not incorporated. Table 4.2 presents the summary statistics for our variables. It is noteworthy that the mean and standard deviations for *OI* and *PI* are relatively similar, however, *NPT* is on average negative and exhibits a high variability. This is in line with Figure 4.1. Table 4.3 presents the pair-wise correlations for variables in our dataset. The *OI* and *NPT* are negatively correlated with bond yield spreads, whereas the *PI* is positively correlated.

Table 4.2 Descriptive Statistics

Variable	Unit	Mean	SD	Min	Max
Bond Yield Spread	Percentage	0.0144	0.0090	-0.0365	0.0587
Optimism Indicator (OI)	Integer	0.0635	0.0393	0	0.2551
Pessimism Indicator (PI)	Integer	0.0653	0.0354	0	0.2474
Net Positive Tone (NPT)	Integer	-0.0385	0.3406	-1	1
Coupon Rate	Percentage	0.0388	0.0102	0.0170	0.0775
Bond Price	U.S.-Dollar	98.85	10.23	64.14	127.79
Transaction volume	U.S.-Dollar	420,592	391,948	10,083	2,989,666
Time to Maturity	Months	92.53	79.61	2	359
Credit Rating Dummy	Percentage	0.6580	0.4744	0	1
Call Status Dummy	Percentage	0.9274	0.2594	0	1
Bond Zero Trading Days	Percentage	0.5871	0.2070	0.2581	0.9677
Firm Zero Trading Days	Percentage	0.3493	0.0836	0.2333	0.9000
Bid-Ask Spread	Percentage	0.0083	0.0009	0.0006	0.0710
Market Capitalization	Mil \$ (log)	9.4392	0.9954	6.7476	11.7325
Return	Percentage	0.0038	0.0776	-0.5274	0.5940
Funds From Operations	Integer	129.72	151.00	-792.30	839.70
Leverage	Percentage	0.5420	0.1230	0.3615	0.9304
Dividends per Share	U.S.-Dollar	0.7963	0.6032	0	3
Market-to-Book	Integer	2.6007	1.8808	0.4246	10.3505
S&P 500	Integer	3,662	984.3	1,173	5,621
VIX	Integer	19.13	6.89	10.08	58.08
Term Spread	Percentage	-0.0166	0.0307	-0.0662	0.0304
Default Risk Spread	Percentage	0.0092	0.0021	0.0055	0.017
Interest Rate Volatility	Percentage	0.0040	0.0043	0.0001	0.0157
Unemployment Rate	Percentage	0.0467	0.0191	0.034	0.148
GDP	Integer	21,168	1,328	17,031	23,400

Notes: This table presents the summary statistics for a sample of 68 equity REITs over the period of 2011 to 2024. Variable definitions are in Table 4.1.

Table 4.3 Pairwise Correlations

Variable	OI	PI	NPT
OI	1.000***	0.287***	0.683***
PI	0.287***	1.000***	-0.423***
Bond Variables			
Bond Yield Spread	-0.024***	0.013***	-0.043***
Coupon Rate	0.009	0.014**	-0.016***
Bond Price	-0.01*	0.027***	-0.012**
Transaction volume	0.039***	-0.033***	0.075***
Time to Maturity	0.048***	-0.018***	0.083***
Liquidity Variables			
Bond Zero Trading Days	0.047***	0.106***	0.012**
Firm Zero Trading Days	0.028***	0.015***	0.020***
Bid-Ask Spread	0.015***	0.049***	0.010*
REIT-Specific Variables			
Market Capitalization	0.177***	-0.125***	0.260***
Return	-0.023***	0.055***	-0.060***
Funds From Operations	0.096***	-0.173***	0.211***
Leverage	0.230***	0.055***	0.060***
Dividends per Share	0.213***	-0.006	0.146***
Market-to-Book	0.376***	0.081***	0.233***
Equity Market Variables			
S&P 500	-0.040***	0.108***	-0.127***
VIX	-0.001	0.097***	-0.046***
Macroeconomic Variables			
Term Spread	0.000	0.023***	-0.005
Default Risk Spread	0.021***	-0.022***	0.059***
Interest Rate Volatility	0.022***	0.038***	0.003
Unemployment Rate	-0.054***	0.104***	-0.093***
GDP	-0.022***	0.071***	-0.085***

Notes: This Table presents the pair-wise correlations between the tone measures and other variables for a sample of 68 equity REITs over the period of 2011 to 2024. Variable definitions are in Table 4.1. '***', '**', and '*' denote significance at the 1%, 5%, and 10% level respectively.

To disaggregate our tone measures, we follow previous studies (see, e.g., Ling et al., 2014 Huang et al., 2014) and use an orthogonalization approach: We first regress each of the tone measures on firm fundamentals (Equity Return, Market Cap, Leverage, Dividends per Share, FFO, Market-to-book) and retain the fitted values. The fitted values represent the component of disclosure tone that can be explained with firm fundamentals ($Tone_{Fundamental}$). The residuals from this regression represent the component of disclosure tone that cannot be explained by REIT fundamentals ($Tone_{Non-Fundamental}$).

4.5 Methodology

To assess the two dimensions of non-linearity in the relationship of disclosure tone and REIT bond yield spread, we employ a combination of machine learning and conventional regression approaches. Previous studies provide evidence for a superior performance of machine learning algorithms over conventional regression methods due to their ability to

address the non-linearity and non-stationary data properties (see, e.g., Jenett et al., 2025; Loo, 2020). However, a crucial limitation of these machine learning approaches is the lack of interpretability of statistical significance. Therefore, we combine both approaches to ensure ideal outputs and an enhancement over an isolated application of either method. In particular, we employ two-way fixed effects regression and XGBoost algorithm in combination with ALE plots.

4.5.1 Fixed Effects Regression

We employ a panel regression with time-fixed and property-type fixed effects to assess the predictive value of aggregated and disaggregated tone measures for REIT bond yield spreads. To address endogeneity concerns and focus on prediction, we use lagged tone variables in our analysis. Control variables are contemporaneous. Our model is shown in Equation 4.4.

$$\begin{aligned}
 r_{i,t}^{REIT} - r_{i,t}^{Treasury} = & c_0 + \beta_1 OI + \beta_2 PI + \beta_3 NPT + \sum_{i=1}^6 \chi_i Bond\ Controls \\
 & + \sum_{i=1}^3 \chi_i Liquidity\ Controls + \sum_{i=1}^7 \chi_i REIT\ Controls \\
 & + \sum_{i=1}^2 \chi_i Equity\ Market\ Controls + \sum_{i=1}^5 \chi_i Macroeconomic\ Controls \\
 & + \gamma_t^{Time} + \gamma_t^{Property\ Type} + \epsilon_{i,t},
 \end{aligned} \tag{4.4}$$

where $r_{i,t}^{REIT} - r_{i,t}^{Treasury}$ is the monthly REIT bond yield spread as the difference between the monthly REIT bond yield of REIT i at time t and the U.S. Treasury yield with the same maturity i at time t . γ_t^{Time} are year-fixed effects, $\gamma_t^{Property\ Type}$ are property type-fixed effects and $\epsilon_{i,t}$ is the residual component (error term). Our analysis using the disaggregated tone measures is based on Equation 4.4, however, REIT fundamentals are accounted for in the fundamentals-based tone measures and are not included in the model.

4.5.2 eXtreme Gradient Boosting (XGBoost)

While fixed effects regression assumes a linear relationship between the independent and dependent variables, we hypothesize that the tone-yield relationship is quadratic. One approach to assess linearity of a relationship is the eXtreme Gradient Boosting (XGBoost) algorithm, which is a tree-based ensemble learning method. The origin of XGBoost can be attributed to Chen and Guestrin (2016), who introduced this algorithm to address the bias-variance trade-off commonly encountered in single decision trees. XGBoost has since demonstrated strong performance in tackling complex issues such as classifications, pattern recognition, and regression-based tasks, often surpassing traditional methods and

alternative machine learning approaches in the field of REITs (see, e.g., Jenett et al., 2025; Habbab et al., 2023).

Tree-based ensemble learning algorithms improve predictive accuracy by aggregating the outcomes of multiple individual decision trees rather than relying on a single one. The basic idea of ensemble learning is that multiple weak learners, in our case, single decision trees, can be combined to form one strong learner. In the case of XGBoost, this relationship can be expressed as:

$$h(y|x) = \sum_{i=1}^M u_m h_m(y|x), \quad (4.5)$$

where h denotes the strong learner, h_m are the weak learners, u_m are their respective weights, M is the number of single trees, x presents the full features space and y is the response variable. In boosting, weak learners h_m are trained sequentially, with every new model focusing on improving the errors made by the previous ones. Gradient boosting extends this principle by iteratively adding trees that minimize the model's loss function using gradient descent. Extreme Gradient Boosting is a computationally efficient and highly scalable version of gradient boosting. Compared with parametric or semi-parametric models, XGBoost is capable of identifying complex, non-linear relationships and higher-order interactions within large datasets. This significantly reduces the need for manual optimization. Moreover, it requires less manual optimization than traditional models.⁸

4.5.3 Accumulated Local Effects (ALE)

However, XGBoost is frequently described as a "black box" model because the internal mechanisms through which it learns the relationships between inputs and outputs are not immediately observable. An emerging branch of research that focuses on making XGBoost and other machine learning methods interpretable is called eXplainable Artificial Intelligence (XAI). To evaluate how a feature influences the model's predictions, one feature is varied while holding everything else constant, and the resulting change in the model's prediction is observed.

