Toward a Sustainable Future in Real Estate: Exploring the Synergy of Environmental Sustainability, Innovation, and Regulation

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

vorgelegt von:
Benedikt Gloria

Berichterstatter:	Prof. Dr. Sven Bienert, MRICS REV
	Prof. Dr. Wolfgang Schäfers
Tag der Disputation:	09. Dezember 2024

UNIVERSITY OF REGENSBURG

Dissertation
submitted in partial fulfillment of the requirements for the degree of
Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

Toward a Sustainable Future in Real Estate: Exploring the Synergy of Environmental Sustainability, Innovation, and Regulation

submitted by
Benedikt Gloria, M.Sc.

Advisors:
Prof. Dr. Sven Bienert, MRICS REV
Prof. Dr. Wolfgang Schäfers

Contents

Abbreviations

AI	Artificial Intelligence
AIF	Alternative Investment Fund
API	Application Programming Interface
BVI	Bundesverband Investment und Asset Management
CapEx	Capital Expenditure
CNN	Convolutional Neural Network
CO_2e	Carbon Dioxide Equivalent
CSRD	Corporate Sustainability Reporting Directive
D/E	Debt to Equity
DSCR	Debt Service Coverage Ratio
EPC	Energy Performance Certificate
EPBD	Energy Performance of Buildings Directive
ESG	Environmental, Social, and Governance
ESRS	European Sustainability Reporting Standards
FFO	Funds From Operations
FM	Fama-MacBeth
GAR	Green Asset Ratio
GDP	Gross Domestic Product
GHG	Greenhouse Gas
GIS	Geographic Information System
GLS	Generalized Least Squares
GPT	Generative Pre-trained Transformer
Grad-CAM	Gradient-weighted Class Activation Mapping
GSV	Google Street View
GVIF	Generalized Variance Inflation Factor
HHI	Herfindahl–Hirschman Index
IPCC	Intergovernmental Panel on Climate Change
IPO	Initial Public Offering
IRR	Internal Rate of Return

KPI	Key Performance Indicator
kWh	Kilowatt-Hour
LLM	Large Language Model
LM	Lagrange Multiplier
LTV	Loan to Value
M/B	Market to Book
ML	Machine Learning
MSA	Metropolitan Statistical Area
MWh	Megawatt-Hour
NGFS	Network for Greening the Financial System
NLP	Natural Language Processing
NOI	Net Operating Income
NZBA	Net Zero Banking Alliance
OLS	Ordinary Least Squares
OpEx	Operating Expenses
PAI	Principal Adverse Impact
P/B	Price to Book
P/NAV	Price to Net Asset Value
PBR	Principles-Based Requirements
PCAF	Partnership for Carbon Accounting Financials
PPA	Power Purchase Agreement
QLoRA	Quantized Low-Rank Adaptation
RAG	Retrieval-Augmented Generation
REIT	Real Estate Investment Trust
ResNet	Residual Network
ROA	Return on Assets
ROE	Return on Equity
RTS	Regulatory Technical Standards
SAP	Standard Assessment Procedure
SDG	Sustainable Development Goal
SFDR	Sustainable Finance Disclosure Regulation
SOTA	State of the Art
TCFD	Task Force on Climate-related Financial Disclosures
VIF	Variance Inflation Factor

List of Figures

List of Tables

1 Introduction

1.1 Introduction and Background

Due to the significant impact of climate change on humanity, environmental sustainability has been on the global agenda for many years, and the economic relevance of climate change is growing. As temperatures and extreme weather events rise, the realized and potential damage to the systems of the global economy increases (Stern, 2008; Andersson et al., 2020). Real estate is a major source of climate-changing greenhouse gases through operational and built-in emissions, and one of the sectors most affected by rising urban temperatures and more frequent and severe extreme weather events (UNEP, 2022). Within the real estate sector, environmental sustainability or the minimization of greenhouse gas emissions is often associated with innovation: After all, innovations in heating, cooling and lighting technologies, among others, have the potential to save a large amount of operational emissions (Han et al., 2010; International Energy Agency, 2022), and optimization of building materials, for example through machine learning, can lead to savings in embodied carbon (Schneider et al., 2023; Zhu et al., 2022). However, because negative externalities or increasing climate risks are not yet adequately priced or internalized, market-based mechanisms tend to favor cost-effective solutions over environmentally beneficial ones (Stavins, 2003; Stern, 2008). Therefore, achieving carbon neutrality in the real estate sector may require regulatory intervention to ensure that market mechanisms are consistent with emission reduction targets. In the context of real estate, practical and feasible measures include, for example, the implementation of minimum standards and the introduction of a carbon price on emissions. Ensuring the availability of data of sufficient quantity and quality is essential for these measures to be effective. Without a robust database, the benchmarks obtained are of limited value and the accuracy of measuring the degree to which goals are being met is reduced. (Christensen & Gabe, 2019).