The global model-agnostic technique of Accumulated Local Effect (ALE) plots provides valuable insights into how an independent variable affects the model's average prediction of the dependent variable across its entire range of observed values. Alternative approaches, such as Shapley Additive Explanations (SHAP) by Lundberg and Lee (2017) and Partial Dependency Plots (PDP) by Friedman (2001), assume that variables are

⁸ For more information about tree-based methods, ensemble learning, and gradient boosting, see Hastie et al. (2009).

uncorrelated, which is rarely the case in real-world finance and real estate applications. ALE plots, introduced by Apley and Zhu (2020), overcome this limitation by explicitly accounting for correlated features, making them more robust in understanding the effects of features on predictions. Recent real estate studies, including Kraemer et al. (2023) and Lorenz et al. (2023), have highlighted the growing relevance of ALE plots.

To construct ALE plots for our tone measures and dependent variable, the full range of each independent variable $x_r \in \mathbb{R}^{N \times 1}$, is divided into K bins. Let $Z_{r,k}$ denote the $\frac{k}{K}$ quantile of the empirical distribution with $Z_{r,0}$ and $Z_{r,K}$ being the minimum and maximum values, respectively. The set $S_{r,k}$ contains all observations that fall within the left open interval from $Z_{r,k-1}$ to $Z_{r,k}$ with $n_{r,k}$ being the number of observations within that bin. We define $k(x_r)$ as an index that identifies the bin corresponding to a given value of x_r . Consequently, the uncentered accumulated local effect is then given by:

$$g_{ALE}(x_r) = \sum_{k=1}^{k(x_r)} n_{r,k}^{-1} \sum_{i \in S_{r,k}} [f(Z_{r,k}, X_{\setminus r,i}) - f(Z_{r,k-1}, X_{\setminus r,i})], \quad (4.6)$$

where $x_{\setminus r} \in \mathbb{R}^{N \times P-1}$ is the set of features without the feature r of P variables and $f(\cdot)$ the network's prediction formula. The quantity in the squared brackets delineates the prediction of $f(\cdot)$ when observation i is replaced with $Z_{r,k}$ in the minuend and the prediction with $Z_{r,k-1}$ instead of the true value i in the subtrahend. These differences are aggregated for all observations within each bin $S_{r,k}$. Since this computation is repeated across all bins k , $g_{ALE}(x_r)$ represents the sum of these local effects weighted by the number of observations in each bin (Apley and Zhu, 2020). The centered ALE main effect has a zero mean with respect to the marginal distribution of x_r and can be written as:

$$\Theta_{ALE}(x_r) = g_{ALE}(x_r) - N^{-1} \sum_{i=1}^N g_{ALE}(x_{r,i}). \quad (4.7)$$

Centering the ALE functions ensures that the resulting plots are interpretable. Specifically, the ALE plot consistently illustrates the impact of a feature as a function of the average prediction of the neural network across the entire feature space. Examining these ALE plots allows us to detect any non-linear relationships between the independent variables and the predicted bond yield spread.

4.6 Results

4.6.1 First Dimension of Non-Linearity

While our independent variable of interest is net positive tone (*NPT*), we also provide the results for positive or optimistic tone (*OI*) and negative or pessimistic tone (*PI*) to give more context. The results of our fixed effects regression for REIT bond yield spread and the aggregated tone measures are presented in Table 4.4. The negative or pessimistic tone in REIT financial statements (*PI*) positively predicts the bond yield spread in the next month. Thus, a negative disclosure tone predicts higher cost of debt for REITs. The positive and net positive tone in REIT financial statements have a negative relationship with the bond yield spread in the next month, however, coefficients are only significant at the 10% level. Our findings for the aggregated tone measures are in line with previous studies (Cao et al., 2025; Jin et al., 2023; Liang et al., 2023; Consoli et al., 2021; Wang, 2020) that find negative disclosure tone to have the highest informative value for bond investors.

Our results for the control variables are in line with previous studies on, for example, the positive co-movement of REIT bond yields and stock returns (Alexander et al., 2000), implied stock market volatility (Collin-Dufresne et al., 2000), default risk captured by market capitalization (Becker and Ivashina, 2015), liquidity risk (Longstaff et al., 2005) as well as credit rating and time to maturity (Nayak, 2010). Please note, *S&P 500*, *Term Spread*, and *GDP* were excluded from the fixed effects regression due to multicollinearity, i.e., VIF values exceeding the critical value of 10. To ensure that multicollinearity is not affecting standard errors, we also estimate the regressions without control variables while keeping the fixed effects. Our results remain robust.

Table 4.4 Regression Results for Aggregated Tone Measures

	OI	PI	NPT
Tone _{t-1}	-0.0008 (0.0004)*	0.0021 (0.0001)***	-0.0007 (0.0004)*
Volume	-0.0008 (0.0001)***	-0.0008 (0.0001)***	-0.0008 (0.0001)***
Bond Price	-0.0123 (0.0002)***	-0.0123 (0.0002)***	-0.0123 (0.0002)***
Coupon Rate	0.0086 (0.0001)***	0.0086 (0.0001)***	0.0086 (0.0001)***
Time to Maturity	0.0015 (0.0001)***	0.0015 (0.0001)***	0.0015 (0.0001)***
Credit Rating	0.0009 (0.0000)***	0.0009 (0.0000)***	0.0009 (0.0000)***
Call Status	0.0018 (0.0001)***	0.0018 (0.0001)***	0.0017 (0.0001)***
Market Cap	-0.0085 (0.0001)***	-0.0085 (0.0001)***	-0.0088 (0.0001)***
Dividends per Share	0.0004 (0.0001)***	0.0006 (0.0001)***	0.0007 (0.0001)***
FFO	-0.0000 (0.0002)	-0.0003 (0.0003)	-0.0004 (0.0002)*
Equity Return	0.0034 (0.0003)***	0.0034 (0.0003)***	0.0036 (0.0003)***
Leverage	0.0011 (0.0001)***	0.0014 (0.0001)***	0.0012 (0.0001)***
VIX	0.0038 (0.0002)***	0.0039 (0.0002)***	0.0038 (0.0002)***
Market-to-Book	-0.0033 (0.0003)***	-0.0029 (0.0003)***	-0.0030 (0.0003)***
Default Risk Spread	0.0036 (0.0002)***	0.0036 (0.0002)***	0.0035 (0.0002)***
IVOL	-0.0006 (0.0001)***	-0.0005 (0.0001)***	-0.0005 (0.0001)***
Unemployment Rate	0.0018 (0.0002)***	0.0016 (0.0002)***	0.0018 (0.0002)***
Bond Zero Trading Days	-0.0031 (0.0001)***	-0.0030 (0.0001)***	-0.0030 (0.0001)***
Firm Zero Trading Days	-0.0006 (0.0002)***	-0.0005 (0.0002)***	-0.0006 (0.0002)***
Bid-Ask Spread	0.0017 (0.0002)***	0.0016 (0.0002)***	0.0016 (0.0002)***
Time FE	Yes	Yes	Yes
Property Type FE	Yes	Yes	Yes
Adjusted R ²	0.602	0.598	0.602
Observations	31,812	31,812	31,812

Notes: This Table presents the fixed effects regression results for the three tone measures and REIT bond yield spread. REIT bond yield spread are in decimals (e.g., 0.012 = 1.2%). Standard errors are reported in brackets. Variable definitions are in Table 4.1. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Next, we present the results of our investigation into whether the relationships of *OI*, *PI*, and *NPT* with bond yields spreads are linear or not using ALE plots derived from an XGBoost model. The ALE plots as shown in Figure 4.2 visualize how individual features influence the model's predictions beyond simple linear dependencies. Specifically, the ALE plots illustrate how changes in a feature's value impact the predicted bond yield spread, as shown by deviations from the average prediction along the y-axis for different values on the x-axis.

The ALE plots for the aggregative positive and negative tone measures (*OI* and *PI*) reveal non-linear patterns, suggesting that the impact of sentiment on REIT bond yield spreads is marked by reversals and saturation points rather than a consistent linear relationship. In particular, the ALE plot for *OI* shows a general negative relationship with bond yield spreads. However, this relationship flattens out at higher levels of optimism (around *OI*=0.15), indicating a saturation point beyond which further increases in optimism no longer reduce the perceived credit risk. This finding suggests that while moderate optimism signals financial health and reduces borrowing costs, an overly positive disclosure tone may be viewed with suspicion by institutional bond investors, which is in line with previous

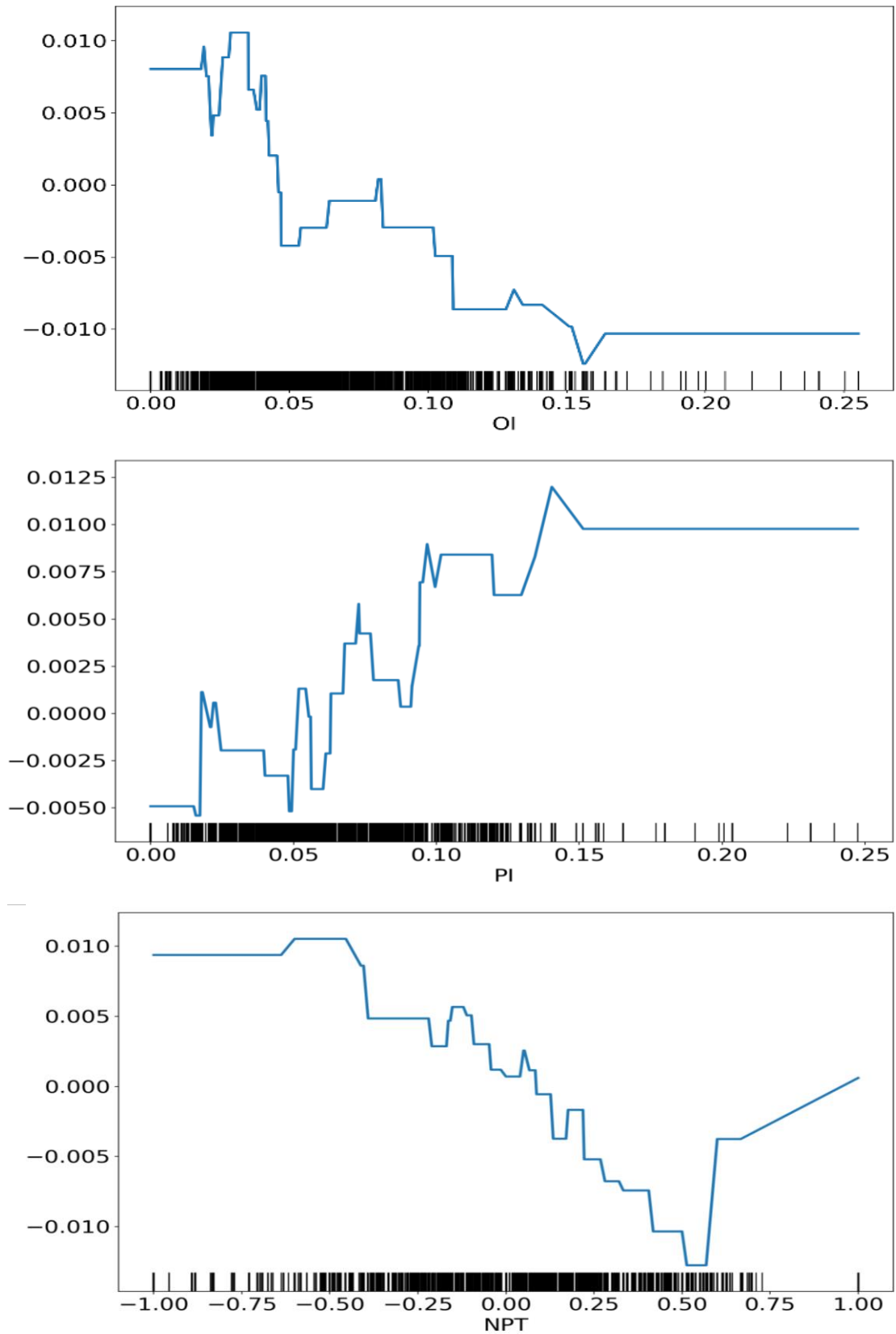
studies on institutional investors in the stock market (Baganski et al., 2018; Blau, Delisle and Price, 2015; Tan, Wang and Zhou, 2014).

The effect of *PI* on bond yield spreads is positive. However, the relationship exhibits even more non-linearity than the *OI*-bond yield relationship. Initially, increases in pessimism sharply raise yield spreads, but this effect flattens beyond a threshold (around $PI = 0.15$). The diminishing marginal impact suggests that once a certain level of negative disclosure tone is reached, additional negativity does not substantially change investor perceptions of risk.

The ALE plot pattern for net positive tone measure (*NPT*) differs noticeably from *OI* and *PI*: At very negative tone levels (*NPT* between -1 and -0.5), the ALE plot shows a flat response in bond yield spreads, suggesting that a highly negative disclosure tone has a limited effect on bond yield spreads. As *NPT* increases from around -0.5 toward moderately positive values, yield spreads decline, consistent with the idea that a somewhat negative to neutral to somewhat positive tone signals lower risk. However, beyond the threshold of $NPT > 0.5$, the trend reverses. An overly positive tone is associated with increasing yield spreads.

While the ALE graphs for *OI* and *PI* suggest a somewhat linear relationship with the bond yield spread, the ALE graph for *NPT* suggests a quadratic relationship in which overly negative and positive disclosure tone increase the bond yield spread while mildly negative to neutral to mildly positive disclosure tone decreases it. Our findings for *NPT* are in line with our expectation that the tone-bond yield relationship is quadratic: A mildly negative/neutral/mildly positive disclosure tone might signal a REIT's commitment to a high level of transparency and absence of tone management, which reduces the risk perception of bond investors and their required yield. However, a higher negative disclosure tone is likely to signal an increasing default risk to bond investors and increases their yield requirements. Similarly, a higher positive disclosure tone may signal higher risks considering that an overly positive disclosure tone has been found to negatively predict firm's financial performance (Jiang et al., 2019; Huang, Teoh and Zhang, 2014) as well as positively predict litigation risk (Rogers, Van Buskirk and Zechman, 2011). Additionally, firms have been found to use a more positive disclosure language if they experience, for example, decreases in operating performance (Schleicher and Walker, 2010). Thus, an overly positive disclosure tone increases the risk perception of REIT bond investors and their required yield.

Figure 4.2 ALE Plots for Tone Measures



Considering our findings for *NPT* based on the ALE plots in Figure 4.2, we estimate our model in Equation 4.4 by including a quadratic term for *NPT*. Our results are presented in Table 4.5 and provide additional evidence for a quadratic relationship: While the coefficient on *NPT* is significantly negative, in line with Table 4.4, the coefficient on *NPT*² is significantly positive. Overall, our multi-methods approach using a combination of machine learning and conventional regression approaches suggest that the relationship of aggregated net positive tone and REIT bond yield spread is non-linear (quadratic) with overly negative and positive tone predicting a higher bond yields spread in the next month while a mildly negative to positive tone predict a lower spread.

Table 4.5 Regression Results for NPT and Bond Yield Spread

	Bond Yield Spread
<i>NPT</i> _{t-1}	-0.0008 (0.0001) ***
<i>NPT</i> ² _{t-1}	0.001 (0.0001)***
Volume	-0.0008 (0.0001)***
Bond Price	-0.0123 (0.0002)***
Coupon Rate	0.009 (0.0001)***
Time to Maturity	0.0015 (<0.0001)***
Credit Rating	0.0009 (<0.0001)***
Call Status	0.0002 (<0.0001)***
Market Cap	-0.0088 (0.0001)***
Dividends per Share	0.0007 (0.0001)***
FFO	-0.0004 (0.0002)*
Equity Return	0.0036 (0.0003)***
Leverage	0.0013 (0.0001)***
VIX	0.0038 (0.0002)***
Market-to-Book	-0.003 (0.0003)***
Default Risk Spread	0.0035 (0.0002)***
IVOL	-0.0005 (0.0001)***
Unemployment Rate	0.0017 (0.0002)***
Bond Zero Trading Days	-0.0031 (<0.0001)***
Firm Zero Trading Days	-0.0007 (0.0002)***
Bid-Ask Spread	0.0016 (0.0002)***
<i>Time FE</i>	Yes
<i>Property Type FE</i>	Yes
<i>Adjusted R</i> ²	0.602
<i>Observations</i>	31,812

Notes: This table presents the results of a two-way fixed effects regression for REIT bond yield spreads. *NPT*² is the quadratic term of the net positive tone measure (*NPT*). Standard errors reported in brackets. Variables are defined in Table 4.1. '***', '**', and '*' denote significance at the 1%, 5%, and 10% level respectively.

4.6.2 Second Dimension of Non-Linearity

Next, we present the results of our analysis the disaggregated tone measures in Table 4.6. In line with our expectations, the fundamental and non-fundamental negative disclosure tone in REIT financial statements (*PI*) positively predict bond yield spread in the next month. Hereby, it is noteworthy that the coefficient on the non-fundamental tone component is

larger than the fundamentals-based tone component. This is in line with our expectation that the information in the fundamentals-based negative tone component is not as novel to investors as it is captured in fundamental variables such as FFO, size, and growth opportunities. The negative tone not explained by firm fundamentals on the other hand has a higher informative value as it provides new information.

The fundamentals-based component of positive tone (*OI*) has no predictive power for bond yield spread while the non-fundamental component positively predicts it. Or put differently, an overly positive tone not explained by REIT fundamentals predicts a higher bond yield spread in the next month. This is in line with our expectation that the risk perception of bond investors increases if a firm uses an overly positive tone and previous studies, which find that institutional investors penalize firms for using an overly positive tone in the stock market (Baganski et al., 2018; Blau, Delisle and Price, 2015; Tan, Wang and Zhou, 2014).

Last, the fundamentals-based component of net positive tone (*NPT*) negatively predicts bond yield spread in the next month while the non-fundamental component positively predicts it. Thus, while an increase in the net positive tone based on REIT fundamentals reduces the bond yield spread in the next month, an increase in non-fundamental net positive tone increases it. The results for *OI* and *PI* in Table 4.6 suggest that the relationship of fundamentals-based net positive tone is driven by fundamental negative tone. On the other hand, the relationship of non-fundamental net positive tone and bond yield spread is driven by non-fundamental positive and negative tone.

Overall, our results in Table 4.6 provide evidence for the presence of a second dimension of non-linearity in the tone-bond yield relationship when positive and negative disclosure tone are jointly accounted for (*NPT*). Additionally, they suggest that non-fundamental disclosure, irrespective of positive, negative or net positive, has a higher predictive value for REIT bond yield spreads than fundamentals-based tone. This is in line with our expectations that non-fundamental tone represents a new source of information to bond investors signaling information about firm risks and improving the information environment around a REIT.