This dissertation examines five individual contributions across three interrelated areas: *environmental sustainability*, *innovation*, and *regulation*, aiming to explore the intersections and impacts within the real estate sector. The critical need for research in these areas is underscored by the following observations:

First, the significance of *environmental sustainability* in the real estate industry is becoming increasingly more important. The sixth report by the Intergovernmental Panel on Climate Change (IPCC) reiterates the need for a substantial intensification of global efforts to reduce greenhouse gas emissions. The scientific consensus is that meeting the Paris Climate Agreement's goal of limiting warming to 1.5 degrees Celsius above pre-industrial levels will require a 43% reduction[1] in GHG emissions from 2019 to 2030. Nevertheless, emissions have continued to rise in recent years (IPCC, 2023). While the future significance of environmental sustainability is widely recognized by industry stakeholders, current research predominantly examines its present-day effects on properties and markets. A large body of literature exists primarily on the price effects of sustainability at the company or property level, or on the effects of climate risks (Devine & Yönder, 2023; Bond & Devine, 2016; Salisu et al., 2024). To date, there has been limited research on the forward-looking analysis of unrealized potentials concerning transition risks. However, as buildings become more electrified and the power grid decarbonizes, the need for this type of forward-looking analysis of opportunities and risks at the individual property level will continue to grow.

Second, some market participants argue that *innovation* is needed to drive decarbonization (Geels et al., 2017). When it comes to real estate, it seems that it is not only the physical innovations related to the built environment that appear to be a nexus for a successful decarbonization. For example, directly measurable GHG savings can be achieved by using new information technologies to control and manage buildings (Naylor et al., 2018). However, quality-assured data as a basis for decision-making appears to be necessary for an efficient handling of sustainability and strategic decision-making processes at all levels of the real estate value chain. In both policy and market terms, strategic decisions need to be based on reliable data in order to invest where the most emissions can be saved, especially in terms of desired retrofit rates (Bienert & Groh, 2022). To date, data on real estate emissions has been subject to a high degree of opacity. The European Union has recently attempted to harmonize European Energy Performance Certificates (EPCs) with a reform of the Energy Performance of Buildings Directive (EPBD) (European Parliament and Council, 2024), as the creation of EPCs in their current form sometimes varies widely among member countries and emission values are not always reported (Semple & Jenkins, 2020). Yet because there has been little progress in the comprehensive collection and assessment of operational real estate emissions, there remains a mismatch between the available data and the decisions that need to be made. In this context, data models can be part of the solution, as well-trained models are capable of producing an out-of-sample prediction that can

[1] Based on a probability of occurrence of >50%.

help fill this data gap. Large Language Models (LLMs) such as ChatGPT could also be a solution—and not just for the estimations of emissions—as they are capable of data collection, analysis and predictive modeling if properly fine-tuned and trained. In addition, the applications in real estate are diverse and could support a wide range of daily tasks such as research, management, or marketing. However, research and application of LLMs in real estate is still scarce.

Third, due to the novelty of the *regulations* introduced, research on their effects seems to be underdeveloped and therefore information asymmetries still exist in the market. In 2018, the EU Commission issued the EU Action Plan on Financing Sustainable Growth related to the Paris Climate Agreement and the UN Sustainable Development Goals (Agenda 2030), which aims to redirect capital towards sustainable investments. The Action Plan results, among other things, in the Taxonomy Regulation (EU 2020/852) and the Sustainable Finance Disclosure Regulation (SFDR) (EU 2019/2088)—both of which are highly relevant in the context of real estate investments. The Taxonomy Regulation now provides, for the first time, a uniform classification system for sustainable investments and economic activities. The SFDR defines transparency requirements for sustainability aspects starting at the level of financial market participants and is particularly relevant in the context of indirect real estate investment. Nevertheless, recent research on the real estate implications of sustainability aspects has focused more on voluntary certificates than on regulatory requirements (Chegut et al., 2014; Holtermans & Kok, 2019).

This research contributes to the existing literature by elucidating the financial implications of *environmental sustainability*, *innovation*, and *regulation* in the real estate sector. Using econometric methods and machine learning algorithms to analyze different datasets, each paper within the dissertation examines a relevant subtopic and contextualizes the findings within the existing literature. The results are both academic and practical in nature, addressing current research gaps and highlighting the potential practical applications of the findings.

1.2 Research Questions

This section briefly describes the objectives of each of the five papers and the corresponding research questions. Figure 1.1 illustrates the assignment of the individual papers to the topic areas and the structure of the work.

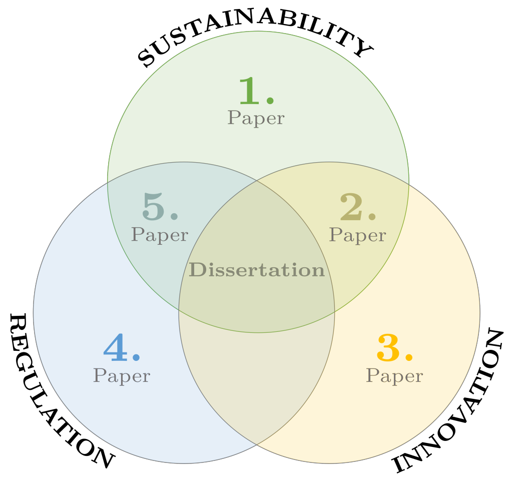


Figure 1.1: Overview of Individual Papers by Topic Areas and Structure

Rooftop Solar Potential and Financial Performance of REITs
Paper 1 takes a U.S. perspective and examines the measurable impact of rooftop solar potential and Real Estate Investment Trust (REIT) financials. For this purpose, we use a unique dataset, which includes the weighted available rooftop solar potential and the weighted carbon offset factor of the individual properties in the portfolio, in addition to the typical financial indicators at the REIT level. The effects are calculated using alternating four-factor models according to Fama and French (1993) and Carhart (1997) as well as fixed-effects models. The focus is on the following research questions:

- To what extent is the rooftop solar potential of commercial real estate in U.S. REIT portfolios financially priced in?
- What is the financial impact of the average carbon intensity of a U.S. REIT portfolio?
- Is there a difference in the impact of U.S. REITs' rooftop solar potential and average carbon intensity on stock performance, operating performance, and firm value?