Table 4.6 Regression Results for Disaggregated Tone Measures

	OI	PI	NPT
Fundamental Tone _{t-1}	-0.0050 (0.0039)	0.0071 (0.0025)***	-0.0026 (0.0010)***
Non-Fundamental Tone _{t-1}	0.0017 (0.0006)***	0.0092(0.0034)***	0.0045 (0.0017)***
Volume	-0.0000 (0.0000)***	-0.0000 (0.0000)***	-0.0000 (0.0000)***
Bond Price	-0.0006 (0.0000)***	-0.0005 (0.0000)***	-0.0005 (0.0000)***
Coupon Rate	0.4201 (0.0048)***	0.3594 (0.0048)***	0.3328 (0.0048)***
Time to Maturity	-0.0000 (0.0000)***	0.0000 (0.0000)***	0.0000 (0.0000)***
Credit Rating	0.0011 (0.0001)***	0.0017 (0.0001)***	0.0011 (0.0001)***
Call Status	0.0050 (0.0002)***	0.0037 (0.0002)***	0.0039 (0.0001)***
VIX	0.0001 (0.0000)***	0.0002 (0.0000)***	0.0002 (0.0000)***
Default Risk Spread	0.3966 (0.0399)***	0.5684 (0.0385)***	0.4976 (0.0378)***
IVOL	-0.0720 (0.0144)***	-0.0670 (0.0139)***	-0.0699 (0.0136)***
Unemployment Rate	0.0519 (0.0034)***	0.0334 (0.0032)***	0.0297 (0.0032)***
Bond Zero Trading Days	-0.0075 (0.0002)***	-0.0082 (0.0002)***	-0.0088 (0.0002)***
Firm Zero Trading Days	0.0066 (0.0005)***	-0.0004 (0.0006)***	-0.0003 (0.0004)***
Bid-Ask Spread	0.0966 (0.0049)***	0.0649 (0.0048)***	0.0611 (0.0047)***
Time FE	Yes	Yes	Yes
Property Type FE	Yes	Yes	Yes
Adjusted R ²	0.526	0.563	0.576
Observations	31,812	31,812	31,812

Notes: This Table presents the results of a two-way fixed effects regression of fundamental and non-fundamental tone measures and bond yield spread. Fundamental tone is the fitted value of the regression of the respective tone measure on the REIT fundamentals while non-fundamental tone is the residual. Variable descriptions are in Table 4.1. Standard errors reported in brackets. '***', '**', and '*' denote significance at the 1%, 5%, and 10% level respectively.

4.6.3 Robustness Checks

We conduct additional analyses to assess the robustness of our results. First, we derive our tone measures based on the entire financial statement as opposed to only the MD&A section. This robustness check accounts for the fact that REITs may disclose additional information in other parts of financial statements (e.g., risk factors, legal proceedings, regulatory disclosures). We estimate our model in Equation 4.4 for these alternative tone measures and present the results in Table 4.7. The results are in line with Table 4.4, with the exception of the coefficient of *OI*, which changes from significant at the 10% level to insignificant. This robustness check suggests they are robust to which textual sections in a REIT financial statement the tone measures are derived from.

Table 4.7 Robustness Check – Tone Measures Based on Entire Financial Report

	OI	PI	NPT
Tone _{t-1}	-0.0018 (0.0011)	0.0014 (0.0001)***	-0.0009 (0.0003) **
Volume	-0.0008 (0.0001)***	-0.0009 (0.0000)***	-0.0009 (0.0000) ***
Bond Price	-0.0125 (0.0002)***	-0.0124 (0.0000)***	-0.0124 (0.0000) ***
Coupon Rate	0.0086 (0.0001)***	0.0086 (0.0000)***	0.0085 (0.0000) ***
Time to Maturity	0.0015 (0.0001)***	0.0016 (0.0001)***	0.0016 (0.0001) ***
Credit Rating	0.0010 (0.0000)***	0.0010 (0.0000)***	0.0010 (0.0000) ***
Call Status	0.0017 (0.0001)***	0.0016 (0.0001)***	0.0016 (0.0001) ***
Market Cap	-0.0084 (0.0001)***	-0.0082 (0.0000)***	-0.0084 (0.0000) ***
Dividends per Share	0.0004 (0.0001)***	0.0004 (0.0001)***	0.0005 (0.0001) ***
FFO	-0.0005 (0.0003)*	-0.0003 (0.0002)	-0.0009 (0.0003) **
Equity Return	0.0033 (0.0003)***	0.0031 (0.0002)***	0.0033 (0.0002) ***
Leverage	0.0012 (0.0001)***	0.0014 (0.0001)***	0.0013 (0.0001) ***
VIX	0.0037 (0.0002)***	0.0038 (0.0002)***	0.0038 (0.0002) ***
Market-to-Book	-0.0035 (0.0003)***	-0.0032 (0.0003)***	-0.0032 (0.0003) ***
Default Risk Spread	0.0037 (0.0002)***	0.0038 (0.0002)***	0.0037 (0.0002) ***
IVOL	-0.0006 (0.0001)***	-0.0006 (0.0001)***	-0.0006 (0.0001) ***
Unemployment Rate	0.0018 (0.0002)***	0.0017 (0.0002)***	0.0018 (0.0002) ***
Bond Zero Trading Days	-0.0028 (0.0001)***	-0.0028 (0.0001)***	-0.0028 (0.0001) ***
Firm Zero Trading Days	-0.0007 (0.0002)***	-0.0008 (0.0002)***	-0.0009 (0.0002) ***
Bid-Ask Spread	0.0017 (0.0002)***	0.0017 (0.0002)***	0.0016 (0.0002) ***
Time FE	Yes	Yes	Yes
Property Type FE	Yes	Yes	Yes
Adjusted R ²	0.592	0.590	0.591
Observations	31,812	31,812	31,812

Notes: This Table presents the fixed effects regression results for the three tone measures derived from the entire REIT financial statement and REIT bond yield spread. Standard errors are reported in brackets. Variable definitions are in Table 4.1. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Second, Wang (2020) shows that the response of investors in the CDS market to pessimistic tone in the MD&A section of financial statements is stronger for firms closer to default. Jin et al. (2024) find that the tone in financial statements has particularly strong predictive power for a firm's default risk if the firm is, amongst others, experiencing financial distress. As a consequence, we assess whether a higher default risk of a REIT, as proxied by a higher leverage and lower credit rating, further increases the predictive power of negative tone for future bond yield spread.

In particular, we create interaction effects of the tone measures and 1) leverage and 2) credit rating and estimate our model in Equation 4.4. Our results are presented in Table 4.8. The main effects and interaction term between *PI* and leverage is significantly positive. Thus, a more negative disclosure tone by higher leveraged REITs additionally increases the bond yield spread. This is in line with Jin et al. (2024) and Wang (2020), who find the informative value of disclosure tone to be particularly high for firms with higher financial challenges.

The interaction effect between *PI* and the credit rating dummy is significantly negative indicating that for investment-grade (non-investment grade) bonds, a negative tone in

financial disclosures is associated with a decrease (increase) in bond yield spreads. On the other hand, the main and interaction effects for *OI* and *NPT* are insignificant or only significant at the 10% level. Overall, our results in Table 4.8 suggest that the informative value of negative disclosure tone in REIT financial statements for bond investors increases for REITs with a higher financial risk.

Table 4.8 Robustness Check – Interaction Effects

	OI	PI	NPT
$Tone_{t-1}$	-0.0005 (0.0004)	0.0029 (0.0006)***	-0.0010 (0.0006)*
$Tone_{t-1} * Leverage_{t-1}$	0.0079 (0.0071)	0.0056 (0.0008)***	0.0092 (0.0061)
$Tone_{t-1} * Credit\ Rating$	-0.0008 (0.0004)*	-0.0008 (0.0002)***	-0.0000 (0.0002)
Volume	-0.0009 (0.0001)***	-0.0008 (0.0001)***	-0.0009 (0.0001)***
Bond Price	-0.0123 (0.0002)***	-0.0124 (0.0002)***	-0.0125 (0.0002)***
Coupon Rate	0.0086 (0.0001)***	0.0087 (0.0001)***	0.0087 (0.0001)***
Time to Maturity	0.0015 (0.0001)***	0.0015 (0.0001)***	0.0015 (0.0001)***
Credit Rating	0.0006 (0.0001)***	0.0007 (0.0001)***	0.0009 (0.0000)***
Call Status	0.0016 (0.0001)***	0.0018 (0.0001)***	0.0016 (0.0001)***
Market Cap	-0.0086 (0.0001)***	-0.0085 (0.0001)***	-0.0088 (0.0001)***
Dividends per Share	0.0004 (0.0001)***	0.0007 (0.0001)***	0.0007 (0.0001)***
FFO	-0.0001 (0.0003)	-0.0003 (0.0002)	-0.0005 (0.0002)**
Equity Return	0.0036 (0.0003)***	0.0034 (0.0003)***	0.0037 (0.0002)***
$Leverage_{t-1}$	-0.0006 (0.0002)***	0.0024 (0.0002)***	0.0011 (0.0001)***
VIX	0.0038 (0.0002)***	0.0039 (0.0002)***	0.0038 (0.0002)***
Market-to-Book	-0.0037 (0.0003)***	-0.0026 (0.0003)***	-0.0034 (0.0003)***
Default Risk Spread	0.0036 (0.0002)***	0.0036 (0.0002)***	0.0035 (0.0002)***
IVOL	-0.0006 (0.0001)***	-0.0005 (0.0001)***	-0.0006 (0.0001)***
Unemployment Rate	0.0017 (0.0002)***	0.0016 (0.0002)***	0.0017 (0.0002)***
Bond Zero Trading Days	-0.0031 (0.0001)***	-0.0030 (0.0001)***	-0.0031 (0.0001)***
Firm Zero Trading Days	-0.0007 (0.0002)***	-0.0005 (0.0002)***	-0.0008 (0.0002)***
Bid-Ask Spread	0.0017 (0.0002)***	0.0017 (0.0002)***	0.0017 (0.0002)***
Time FE	Yes	Yes	Yes
Property Type FE	Yes	Yes	Yes
Adjusted R^2	0.604	0.599	0.605
Observations	31,812	31,812	31,812

Notes: This Table presents the fixed effects regression results for the three tone measures, its interaction effects with leverage and credit rating, and REIT bond yield spread. Standard errors are reported in brackets. Variable definitions are in Table 4.1. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

4.7 Conclusion

The qualitative or soft information in corporate disclosures has been found to improve the information environment around firms for bond investors (Cao et al., 2025; Jin et al., 2024; Gupta et al., 2025; Lou, 2019; Roberts and Zhang, 2018). Hereby, not only what a firm discloses has informative value for bond investors, but also how it discloses information in terms of tone or readability. We use the findings of previous studies on the relationship of tone and bond yields (Schmeling and Wagner, 2025; Hirsch et al., 2024; Wang, 2020) as

a starting point and investigate whether the assumption made by these studies about the linearity of this relationship is appropriate.

We hypothesize that two dimensions of non-linearity exist in the tone-bond yield relationship: First, we expect a quadratic relationship of tone, defined as net positive tone, which jointly considers negative and positive tone in corporate disclosures, to have a quadratic relationship with the bond yield spread. While a mildly negative or neutral tone may signal information about a firm's disclosure practices and credibility to investors, which reduces their risk perception and yield requirements, a higher positive or negative tone signals higher risks that impact a firm's default risk and results in higher bond investor risk perception and yield requirements.

Second, we disaggregate tone into a component that can be explained with firm characteristics and a non-fundamental component. We hypothesize that these components of net positive tone impact bond yield spreads differently. While we expect the fundamentals-based tone to have a negative relationship with bond yield spread, we expect the non-fundamental tone to have a positive relationship one. Additionally, the non-fundamental tone is expected to have a higher predictive value for bond yield spread as it contains new information for bond investors.

In our empirical analysis, we use equity REITs as a laboratory and apply a multi-method approach. In particular, we use a large language model (RoBERTa) to extract tone measures from the financial statements of 68 bond-issuing REITs over the period of 2011 to 2024 and employ an XGBoost algorithm in combination with ALE plots to assess the linearity of the tone-bond yield relationship. We also use conventional fixed effects regression to assess the relationships of aggregated and disaggregated tone measures with REIT bond yields.

Our results provide evidence for the two dimensions of non-linearity in the tone-yield relationship. We firstly show that net positive tone has a quadratic relationship with REIT bond yield spread in the next month. We secondly find the fundamental and non-fundamental net positive tone to predict the future REIT bond yield spread in different directions. While fundamentals-based net positive tone negatively predicts the REIT bond yield, the non-fundamental net positive tone positively predicts it. Our results suggest that the non-fundamental component of disclosure tone has a higher informative value to bond investors than the fundamentals-based component. Overall, our results suggest that the relationship of disclosure tone and bond yield is more complex than assumed in previous studies, and future investigations need to account for aspects of non-linearity identified in this study.

Practical implications arise for both REIT managers and bond investors. Managers should avoid overly positive or negative tone and ensure alignment with fundamentals to limit financing costs. For investors, disclosure tone, especially its non-fundamental component, provides useful additional signals for predicting bond yield spreads.

4.8 References

- Alexander, G. J., Edwards, A. K. and Ferri, M. G. (2000).** What Does Nasdaq's High-Yield Bond Market Reveal about Bondholder-Stockholder Conflicts? *Financial Management*, 29(1), 23–39.
- Allen, L. and Letdin, M. (2020).** The Cost of Debt for REITs: The Mortgage Puzzle. *Journal of Real Estate Research*, 42(2), 239–260.
- Antweiler, W. and Frank, M. Z. (2004).** Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, 59(3), 1259–1294.
- Apley, D. W. and Zhu, J. (2020).** Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 82(4), 1059–1086.
- Aroul, R., Carstens, R. and Freybote, J. (2025).** A Home is More than a House – Does the Tone in Listing Texts Matter to Homebuyers? *Journal of Real Estate Research*, forthcoming.
- Baganski, S.P., Demers, E., Kausar, A. and Yingri, J.Y. (2018).** Linguistic Tone and the Small Trade. *Accounting, Organizations and Society*, 68–69(1), 21–37.
- Becker, B. and Ivashina, V. (2015).** Reaching for Yield in the Bond Market. *The Journal of Finance*, 70(5), 1863–1902.
- Beracha, E., Hausler, J. and Lang, M. (2019).** On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination. *Journal of Real Estate Research*, 41(4), 605–637.
- Bessembinder, H., Kahle, K. M., Maxwell, W. F. and Xu, D. (2009).** Measuring abnormal bond performance. *The Review of Financial Studies*, 22(10), 4219–4258.
- Blau, B.M., DeLisle, R.J. and Price, S.M. (2015).** Do Sophisticated Investors Interpret Earnings Conference Call Tone Differently than Investors at Large? Evidence from Short Sales. *Journal of Corporate Finance*, 31(1), 203–219.
- Bonsall, S.B. IV and Miller, B.P. (2017).** The Impact of Narrative Disclosure Readability on Bond Ratings and the Cost of Debt. *Review of Accounting Studies*, 22(1), 608–643.
- Bozanic, Z., Kraft, P. and Tillet, A. (2023).** Qualitative Disclosure and Credit Analysts' Soft Rating Adjustments. *European Accounting Review*, 32(4), 779–807.
- Brockman, P., Li, X. and Price, S.M. (2015).** Differences in Conference Call Tones: Managers vs. Analysts. *Financial Analysts Journal*, 71(4), 24–42.
- Cao, J., Li, G., Wermers, R., Zhan, X. and Zhou, L. (2025).** Do Insurers Listen to Earnings Conference Calls? Evidence from the Corporate Bond Market. Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5232301.

- Carstens, R. and Freybote, J. (2021).** Can the Textual Tone in REIT Financial Statements Improve the Information Environment for Commercial Real Estate Investors? An Investigation. *Journal of Real Estate Research*, 43(3), 335–354.
- Chen, T.-K., Tseng, Y., Lin, R.-C. and Hung, Y.-S. (2024).** Readability of Pension Narrative Disclosures, Pension Regulatory Changes, and Corporate Credit Risk. *European Accounting Review*, forthcoming.
- Chen, T.-K. and Tseng, Y. (2023).** Readability of Notes to Consolidated Financial Statements and Corporate Bond Yield Spread. *European Accounting Review*, 30(1), 83–113.
- Chen, T. and Guestrin, C. (2016).** XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Chung, R., Fung, S. and Hung, S.-Y. K. (2012).** Institutional Investors and Firm Efficiency of Real Estate Investment Trusts. *Journal of Real Estate Finance & Economics*, 45(1), 171–211.
- Ciochetti, B.A., Craft, T.M. and Shilling, J.D. (2002).** Institutional Investors' Preference for REIT Stocks. *Real Estate Economics*, 30(4), 567–593.
- Collin-Dufresne, P., Goldstein, R. S. and Martin, J. S. (2000).** The Determinants of Credit Spread Changes. *SSRN Electronic Journal*.
- Consoli, S., Pezzioli, L.T. and Tosetti, E. (2021).** Emotions in Macroeconomic News and Their Impact on the European Bond Market. *Journal of International Money and Finance*. 118(1), 102472.
- De Franco, G., Vasvari, F.P. and Wittenberg-Moerman, R. (2009).** The Informational Role of Bond Analysts. *Journal of Accounting Research*, 47(5), 1201–1248.
- Danielsen, B.R., Harrison, D.M., Van Ness and R.A., Warr, R.S. (2014).** Liquidity, accounting transparency, and the cost of capital: Evidence from real estate investment trusts. *Journal of Real Estate Research*, 36(2), 221–251.
- Davis, A.K., Piger, J.M. and Sedor, L.M. (2012).** Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary Accounting Research*, 29(3), 845–868.
- Davis, A.K. and Tama-Sweet, I. (2012).** Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. *Contemporary Accounting Research*, 29(3), 804–837.
- Deng, Y., Devos, E., Rahman, S. and Tsang, D. (2016).** The Role of Debt Covenants in the Investment Grade Bond Market - The REIT Experiment. *Journal of Real Estate Finance and Economics*, 52(1), 428–448.

- Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019).** Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North, 4171–4186.
- Devos, E., Ong, S.-E., Spieler, A.C. and Tsang, D. (2013).** REIT Institutional Ownership Dynamics and the Financial Crisis. *Journal of Real Estate Finance and Economics*, 47(1), 266–288.
- Doran, J. S., Peterson, D. R. and Price, S. M. (2012).** Earnings Conference Call Content and Stock Price: The Case of REITs. *The Journal of Real Estate Finance and Economics*, 45(2), 402–434.
- Duffie, D. and Lando, D. (2001).** Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica*, 69(3), 633–664.
- Eichholtz, P., Holtermans, R., Kok, N. and Yoender, E. (2019).** Environmental Performance and the Cost of Debt: Evidence from Commercial Mortgages and REIT Bonds. *Journal of Banking & Finance*, 102(1), 19–32.
- Farrell, M., Murphy, D., Painter, M. and Zhang, G. (2024).** Disclosure Complexity, Regulatory Burden, and the Cost of Capital. Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4507535.
- Feng, Z., Ghosh, C., He, F. and Sirmans, C.F. (2010).** Institutional Monitoring and REIT CEO Compensation. *Journal of Real Estate Finance & Economics*, 40(1), 446–479.
- Freybote, J. (2016).** Real estate sentiment as information for REIT bond pricing. *Journal of Property Research*, 33(1), 18–36.
- Friedman, J. H. (2001).** Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Gilstrap, C., Petkevich, A., Sezer, O. and Teterin, P. (2022).** REIT Debt Pricing and Ownership Structure. *Journal of Real Estate Finance & Economics*, 64(1), 546–589.
- Granger, C. W. J. (1969).** Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424–438.
- Gu, S., Kelly, B. and Xiu, D. (2020).** Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273.
- Gupta, P., Wang, J.J., Zhao, R. and Zhu, L. (2025).** Corporate Bond Market Reaction to Voluntary Disclosure of Firms' Usage of Derivatives. Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5170899.
- Han, S. and Zhou, X. (2014).** Informed Bond Trading, Corporate Yield Spreads, and Corporate Default Prediction. *Management Science*, 60(3), 675–694.