Picture This: A Deep Learning Model for Operational Real Estate Emissions
Paper 2 focuses on developing a scientific and innovative solution to address data gaps in Carbon Dioxide Equivalent (CO_2e) emissions from residential buildings. Using a deep learning model, the aim is to find out whether building images and corresponding emission values can help to estimate emissions on a large scale. The topic is particularly relevant to practitioners because decision-making in the context of sustainability is sometimes based on only a small fraction of the actual data available. The research questions can be defined as follows:

- Can operational carbon emissions of residential properties be estimated using available building imagery and deep learning methods?
- Which Residual Networks (ResNets) model size, architecture, and hyperparameters provide the best results for estimating carbon emissions?
- Can the inclusion of additional information (e.g., energy source) further improve the accuracy of the model?

Real-GPT: Efficiently Tailoring LLMs for Informed Decision-Making in the Real Estate Industry
Paper 3 also deals with the handling of real estate-related data gaps. Since established LLMs currently only cover real estate topics to a limited extent, the goal of this paper is to develop a framework that allows real estate-specific topics to be represented as accurately as possible in smaller, open-source models in order to enable resource-saving, data-protected, and user-friendly application. Unstructured texts are utilized as a dataset, which are implemented in pre-trained models with the help of fine-tuning. Among others, the following research questions are analyzed:

- Can a suitable dataset and fine-tuning process adequately address the knowledge gap regarding real estate in established LLMs?
- How can unstructured data be transformed into a suitable and quality-assured fine-tuning dataset, and how can this data then be used for an effective fine-tuning process of established LLMs?
- How does a fine-tuned LLM compare to other State of the Art (SOTA) models with respect to real estate issues, and can any informational advantages impact the ability to make investment decisions?

Unveiling the Impact of SFDR on Unlisted Real Estate Funds: A J-Curve and Panel Regression Analysis

The SFDR is currently a much discussed topic in practice, especially with regard to indirect real estate investments. As data and research on this topic are limited, Paper 4 aims to shed light on possible financial relationships between the SFDR and non-listed real estate. In this context, the typical phasing of a fund, the so-called "J-Curve", is of particular importance for market participants. Methodologically, the paper follows a two-step approach by first deriving and analyzing the J-Curves using cubic spline interpolation and then applying a random effects model. The results of the paper answer the following research questions:

- What financial implications can be identified for unlisted funds that are subject to different SFDR categorizations?
- Especially in relation to the different SFDR categories, what is the typical J-Curve development phase for unlisted real estate?
- What financial aspects should stakeholders consider when dealing with SFDR for unlisted real estate?

Is there a Green Discount in Commercial Real Estate Lending?

While the analysis of the economic impact of sustainability on the real estate industry often focuses on prices and rents, the aim of the last paper is to investigate the impact of sustainable properties on the formation of loan interest rates. We use a hedonic pricing model to analyze a European bank loan portfolio with a volume of more than EUR 30 billion for the period 2018 to 2023. The research questions examine the relationship between the bank's profit margin—that is, the amount by which it exceeds the purchase rate—and the sustainability characteristics of the financed properties:

- Can green commercial real estate be financed at a lower cost than its non-green counterparts?
- Do lenders incentivize the financing of green real estate by foregoing margin, and what might be the rationale for a targeted incentive?
- Does the green building classification have an impact on the loan-to-value ratio of the financed property?

1.3 Co-Authors, Submissions and Conference Presentations

The following is an overview of individual co-authors, submissions, and conference presentations.

Paper 1

Rooftop Solar Potential and Financial Performance of REITs

Authors: Sebastian Leutner, Benedikt Gloria, Ben Höhn, Sven Bienert

Publication Details: Under review (23rd of June 2024) in the Journal of Real Estate Finance and Economics

Conference Presentation: Presented at the 30th European Real Estate Society Annual Conference in Sopot & Gdańsk 2024

Award: Received the 2024 European Real Estate Society Best PhD Paper Award

Paper 2

Picture This: A Deep Learning Model for Operational Real Estate Emissions

Authors: Benedikt Gloria, Ben Höhn

Publication Details: Accepted (22nd of August 2023) and published online (21st of September 2023) in the Journal of Sustainable Real Estate

Conference Presentation: Presented at the 39th American Real Estate Society Annual Conference in San Antonio 2023, the European Real Estate Society PhD-Conference 2023 (online) and the ME100 Conference in Prague 2023

Award: Received the 2023 American Real Estate Society Sustainable Real Estate Award

Paper 3

Real-GPT: Efficiently Tailoring LLMs for Informed Decision-Making in the Real Estate Industry

Authors: Benedikt Gloria, Johannes Melsbach, Sven Bienert, Detlef Schoder

Publication Details: Accepted (21st of June 2024) and pre-published online (21st of August 2024) in the Journal of Real Estate Portfolio Management

Conference Presentation: Presented at the 30th European Real Estate Society Annual Conference in Sopot & Gdańsk 2024

Paper 4

Unveiling the Impact of SFDR on Unlisted Real Estate Funds: A J-Curve and Panel Regression Analysis

Authors: Benedikt Gloria, Sebastian Leutner, Sven Bienert

Publication Details: Accepted (14th of March 2024) and pre-published online (4th of April 2024) in the Journal of Property Investment & Finance

Conference Presentation: Presented at the 29th European Real Estate Society Annual Conference in London 2023

Paper 5

Is there a Green Discount in Commercial Real Estate Lending?