- Habbab, F. Z., Kampouridis, M. and Papastylianou, T. (2023).** Improving REITs Time Series Prediction Using ML and Technical Analysis Indicators. 2023 International Joint Conference on Neural Networks (IJCNN), 1–8.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009).** The Elements of Statistical Learning. Springer New York.
- Hartzell, J.C., Sun, L. and Titman, S. (2006).** The effect of corporate governance on investments: evidence from real estate investment trusts. *Real Estate Economics*. 34 (3), 343–376.
- Hartzell, J.C., Sun, L. and Titman, S. (2014).** Institutional investors as monitors of corporate diversification decisions: Evidence from real estate investment trusts. *Journal of Corporate Finance*. 25, 61–72.
- Hausler, J., Ruscheinsky, J. and Lang, M. (2018).** News-Based Sentiment Analysis in Real Estate: A Machine Learning Approach. *Journal of Property Research*, 35(2), 344–371.
- Henry, E. (2008).** Are Investors Influenced by How Earnings Press Releases are Written? *Journal of Business Communication*, 45(4), 363–407.
- Hildebrandt, H.W. and Snyder, R.D. (1981).** The Pollyanna Hypothesis in Business Writing: Initial Results, Suggestions for Research. *Journal of Business Communication*, 18(1), 5–15.
- Hirsch, P., Feld, L.P., Koehler, E.A. and Thomas, T. (2024).** "Whatever It Takes!" How Tonality of TV-news Affected Government Bond Yield Spreads During the European Debt Crisis. *European Journal of Political Economy*, 82(1), 102511.
- Huang, X., Teoh, S.H. and Zhang, Y. (2014).** Tone Management. *The Accounting Review*, 89(3), 1083–1113.
- Hyland, K. (1998).** Exploring Corporate Rhetoric: Metadisclosure in the CEO's Letter. *Journal of Business Communication*, 35(2), 224–245.
- Iatridis, G.E. (2016).** Financial Reporting Language in Financial Statements: Does Pessimism Restrict the Potential for Managerial Opportunism? *International Review of Financial Analysis*, 45(1), 1–17.
- Jain, A., Manchiraju, H. and Sunder, S.V. (2023).** Institutional Ownership and the Informativeness of Disclosure Tone. *Journal of Business Finance & Accounting*, 50(1), 61–90.
- Jenett, H., Nagl, C., Nagl, M., Price, S. M. and Schaefers, W. (2025).** Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach. *The Journal of Real Estate Finance and Economics*, forthcoming.

- Jiang, F., Lee, J., Martin, X. and Zhou, G. (2019).** Manager Sentiment and Stock Returns, *Journal of Financial Economics*, 132(1), 126–149.
- Jin, S., Lee, H.-W. and Xie, Y.A. (2024).** Risk Disclosures in Annual Reports and Firms' Default Risk. *Journal of Corporate Accounting & Finance*, 35(1), 192–212.
- Karanikola, A., Davrazos, G., Liapis, C. M. and Kotsiantis, S. (2023).** Financial sentiment analysis: Classic methods vs. deep learning models. *Intelligent Decision Technologies*, 17(4), 893–915.
- Kim, D., Lim, D. and Wiley, J.A. (2023).** Narrative Investment-Risk Disclosure & REIT Investment. *Journal of Real Estate Finance and Economics*, 66(1), 542–567.
- Koelbl, M. (2020).** Is the MD&A of US REITs informative? A textual sentiment study. *Journal of Property Investment & Finance*, 38(3), 181–201.
- Kozak, J., Nagl, C., Nagl, M., Beracha, E. and Schaefer, W. (2025).** Does Real Estate Determine REIT Bond Risk Premia? *Journal of Real Estate Finance and Economics*, forthcoming.
- Kraemer, B., Nagl, C., Stang, M. and Schaefer, W. (2023).** Explainable AI in a Real Estate Context – Exploring the Determinants of Residential Real Estate Values. *Journal of Housing Research*, 32(2), 204–245.
- Kravet, T., & Muslu, V. (2013).** Textual Risk Disclosures and Investors' Risk Perceptions. *Review of Accounting Studies*, 18(1), 1088–1122.
- Liang, P., Hu, N., Liu, L. and Zhang, T. (2023).** Managerial Tone and Investors' Hedging Activities: Evidence from Credit Default Swaps. *Accounting & Finance*, 63(1), 3971–3998.
- Ling, D. C., Naranjo, A. and Scheick, B. (2014).** Investor Sentiment, Limits to Arbitrage and Private Market Returns. *Real Estate Economics*, 42(3), 531–577.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V. (2019).** RoBERTa: A Robustly Optimized BERT Pretraining Approach (Version 1). arXiv.
- Liu, Y. and Wu, J. C. (2021).** Reconstructing the yield curve. *Journal of Financial Economics*, 142(3), 1395–1425.
- Longstaff, F. A., Mithal, S. and Neis, E. (2005).** Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *The Journal of Finance*, 60(5), 2213–2253.
- Loo, W. K. (2020).** Predictability of HK-REITs returns using artificial neural network. *Journal of Property Investment & Finance*, 38(4), 291–307.
- Lorenz, F., Willwersch, J., Cajias, M. and Fuerst, F. (2023).** Interpretable machine learning for real estate market analysis. *Real Estate Economics*, 51(5), 1178–1208.

- Lou, Y. (2019).** Disclosure of Pending Lawsuits and Bond Terms. *Management Science*, 65(4), 1926–1947.
- Lundberg, S. and Lee, S.-I. (2017).** A Unified Approach to Interpreting Model Predictions.
- McCracken, M. W. and Ng, S. (2016).** FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, 34(4), 574–589.
- Nayak, S. (2010).** Investor Sentiment and Corporate Bond Yield Spreads. *Review of Behavioural Finance*, 2(2), 59–80.
- Nowak, A.D. and Smith, P. (2017).** Textual Analysis in Real Estate. *Journal of Applied Econometrics*, 32(1), 896–918.
- Nowak, A.D., Price, B.S. and Smith, P.S. (2021).** Real Estate Dictionaries Across Space and Time. *Journal of Real Estate Finance and Economics*, 62(1), 139–163.
- Ott, S.H., Riddiough, T.J. and Yin, H.-C. (2005).** Finance, Investment and Investment Performance: Evidence from the REIT Sector. *Real Estate Economics*, 33(1), 203–235.
- Paulus, N. M., Lautenschlaeger, L. and Schaefers, W. (2024).** Social Media and Real Estate: Do Twitter Users Predict REIT Performance? *Journal of Real Estate Research*, 1–34.
- Price, S.M., Doran, J.S., Peterson, D.R. and Bliss, B.A. (2012).** Earnings Conference Calls and Stock Returns: The Incremental Informativeness of Textual Tone. *Journal of Banking & Finance*, 36(1), 992–1011.
- Rich, K.T., Roberts, B.L. and Zhang, J.X. (2021).** Linguistic Tone of Management Discussion and Analysis Disclosures and the Municipal Debt Market. *Journal of Public Budgeting, Accounting & Financial Management*, 33(4), 427–446.
- Riddiough, T. and Steiner, E. (2020).** Financial Flexibility and Manager-Shareholder Conflict: Evidence from REITs. *Real Estate Economics*, 48(1), 200–239.
- Rogers, J.L., Van Buskirk, A. and Zechman, S.L.C. (2011).** Disclosure Tone and Shareholder Litigation. *The Accounting Review*, 86(6), 2155–2183.
- Rutherford, B.A. (2005).** Genre Analysis of Corporate Annual Report Narratives – A Corpus Linguistics-Based Approach. *Journal of Business Communication*, 42(4), 349–378.
- Schleicher, T. and Walker, M. (2010).** Bias in the tone of forward-looking narratives. *Accounting and Business Research*, 40(4), 371–390.
- Schmeling, M. and Wagner, C. (2025).** Does Central Bank Tone Move Asset Prices? *Journal of Financial and Quantitative Analysis*, 60(1), 36–67.
- Shen, L.S. and Ross, S. (2021).** Information Value of Property Description: A Machine Learning Approach. *Journal of Urban Economics*, 121(1), 103299.

Sinha, A., Kedas, S., Kumar, R. and Malo, P. (2022). SENTFIN 1.0: ENTITY-AWARE sentiment analysis for financial news. *Journal of the Association for Information Science and Technology*, 73(9), 1314–1335.

Tan, H.-T., Wang, E.Y. and Zhou, B. (2014). When the Use of Positive Language Backfires: The Joint Effect of Tone, Readability, and Investor Sophistication on Earnings Judgments. *Journal of Accounting Research*, 52(1), 273–302.

Wang, K. (2020). Is the Tone in Risk Disclosures in MD&As Relevant to Debt Markets? Evidence from the Pricing of Credit Default Swaps. *Contemporary Accounting Research*, 38(2), 1465–1501.

5 Conclusion

5.1 Executive Summary

This section provides a comprehensive summary of the three research papers that comprise the cumulative thesis. It discusses the objectives of each study, the data and methodologies used, as well as the main results and implications for science and practice.

Paper 1: Dynamics of REIT Returns and Volatility: Analyzing Time-Varying Drivers Through an Explainable Machine Learning Approach

Problem and Objective

Understanding the determinants of REIT returns and the underlying volatility is crucial for investors, policymakers, and academics, particularly in times of economic uncertainty. Despite a large existing body of literature, most studies have analyzed REIT returns and volatility separately, often relying on linear, cross-sectional frameworks that may overlook their time-varying, dynamic, and interconnected nature (see, e.g., Letdin et al., 2019; Chui et al., 2003; Clayton and MacKinnon, 2003). This separation creates a fragmented view, limiting the ability to capture how risk and return evolve jointly across various market states and crisis periods. In addition, linear models struggle to address the inherent non-linearities, non-stationarity, and complex interactions typical of financial time series, raising concerns about mis-specified models and underestimated risks (Loo, 2019).

To address these limitations of linear models, the objective of this paper is to simultaneously analyze REIT returns and volatility over time using machine learning algorithms. By combining REIT-specific financials with macroeconomic factors, the study aims to determine how the importance of these drivers shifts across market cycles, with a particular focus on crisis periods. Moreover, the study seeks to identify potential non-linearities between predictors and REIT returns and volatility.

Methodology and Data

The sample for this study comprises 114,806 weekly return and volatility observations of 151 U.S. equity REITs that were constituents of the NAREIT index at any time during the sample period from 1991 to 2022. This long horizon encompasses multiple crisis periods, including the Asian financial crisis, the dot-com bubble, the Global Financial Crisis (GFC),

and the COVID-19 pandemic, all of which were marked by heightened market uncertainty. The study uses REIT-specific market and accounting data (from Eikon) as well as macroeconomic data (from FRED).

To overcome the shortcomings of traditional linear models, this study employs a robust methodology that combines machine learning models with explainable artificial intelligence (XAI). Specifically, the eXtreme Gradient Boosting (XGBoost) algorithm is employed to forecast REIT returns and volatility. This machine learning algorithm builds predictive models by systematically combining weak predictive models (i.e., decision trees) into a strong learner. To enhance transparency and interpretability and to overcome the often-cited “black box” nature of machine learning models, Shapley Additive Explanations (SHAP) and Accumulated Local Effects (ALE) are utilized to decode the internal processings of the predictive model.

Results and their Contribution to Science and Practice

The results reveal that the characteristics influencing REIT returns and volatility display time-varying properties, indicating that different factors dominate various market phases. While macroeconomic indicators played a major role in explaining returns in earlier years, in recent decades, REIT-specific fundamentals have become increasingly influential. Volatility is more evenly driven by both macroeconomic and REIT-specific variables, yet economic conditions were particularly influential during the 2010s. Importantly, crises lead to substantial deviations from long-run patterns for both return and risk, while the correlation between risk and return drivers increases during these periods. From a scientific perspective, the analysis uncovers significant non-linearities that traditional models fail to capture, underscoring the advantage machine learning methods. From a practical perspective, the results enhance transparency, allowing investors to tailor portfolio strategies to shifting market states, while policymakers can better assess the impact of regulation on REITs.

Paper 2: Location Matters: Local Real Estate Market Risk and Geographic Diversification in REIT Public Debt

Problem and Objective

Corporate bond risk premia and their determinants have been widely studied in relation to default risk, tax premium, liquidity risk premium, and equity volatility premium (Driessen, 2005; Huang and Huang, 2012; Bai et al., 2020). However, despite the increasing reliance of REITs on public debt markets, relatively little is known about how real estate-specific factors influence the pricing of REIT bonds. Unlike other firms that issue corporate bonds, REITs derive their cash flows directly from income-generating properties, which makes their credit risk inherently tied to spatial factors. While earlier research has investigated investor sentiment (Freybote, 2016) and institutional ownership (Gilstrap et al., 2022) as determinants of REIT bond pricing, the role of location-based risks remains underexplored. Ibrahim and Falkenbach (2024) address geographic diversification of a REIT's property portfolio within a linear framework, which may not adequately capture the complex and potentially non-linear effects.

On the equity side, Zhu and Lizieri (2022) show that REITs with a high local beta earn higher returns as compensation for location-based risk, whereas evidence on geographic diversification remains inconclusive (Feng et al., 2021; Huerta and Mothorpe, 2024). This raises the question of whether debt markets evaluate these factors differently, given bondholders' higher claim priority and their stronger focus on downside risks.

The objective of this paper is to analyze how two locational factors, namely local real estate market risk and geographic diversification, influence REIT bond risk premia. Bondholders are expected to price local real estate market risk differently from equity holders due to their claim priority and limited upside potential. Accordingly, investors are expected to demand higher compensation for holding bonds issued by REITs with property portfolios in locations that are more sensitive to real estate market shocks. However, this effect is expected to be weaker compared to the one found for stock returns, since bondholders are repaid before equity holders in case of liquidation and therefore might be less sensitive to geographic risks. Geographic diversification is assessed with a focus on potential non-linear effects consistent with the diminishing marginal benefits suggested by Modern Portfolio Theory. Additionally, the study examines how these two dimensions interact to jointly influence the pricing of REIT bonds.

Methodology and Data

The dataset covers 30,186 monthly bond observations of 60 U.S. REITs from 2010 to 2023, along with 27 explanatory variables. These variables include bond information, property location data, accounting data and macroeconomic indicators sourced from TRACE, S&P Capital IQ, Eikon, and FRED. Local real estate market risk is captured through a local beta, following Zhu and Lizieri (2022), which measures each REIT's sensitivity to local property market fluctuations based on the NCREIF Total Return Index. Geographic diversification of a REIT's property portfolio is quantified using the Herfindahl-Hirschman Index (HHI).

The methodological approach combines machine learning with traditional econometric models, driven by the need to capture complex, non-linear relationships while maintaining interpretability and statistical rigor. An artificial neural network (ANN) is employed to model non-linearities and interactions that conventional linear specifications may not capture. To address interpretability, ALE plots are applied, which provide systematic insights into the marginal impacts of predictors on bond risk premia. Building on these insights, a fixed effects regression with polynomial transformations is estimated to validate the patterns identified by the ANN and ground them in an econometric setting.

Results and their Contribution to Science and Practice

The results indicate that both local real estate market risk and geographic diversification significantly influence REIT bond risk premia. First, higher exposure to local market risk is associated with higher bond risk premia, indicating that bondholders demand compensation for REIT portfolios that are more sensitive to regional property market shocks. Second, geographic diversification does not consistently reduce the cost of debt, but rather exhibits a non-linear impact, where the benefits diminish as portfolios become more widely dispersed. Finally, the interaction between local risk and diversification indicates that REITs operating in more volatile markets benefit disproportionately from diversification, as spreading assets across regions reduces bond risk premia.

From a scientific perspective, the paper contributes to the bond pricing literature by incorporating real estate-specific locational factors and by demonstrating the added value of combining machine learning with traditional regression methods to uncover non-linear effects and interactions. From a practical perspective, the impact of geographic allocation decisions on the cost of debt should motivate REIT managers to optimize long-term financing strategies. For investors, the findings indicate that local market risk is a major driver of bond premia, offering a valuable complement to firm-level financials.

Paper 3: The Non-Linear Dynamics of Disclosure Tone and REIT Bond Yields

Problem and Objective

When evaluating corporate credit risk, bond investors face informational asymmetries with issuing firms, which can translate into higher bond yields (Han and Zhou, 2014). A major challenge is that accounting data is backward-looking and disclosed infrequently, leaving gaps in assessing default risk (Duffie and Lando, 2001). These limitations have led to a growing body of research on whether qualitative information can complement financial indicators and enhance credit risk assessment (Lou, 2019). Evidence from the broader finance literature shows that disclosure tone predicts default risk and bond pricing (Jin et al., 2024), suggesting that tone carries important risk signals and improves the information environment around a firm. Yet most studies assume a linear effect (Schmeling and Wagner, 2025; Wang, 2020), potentially missing that both overly negative and overly positive tone may increase risk perception. Bond investors, who are primarily institutional, tend to weigh negative information more heavily, which may explain why a moderately negative tone can lower yields by signaling transparency (Iatridis, 2016). In contrast, excessive pessimism or optimism may instead raise concerns about default risk or the potential for litigation. Moreover, most studies use aggregated tone measures, although tone not tied to fundamentals may provide new information of greater value to bond investors.

Therefore, the objective of this paper is to examine whether the relationship between disclosure tone and REIT bond risk premia is non-linear. The first dimension addresses the possibility of a quadratic relationship, where a mildly negative or neutral tone reduces risk premia, while overly optimistic or pessimistic tone increases them. A second dimension relates to the distinction between fundamentals-based and non-fundamental tone. The former is expected to predict yields negatively, while the latter is expected to positively predict them and provide greater informational value by capturing incremental signals about managerial disclosure practices not contained in existing financial data.

Methodology and Data

The sample for this study consists of 31,812 monthly bond observations from 68 U.S. equity REITs between 2011 and 2024. Disclosure tone is measured using the net positive tone (*NPT*) derived from quarterly 10-Q/K filings. The tone measures are generated through the RoBERTa model (Liu et al., 2019), a state-of-the-art large language model in natural language processing, where *NPT* reflects the balance of positive and negative language.

Alongside the sentiment measures, the analysis incorporates bond, liquidity, accounting, and macroeconomic variables from TRACE, Compustat, CRSP, and FRED. Following Ling et al. (2014), an orthogonalization approach is used to separate tone grounded in firm fundamentals from tone independent of fundamentals.

The methodological framework applies a multi-method design that integrates traditional econometric techniques with machine learning. A fixed effects regression is estimated to provide a linear baseline. To capture non-linearities, an XGBoost model is employed, complemented by ALE plots. Finally, quadratic terms are incorporated into the regression model to validate the non-linear patterns identified by the machine learning approach.

Results and their Contribution to Science and Practice

The results indicate a clear non-linear relationship between disclosure tone and REIT bond yields across both dimensions. First, a quadratic pattern emerges for net positive tone (*NPT*). A mildly negative to neutral tone is associated with lower bond risk premia, while both highly optimistic and highly pessimistic tone predict higher premia. This suggests that bond investors reward moderate transparency but interpret extreme sentiment as a signal of increased risk. Second, a fundamentals-based tone negatively predicts risk premia, while a non-fundamentals-based tone positively predicts them and exhibits stronger predictive power. This indicates that non-fundamental tone captures incremental information not embedded in standard financial indicators. Tone explained by fundamentals can signal managerial dedication to transparency and disclosure quality, thereby reducing required yields. In contrast, non-fundamental tone conveys additional information about managerial sentiment or risk perception that is not captured by quantitative variables and thus carries higher informational value for bond investors.