Authors: Sebastian Leutner, Benedikt Gloria, Sven Bienert

Publication Details: Accepted (03rd of January 2024) and pre-published online (23rd of January 2024) in the Journal of Property Investment & Finance

Conference Presentation: Presented at the 23rd African Real Estate Society Annual Conference in Livingstone 2024

1.4 References

Andersson, M., Morgan, J., & Baccianti, C. (2020). Climate change and the macro economy. *ECB Occasional Paper, June 2020/No 243.*

Bienert, S., & Groh, A. M. (2022). Klimaneutralität vermieteter Mehrfamilienhäuser? Aber wie? [Climate neutrality for rented apartment buildings? But how?]. *Beiträge zur Immobilienwirtschaft*, *28*.

Bond, S. A., & Devine, A. (2016). Certification matters: Is green talk cheap talk? *The Journal of Real Estate Finance and Economics*, *52*(2), 117–140.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Chegut, A., Eichholtz, P., & Kok, N. (2014). Supply, demand and the value of green buildings. *Urban studies*, *51*(1), 22–43.

Christensen, P. H., & Gabe, J. (2019). Public regulatory trends in sustainable real estate. In T. Walker, C. Krosinsky, L. N. Hasan, & S. D. Kibsey (Eds.), *Sustainable real estate: Multidisciplinary approaches to an evolving system* (pp. 35–76). Cham: Springer International Publishing.

Devine, A., & Yönder, E. (2023). Impact of environmental investments on corporate financial performance: Decomposing valuation and cash flow effects. *The Journal of Real Estate Finance and Economics*, *66*(4), 778–805.

European Parliament and Council. (2024). Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance). *Official Journal of the European Union*.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Geels, F. W., Sovacool, B. K., Schwanen, T., & Sorrell, S. (2017). Sociotechnical transitions for deep decarbonization. *Science*, *357*(6357), 1242–1244.

Han, H., Jeon, Y., Lim, S., Kim, W., & Chen, K. (2010). New developments in illumination, heating and cooling technologies for energy-efficient buildings. *Energy*, *35*(6), 2647-2653.

Holtermans, R., & Kok, N. (2019). On the value of environmental certification in the commercial real estate market. *Real Estate Economics*, *47*(3), 685–722.

International Energy Agency. (2022). Technology and innovation pathways for zero-carbon-ready buildings by 2030. *Part of Net Zero Emissions Online Publications*.

IPCC. (2023). Climate change 2023: Synthesis report. *Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 35–115.

Naylor, S., Gillott, M., & Lau, T. (2018). A review of occupant-centric building control strategies to reduce building energy use. *Renewable and Sustainable Energy Reviews*, *96*, 1-10.

Salisu, A. A., Ogbonna, A. E., & Vo, X. V. (2024). Climate risks and the REITs market. *International Journal of Finance & Economics*.

Schneider, M., Hoenig, V., Ruppert, J., & Rickert, J. (2023). The cement plant of tomorrow. *Cement and Concrete Research*, *173*, Article 107290.

Semple, S., & Jenkins, D. (2020). Variation of energy performance certificate assessments in the European Union. *Energy Policy*, *137*, Article 111127.

Stavins, R. N. (2003). Chapter 9 - experience with market-based environmental policy instruments. In K.-G. Mäler & J. R. Vincent (Eds.), *Environmental degradation and institutional responses* (Vol. 1, p. 355-435). Elsevier.

Stern, N. (2008). The economics of climate change. *American Economic Review*, *98*(2), 1-37.

UNEP. (2022). 2022 Global Status Report for Buildings and Construction. Towards a zero-emissions, efficient and resilient buildings and construction sector. *United Nations Environment Programme*.

Zhu, F., Wu, X., Zhou, M., Sabri, M. M. S., & Huang, J. (2022). Intelligent design of building materials: Development of an AI-based method for cement-slag concrete design. *Materials*, *15*(11), Article 3833.

2 Rooftop Solar Potential and Financial Performance of REITs

2.1 Abstract

Extensive research on REITs and their engagement in sustainable practices has shown that firms' environmental performance has an increasingly significant impact on financial outcomes. However, there is little research on how REITs can leverage their own property holdings to mitigate corporate transition risks. Using recently available rooftop data for a sample of U.S. REITs, we examine how the solar potential of properties affects financial performance at the firm level. While we find only weak evidence of an impact on market betas, abnormal returns, and operating performance, there is a significant valuation effect on firm values that extends beyond REITs' underlying properties. These results suggest that the market is placing a higher value on the future earnings of REITs with higher solar potential, reflecting the expected financial benefits and the potential to mitigate transition risks. We also provide evidence that firm values are affected by upstream emissions through local grid dependency outside REITs' direct control.