The study makes three key contributions to science. First, it challenges the prevailing assumption of linearity in the tone-yield relationship by demonstrating that disclosure tone operates in a quadratic manner, thereby enriching the finance and accounting literature on sentiment in corporate disclosures. Second, it provides evidence that non-fundamental tone carries more substantial informational value for bond investors than fundamentals-based tone, extending existing research on disclosure analysis. Third, it applies textual tone analysis to the REIT bond market, advancing the relatively small body of literature on REIT debt and contributing to the broader field of real estate finance. From a practical perspective, the findings highlight that bond investors can utilize disclosure tone, particularly its non-fundamental component, as an additional risk signal. Meanwhile, REIT managers should avoid excessively extreme tone and ensure alignment with fundamentals to minimize financing costs.

5.2 Final Remarks

The listed real estate universe is currently undergoing a profound shift as fluctuating interest and inflation rates trigger tighter regulatory requirements that make capital more selective. At the same time, the rapid growth of big data and the rise of digital tools, especially machine learning algorithms, are offering new opportunities to study, monitor, and ultimately price risk in REIT equity and debt markets. In this dynamic environment, accurately pricing risk in listed real estate has become both more urgent and more complex (Letdin et al., 2019). A key challenge is that REITs occupy a unique position at the interface between physical property markets and financial markets, which makes their analysis distinct from that of other firms. As listed vehicles, they channel property-level cash flows into two distinct layers of the capital structure, equity and debt, each with its own mechanisms of risk transmission and pricing. While the distinction between equity investors (who are exposed to volatility and upside potential) and bondholders (who are focused on downside protection and default risk) has been well established in the general finance literature (Choi et al., 2022), systematic analysis in the REIT sector remains limited.

These gaps call for approaches that move beyond conventional frameworks and exploit the growing availability of large-scale data and advanced methodologies (Breiman, 2001). However, much of the existing finance and real estate research relies on linear regression models, which implicitly assume proportional relationships between predictors and outcomes. This is problematic, as financial time series are inherently non-linear, non-stationary and dynamic (Gu et al., 2020; Bianchi et al., 2021).

Moreover, quantitative indicators alone do not fully capture the forward-looking and qualitative dimensions of risk. Corporate disclosure, management tone, and sentiment embedded in textual data transmit additional signals that shape investor perceptions and financing costs (Lou, 2019). These soft signals are particularly relevant in the presence of persistent information asymmetries, where investors lack the full understanding of a company's situation (Baker and Wurgler, 2007; Feng, 2021). Taken together, these considerations necessitate an integrated approach to studying REIT equity and debt markets that acknowledges their joint and distinct characteristics, accounts for non-linearities, and incorporates both quantitative and qualitative sources of information.

This dissertation addresses these challenges by advancing the methodological and informational toolkit available for studying REIT equity and debt markets. By applying modern machine learning methods in combination with XAI, this dissertation demonstrates how advanced algorithms can uncover dynamics that remain hidden in traditional linear frameworks. At the same time, XAI techniques help to open the "black

box” of machine learning, making it possible to interpret model logic and connect predictive patterns to economic reasoning (Mullainathan and Spiess, 2017).

A central insight is that relationships in REIT markets are not static but time-varying and non-linear, requiring approaches that can capture threshold effects, regime shifts, and interactions. Accounting for these dynamics is crucial for developing investment strategies and assessing risk under various market conditions. At the same time, the thesis shows that combining machine learning with econometric methods enhances predictive power without losing theoretical clarity and interpretability. Moreover, the importance of spatial factors in REIT pricing is demonstrated, as local market conditions and geographic diversification shape how real estate shocks transmit into risk premia. Lastly, the dissertation demonstrates that soft information from corporate disclosures can improve the understanding of risk drivers beyond what is contained in financial and macroeconomic data. Incorporating both quantitative and qualitative sources of information helps to mitigate information asymmetries and reduces reliance on purely backward-looking indicators. Collectively, these findings may reduce information frictions, and enable investors, managers, and policymakers to make more informed decisions.

Nonetheless, several limitations and areas require further research to fully understand REIT equity and debt markets using machine learning. From a methodological perspective, machine learning methods lack the economic intuition and judgment necessary to understand the problems they are solving. Their outputs must therefore be critically assessed and guided by human expertise. XAI tools can enhance transparency and help reveal model logic, but they also come with their own limitations, such as sensitivity to collinearity or dependencies between features, which can bias interpretation (Aas et al., 2021). Future research should therefore continue to explore the combination of machine learning and traditional econometric approaches to ensure that predictive advances translate into meaningful knowledge for both research and practice.

From a market and pricing perspective, future research could examine how risk premia across the capital structure respond under different market cycles. Location-based risk measures could be extended with more granular data on property exposures, climate risks, or tenant structures to capture the channels through which location affects bond and equity investors. Finally, disclosure analysis could benefit from incorporating other textual sources, such as earnings calls, analyst reports, or sustainability disclosures to offer more profound insights into how qualitative signals affect financing costs. While this dissertation cannot provide final answers to the complexities of REITs, it marks a step forward in understanding their dynamics and serves as a catalyst for future research in this area.

5.3 References

- Aas, K., Jullum, M. and Loland, A. (2021).** Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artificial Intelligence*, 298(1), 103502.
- Bai, J., Goldstein, R. S. and Yan, F. (2020).** Is the credit spread puzzle a myth? *Journal of Financial Economics*, 137(2), 297–319.
- Baker, M. and Wurgler, J. (2007).** Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, 21(2), 129–151.
- Bianchi, D., Buechner, M. and Tamoni, A. (2021).** Bond Risk Premiums with Machine Learning. *The Review of Financial Studies*, 34(2), 1046–1089.
- Breiman, L. (2001).** Statistical Modeling: The Two Cultures. *Statistical Science*, 16(3), 199–231.
- Choi, J., Richardson, M. and Whitelaw, R. F. (2022).** Capital Structure Priority Effects in Durations, Stock-Bond Comovements, and Factor Pricing Models. *The Review of Asset Pricing Studies*, 12(3), 706–753.
- Chui, A. C. W., Titman, S. and Wei, K. C. J. (2003).** The Cross Section of Expected REIT Returns. *Real Estate Economics*, 31(3), 451–479.
- Clayton, J. and MacKinnon, G. (2003).** The Relative Importance of Stock, Bond and Real Estate Factors in Explaining REIT Returns. *Journal of Real Estate Finance and Economics*, 27(1), 39–60.
- Driessen, J. (2005).** Is Default Event Risk Priced in Corporate Bonds? *Review of Financial Studies*, 18(1), 165–195.
- Duffie, D. and Lando, D. (2001).** Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica*, 69(3), 633–664.
- Feng, Z. (2021).** How Does Information Asymmetry Affect REIT Investments? Cost of Capital, Performance, and Executive Compensation. *Journal of Real Estate Portfolio Management*, 27(1), 1–19.
- Feng, Z., Pattanapanchai, M., Price, S. M. and Sirmans, C. F. (2021).** Geographic diversification in real estate investment trusts. *Real Estate Economics*, 49(1), 267–286.
- Freybote, J. (2016).** Real estate sentiment as information for REIT bond pricing. *Journal of Property Research*, 33(1), 18–36.
- Gilstrap, C., Petkevich, A., Sezer, O. and Teterin, P. (2022).** REIT Debt Pricing and Ownership Structure. *Journal of Real Estate Finance and Economics*, 64(4), 546–589.
- Gu, S., Kelly, B. and Xiu, D. (2020).** Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273.

- Han, S. and Zhou, X. (2014).** Informed Bond Trading, Corporate Yield Spreads, and Corporate Default Prediction. *Management Science*, 60(3), 675–694.
- Huang, J. and Huang, M. (2012).** How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk? *Review of Asset Pricing Studies*, 2(2), 153–202.
- Huerta, D. and Mothorpe, C. (2024).** The Impact of Property Clustering on REIT Operational Efficiency and Firm Value. *Journal of Real Estate Finance and Economics*, 71(1), 476–509.
- Iatridis, G. E. (2016).** Financial Reporting Language in Financial Statements: Does Pessimism Restrict the Potential for Managerial Opportunism? *International Review of Financial Analysis*, 45(1), 1–17.
- Ibrahim, I. and Falkenbach, H. (2024).** Diversification and Cost of Public Debt for REITs: Evidence from the US. *Journal of Real Estate Portfolio Management*, 30(1), 36–53.
- Jin, S., Lee, H. W. and Xie, Y. A. (2024).** Risk disclosures in annual reports and firms' default risk. *Journal of Corporate Accounting & Finance*, 35(3), 192–212.
- Letdin, M., Sirmans, C. S., Sirmans, S. S. and Zietz, E. N. (2019).** Explaining REIT Returns. *Journal of Real Estate Literature*, 27(1), 1–25.
- Ling, D. C., Naranjo, A. and Scheick, B. (2014).** Investor Sentiment, Limits to Arbitrage and Private Market Returns. *Real Estate Economics*, 42(3), 531–577.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V. (2019).** RoBERTa: A Robustly Optimized BERT Pretrained Approach. *arXiv preprint*.
- Loo, W. K. (2019).** Predictability of HK-REITs returns using artificial neural network. *Journal of Property Investment & Finance*, 38(4), 291–307.
- Lou, Y. (2019).** Disclosure of Pending Lawsuits and Bond Terms. *Management Science*, 65(4), 1926–1947.
- Mullainathan, S. and Spiess, J. (2017).** Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, 31(2), 87–106.
- Schmeling, M. and Wagner, C. (2025).** Does Central Bank Tone Move Asset Prices? *Journal of Financial and Quantitative Analysis*, 60(1), 371–390.
- Wang, K. (2020).** Is the Tone in Risk Disclosures in MD&As Relevant to Debt Markets? Evidence from the Pricing of Credit Default Swaps. *Contemporary Accounting Research*, 38(2), 1465–1501.
- Zhu, B. and Lizieri, C. (2022).** Local Beta: Has Local Real Estate Market Risk Been Priced in REIT Returns? *Journal of Real Estate Finance and Economics*, 69(1), 682–718.