Keywords: *REITs; Financial Performance; Solar; Renewable Energy; Transition Risk.*

2.2 Introduction and Related Literature

Although global financial markets have seen significant growth in sustainable investment over the past decade, during periods of economic stress and challenging market environments, management may perceive investments in sustainability as a distraction and prioritize thus core business operations (Chacon et al., 2024). However, effectively managing climate risks could be a key factor in enhancing resilience against wider systematic risks (Albuquerque et al., 2019). Existing research provides ample evidence that asset prices are already strongly influenced not only by a company's environmental disclosure, but also by its actual carbon emissions as well as the uncertainty associated with transition risks (Bolton & Kacperczyk, 2021; Choi et al., 2020; Matsumura et al., 2014; Ilhan et al., 2021; Seltzer et al., 2022; Bolton & Kacperczyk, 2023). Although the magnitude and timing of transition risks, such as declining market attractiveness, reputational damage, and regulatory pressure, are highly uncertain, carbon-intensive business models are particularly exposed to these risks (Ilhan et al., 2021). The financial implications at the firm level increasingly capture the true sensitivity of investors and lenders to corporate environmental performance, rather than simply reflecting other omitted corporate risks (Chava, 2014; Krueger et al., 2020).

Similar findings regarding capital market preferences and pricing mechanisms in relation to transition risks can be found in the literature focused on REITs. The profound impact of transition risks on both global and regional REIT markets underscores the critical importance of aligning investment decisions (Salisu et al., 2024). Given the limited availability of comprehensive whole-building energy data and corresponding emissions data at the firm level, studies in this area focus on two primary metrics to elucidate the potential impact of environmentally-sustainable investment on financial performance of REITs: (1) either examining the proportion of environmentally certified properties within REIT portfolios or (2) assessing GRESB scores and disclosure. Eichholtz et al. (2012) and Fuerst (2015) show that both the share of LEED- and Energy Star-certified properties and the GRESB score have a positive impact on REITs' operating performance and reduced systematic risk, suggesting reduced exposure to energy price fluctuations and regulatory risks. There is also evidence that greener REITs are associated with superior stock performance, as highlighted by Eichholtz et al. (2012) and Sah et al. (2013). Moreover, REITs that engage in environmental-sustainable practices benefit from lower debt financing costs. This is evident at both the property and the firm level (Eichholtz, Holtermans, & Kok, 2019; Feng & Wu, 2023). While finding evidence supporting a reduced cost of capital for greener REITs, Devine et al. (2024) also show that green bond issuance improves corporate environmental performance as well. By isolating the income and capital channels that drive the positive impact of green property investment at the firm level, Devine and Yönder (2023) examine the underlying pricing mechanisms of green building investments of REITs identified in previous studies. Beyond the benefits of better operating performance, lower borrowing costs, and reduced systematic risk at the property level, their research highlights that the relative premium in equity markets exceeds the premium observed in property markets. They find that the mere participation of REITs in such certification schemes has an impact on investor and lender pricing, confirming the signaling effects of certifications found by (Bond & Devine, 2016).[1]

While the benefits to REITs of holding a higher fraction of properties with green building certifications or ratings are well documented in the literature, it is critical to recognize that these frameworks alone, in their current form, are not sufficient to optimize both environmental and financial performance (Clayton et al., 2021; Newsham et al., 2009). Using GRESB performance scores for REITs, Chacon et al. (2024) provide evidence that capital of REITs might be overallocated to Environmental, Social, and Governance (ESG) investments, resulting in lower firm values and cash flow. Although all three dimensions of ESG are taken into account, the study indicates that a careful balance should be struck between the short-term costs and long-term benefits of environmental-sustainable investments. This underscores that adherence to the box-ticking approach prevalent in certification and rating systems may not always be in the best interests of investors, lenders, and tenants in terms of capital allocation and thus also of efficient transition risk management.

Focusing on actual carbon and energy intensity is essential to effectively managing transition risks in property portfolios, as suggested by findings of Akhtyrska and Fuerst (2024). In this context, operational energy plays a key role not only from an environmental perspective, but also from an economic one. This importance, particularly for electricity, is expected to grow in the coming years. First, the phase-out of on-site fossil fuel combustion is driving a shift toward electrification of heating systems. Second, electric cooling loads, already a significant component of today's commercial building energy use, are expected to increase, particularly during periods of extreme heat. Third, geopolitical tensions combined with uncertainty about carbon pricing are expected to put additional inflationary pressure on electricity prices (Fuss et al., 2009). Despite being historically the largest portion of a commercial property's Operating Expenses (OpEx), energy costs are proving to be a highly manageable factor (Wiley et

[1] There is also an extensive body of literature highlighting the benefits of environmentally sustainable investments at property level. See An and Pivo (2020); Chegut et al. (2014); Devine et al. (2022); Eichholtz et al. (2010, 2013); Fuerst and McAllister (2011); Holtermans and Kok (2019); Reichardt et al. (2012); Wiley et al. (2010).

al., 2010).

Krueger et al. (2020) note, that institutional and long-term investors in particular are prioritizing resilience against transition risk over divesting from assets with heightened risk exposure. REITs that focus on long-term value creation are likely to benefit from a performance premium compared to their peers (Feng et al., 2022). In addition to adopting energy efficiency measures, investing in on-site renewable energy sources can significantly reduce the transition risk of property portfolios. This is because REITs are becoming less dependent on the decarbonization of the local grid and on-site renewable energy enables a more efficient switch to electrified heating systems. In many cases, REIT portfolios encompass extensive rooftop space, ideally suited for the installation of solar systems.[2] By utilizing this infrastructure, REITs can decrease their carbon footprint and attain financial benefits, particularly in states where the local grid is heavily reliant on fossil fuels and electricity rates are elevated. The electricity generated through rooftop solar systems can either be sold back to the grid, provided directly to tenants via Power Purchase Agreements (PPAs), or utilized in a hybrid approach.[3] Although the solar potential among REITs varies with their sector focus and the corresponding roof space availability, the benefits of such renewable energy investments are not confined to large-scale warehouses or retail properties. They also extend to property types with more limited rooftop space, including office buildings (Jurasz & Campana, 2019).

REITs may experience financial benefits from investing in rooftop solar through both income and capital channels. One of these benefits is the reduction in energy costs by switching from grid electricity to solar power, thereby reducing operating expenses.[4] In addition, Jurasz and Campana (2019) find that solar systems can effectively reduce peak loads in commercial properties, resulting in decreased energy costs. Solar systems are most beneficial in triple-net leases, where tenants directly benefit from energy cost savings, allowing landlords to charge higher effective gross rents (Reichardt, 2014). Beyond the direct impact on properties' cash flow, solar systems could also create indirect positive cash flow effects. Devine and Kok (2015) document that environ-

[2]See an article from Nareit on April 22, 2022, at the following website: https://www.reit.com/news/blog/market-commentary/reit-solar-opportunities-35-billion-square-feet-roof-space.

[3]Throughout this paper, we refer to solar systems as a generic term for ease of discussion, although technically there is a difference between solar thermal (conversion of solar radiation to heat) and photovoltaic (PV) systems (conversion of solar radiation to electricity).

[4]For a comparison of the cost of solar power and grid power in the U.S., see the Photovoltaic Energy Factsheet 2023 from the University of Michigan's Center for Sustainable Systems at the following website: https://css.umich.edu/publications/factsheets/energy/photovoltaic-energy-factsheet.

mentally efficient buildings benefit from higher occupancy rates, improved tenant retention, and shorter re-letting periods. This not only leads to more stable cash flows, but also reduces operating costs through lower brokerage commissions and tenant incentives, which in turn could lead to lower overall management costs for REITs. As shown by Beracha et al. (2019), REITs with lower operating costs demonstrate improved operational performance, which is reflected in higher Return on Assets (ROA) and Return on Equity (ROE), as well as reduced market volatility. Existing research on the positive impact of rooftop solar on property values centers primarily on the residential sector (Aydin et al., 2018; Dastrup et al., 2012; Hoen et al., 2013; Qiu et al., 2017). However, Leskinen et al. (2020) note that investments in solar systems, can also favorably influence commercial property yields. They suggest that the rapid growth of electricity prices outpacing rental growth, coupled with reduced obsolescence of on-site energy investments, is lowering the risk premium for property investments.

While there is some limited empirical evidence supporting the potential benefits of solar for commercial real estate at the property level, to our knowledge, no attention has been paid to the relationship between rooftop solar and the financial performance of REITs. This paper aims to fill this gap by using a unique dataset to examine the impact of REITs' solar potential on stock performance, operating performance, and firm values. REITs provide an appropriate framework for the identification of contemporaneous effects on market returns and cash flows, since profitability ratios and stock returns can be calculated on a periodic basis. In addition, REITs' underlying properties are actively traded in a secondary market, allowing for a direct comparison of asset and shareholder value, which can provide valuable insight into future market expectations.

To pursue our analysis, we collect financial data and company characteristics from listed U.S. equity REITs and match these datasets with recently available data on the rooftop area suitable for solar installations on each property owned by a REIT. Additionally, we consider the carbon intensity of the local electric grids to which each property is connected. This approach serves as a proxy for both the potential to mitigate transition risks by switching from grid electricity to solar, as well as the exposure to upstream emissions. We aggregate the property-level data sets for each REIT in our panel to quantify the company-level financial impact. To ensure the reliability of our research methodology and thus our empirical results, we perform a series of robustness checks.

Our findings suggest no significant relationship between REITs' solar potential and stock performance, and only limited evidence of a relationship with operating performance. Given the barriers REITs face in adopting rooftop solar to date, we attribute the lack of widespread industry adoption as the primary driver of these results. However, our analysis provides robust statistical and economic evidence that both solar and carbon offset potential significantly impact firm values. This implies that the market incorporates the financial benefits and reduced risks into the pricing of REITs' future cash flows. Furthermore, we find that firm values are affected not only by direct carbon emissions, but also by upstream emissions that depend on the local grid. Overall, our results have important implications for REIT management, stakeholders, and policymakers. Consistent with previous studies, the market values management's commitment to environmental practices, which then translates into financial benefits. REITs that do not adhere to these practices may face market penalties. Policymakers can leverage these market dynamics to unlock the solar potential of the REIT sector and support the decarbonization of the property sector by removing barriers and providing incentives for solar installations. Our findings contribute to the existing literature by providing a new perspective that refines the understanding of environmentally-sustainable property investment, firms' exposure to transition risks and how property portfolios can be actively leveraged to mitigate these risks while generating financial benefits.

The remainder of the paper is structured as follows. The next section presents our data sources, variables, and summary statistics, followed by a description of our research methodologies. We then discuss our results and robustness checks. The final section of the paper provides our conclusions.

2.3 Data and Summary Statistics

We collect data on firm characteristics of all available U.S. listed equity REITs from 2012 to 2023 at an annual frequency, sourced from S&P Capital IQ Pro (formerly SNL Financial). Daily stock market returns are obtained from Refinitiv Datastream (formerly Reuters). Research factors and risk-free rate data are downloaded from Kenneth R. French's Data Library.[5] Our time series data covers a full market cycle and captures potential shifts in REITs' adoption of solar systems. To account for potential selection bias, we consider all REITs regardless of solar potential.

In our analysis of the impact of REITs' solar potential on financial performance, we calculate annualized alphas and betas based on daily stock returns, using a standard 4-factor model according to Fama and French (1993) and Carhart (1997), ROA, ROE, Market to Book (M/B) ratio, and Price to Net Asset Value (P/NAV) ratio.[6]

[5] See Kenneth R. French's website: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data-library.html.

[6] A more detailed description of our dependent variables and the estimation strategy we employ can be found in the next section of this paper.

Consistent with previous studies, we include the following explanatory variables: real estate investment growth, measured as the growth rate of REITs' real estate investments based on the previous reporting period; total assets, defined as total assets owned by the REIT and reported on the balance sheet; Debt to Equity (D/E) ratio, defined as debt as a percentage of total equity; firm age, calculated as one plus years since Initial Public Offering (IPO) or establishment of REIT status; and percentage held by institutional owners, which might add explanatory power with regard to adoption of sustainability practices, as suggested by Dyck et al. (2019) and Krueger et al. (2020). In line with Devine and Yönder (2023), we calculate the proportion of properties that are less than 10 years old. This allows us to control for the quality of property portfolios and the potential suitability of the buildings for the installation of solar systems. Real Gross Domestic Product (GDP) growth by Metropolitan Statistical Area (MSA), considered a typical leading indicator of real estate returns, is included as an external variable to account for potential omitted economic drivers. Following Capozza and Seguin (1999), we use the Herfindahl–Hirschman Index (HHI) to measure both the geographic and property type diversification of the REITs in our sample, where s_i denotes a REIT's market concentration, either in terms of its portfolio share by MSA or real estate property type. Index values range from zero to one, with values close to zero suggesting high diversification. The HHI may capture both the heterogeneity of solar potential across property types and local environmental awareness, such as mandatory solar requirements for properties.

$$HHI = \sum_{i=1}^{N} s_i^2 \tag{2.1}$$

To determine the rooftop solar potential of REITs, we first derive latitude and longitude data from the property addresses of all REIT-owned properties by applying Geographic Information System (GIS) methods. We then use Google's Solar Application Programming Interface (API) to match the latitude, longitude and building size data with the roof area data. The Solar API applies machine learning to assess the available roof area for solar arrays on properties based on satellite images, shading analysis, roof elevations, building outlines, and built-up roof areas. In particular, it identifies potential obstacles on the roof, such as ventilation systems, that could prevent the installation of solar panels.[7] By focusing on the specific characteristics of individual properties, our dataset provides a more accurate determination of solar potential than generalized assumptions based on building footprint or total roof area of buildings applied in

[7] See the detailed documentation of the Solar API on the Google Maps Platform: https://developers.google.com/maps/documentation/solar.

existing literature. We define a REIT's solar potential as the average maximum solar array area (in square feet) that can be installed on the roofs of each property n owned by a REIT i in year t. By weighting each property's maximum solar array area relative to the REIT's total area, we account for differences in portfolio size and property type focus.

To further address the impact of solar installations on operational carbon emissions and to proxy the potential to mitigate transition risks by switching from grid electricity to solar, we include a REIT-specific carbon offset factor. This factor measures the equivalent carbon emissions in kg of CO_2 per Megawatt-Hour (MWh) of grid electricity that a solar system would displace, reflecting each REIT's unique carbon displacement potential. We consider this variable to have additional explanatory power with respect to solar potential, as the results of Bolton and Kacperczyk (2023) show that the financial performance of firms is also affected by indirect upstream emissions. For REITs, this suggests that not only the environmental performance of properties, but also the energy mix of the electrical grid is a critical factor influencing transition risk exposure.[8] Since grid electricity consumption—and therefore carbon emissions—is directly proportional to property size, the carbon offset factor for each REIT is again calculated by weighting the property-specific offset factor by its square footage relative to the REIT's total floor area.

$$Solar\ Potential_{i,t} = \sum_{n=1}^{N} \frac{Floor\ Area_{i,t}^{n} \times Max\ Array\ Area_{i,t}^{n}}{Floor\ Area_{i,t}} \tag{2.2}$$

$$Carbon\ Offset_{i,t} = \sum_{n=1}^{N} \frac{Floor\ Area_{i,t}^{n} \times Max\ Carbon\ Offset_{i,t}^{n}}{Floor\ Area_{i,t}} \tag{2.3}$$

REITs with incomplete data on the latitude and longitude of their properties or other missing control variables are excluded from our analysis. Moreover, we only consider firms for which we can obtain at least 90% of the solar potential data across the entire portfolio via Google's Solar API. Our final sample consists of 103 REITs. To reduce the influence of outliers, we apply winsorization to the variables at the 5th and 95th percentiles. The summary statistics are presented in Table 2.1.

The average estimated market beta is 0.76, while the average annualized alpha is 0.4%. REITs have an average ROA of 1.95% and a ROE of 4.75%. The mean M/B ratio is 2.29 and the mean P/NAV ratio is 0.91. REITs hold on average \$5.11 billion in real estate assets, with an average D/E ratio of 1.38. REITs' real estate investments are growing

[8] Due to multicollinearity between carbon offset factors and local utility rates, we do not include the latter in our dataset.

Table 2.1: Descriptive Statistics

Statistic	Mean	Median	St. Dev.	Min	Max
Alpha (%)	0.37	0.90	6.78	−13.29	13.93
Beta	0.77	0.74	0.28	0.00	2.66
ROA (%)	1.95	1.97	1.34	−0.90	4.93
ROE (%)	4.75	3.98	8.75	−13.00	24.64
M/B ratio	2.29	1.89	1.49	0.78	6.83
P/NAV ratio	0.91	0.93	0.11	0.67	1.17
Solar potential (sq ft)	37,889.33	29,235.59	29,819.40	3,911.27	130,619.10
Carbon offset (kgCO_2/MWh)	556.00	545.24	64.85	391.20	660.76
Total assets ($B)	5.11	3.22	5.18	0.08	19.60
Real estate Inv. growth (%)	7.55	3.88	17.37	−10.67	71.21
D/E ratio	1.38	1.04	0.97	0.38	4.06
Institutional ownership (%)	14.60	14.93	3.20	3.57	19.14
Share buildings < 10 years (%)	11.49	8.49	8.16	1.38	32.68
Firm age	18.58	19.00	9.52	2.00	44.00
HHI property type	0.80	0.90	0.21	0.28	1.00
HHI region	0.19	0.08	0.21	0.02	0.73
GDP growth (%)	2.43	2.11	0.77	0.43	4.14

The table presents the descriptive statistics for the variables used in this analysis covering the sample period from 2012 to 2023. Firm characteristics are sourced from S&P Capital IQ Pro. Annualized alphas and betas are calculated using a standard Fama-French-Carhart 4-factor model. The key independent variables, solar potential and carbon offset, are obtained from the Google Solar API. To mitigate the influence of outliers, the variables are winsorized at the 5^{th} and 95^{th} percentiles.

at an average annual rate of 7.55%. The mean firm age is 18.58 years, with 11.00% of properties less than 10 years old. With an average HHI of 0.19 across regions compared to 0.80 for property types, REITs exhibit greater geographic diversification than sector diversification. Institutional owners hold on average 14.60% of REIT shares. In terms of solar potential, the average solar panel array area per property is 37,889.33 square feet, with a standard deviation of 29,235.59 square feet. By using solar power instead of grid electricity, REITs could offset an average of 556.00 kg of carbon per MWh of solar energy generated, with a standard deviation of 64.85 kg of carbon per MWh. This suggests that while there is some variation in the solar potential of REITs, their ability to displace carbon is relatively consistent across our sample.

2.4 Research Methodologies

We first estimate the impact of solar potential and carbon offset factors on REIT stock performance by calculating alphas (abnormal return) and market betas (systematic risk) for each REIT in our sample using a standard 4-factor model according to Fama and French (1993) and Carhart (1997).[9]

$$\left(R_{i,t} - R_{f,t}\right) = \alpha_i + \beta_i \left(R_m - R_f\right)_t + \gamma_{i1} SMB_t + \gamma_{i2} HML_t + \gamma_{i3} MOM_t + \varepsilon_{it} \tag{2.4}$$

$R_{i,t}$ is the daily stock return of REIT i on day t. SMB (small minus big), HML (high minus low) and MOM (momentum) are the Fama-French-Carhart factors. While some studies include the S&P 500 index as a proxy for market return R_m, we follow the large body of existing REIT literature by using the NAREIT index in our analysis.[10] Abnormal returns may reflect a positive market perception of REITs' solar potential and the associated financial benefits beyond the stock price. As suggested by Eichholtz et al. (2012), betas may provide valuable insights into REITs' exposure to systematic risk factors, such as fluctuating energy prices or environmental policies, which could potentially be mitigated by solar.

Next, we examine the impact of our key independent variables on REITs' operating performance. We calculate ROA and ROE by dividing net income by lagged total assets and total equity, respectively. Although Funds From Operations (FFO) is often considered a more accurate measure of cash flow from core business operations as non-cash expenses are added, we use net income for consistency with previous studies. The results of Vincent (1999) show that both net income and FFO provide valuable information content, with some evidence suggesting that net income may be relatively more informative. We include both ROA and ROE as measures of operating performance because of their different characteristics. While ROA may not be distorted by potential leverage, as noted by Core et al. (2006), ROE is widely considered to be a superior measure of operating performance from a shareholder perspective, providing a clearer assessment of how efficiently invested capital is being used (Brown & Caylor, 2009). Particularly for REITs that are already leveraging their solar potential, the aforementioned cash flow effects as well as reduced operating and management expenses may be reflected in both ROA and ROE.

Finally, we estimate the impact of *Solar Potential* and *Carbon Offset* on corporate valuation. For this purpose, we calculate the M/B ratio which is defined as a REIT's equity market capitalization as a multiple of its book value. Market capitalization incorporates the fair price of a stock, determined in standard valuation models as the present value of firm's future dividends or earnings. Thus, according to Fama and French (1995), the

[9] For additional robustness, we also run the estimation using a Fama and French (2015) five-factor model. The results show consistent values for alphas and betas.

[10] We derive annual alphas by assuming that a year has 252 trading days.

price of a stock and therefore our dependent variables may be affected by both growth rates of future cash flows and changes in investors' required rate of return. Boudry et al. (2010) note that REITs face a double cost of equity capital, which may affect REIT values not only through public markets, but also through private markets given the active secondary market trading of REITs' underlying properties. Since the book value as a past oriented accounting measure is potentially distorted by depreciation and may not fully reflect private market influences, we also use P/NAV as an independent variable, which is calculated by dividing a REIT's share price by its net asset value relative to shares outstanding.

In our model selection, we face methodological challenges similar to those highlighted by Bauer et al. (2010). Our main predictor variables—*Solar Potential* and *Carbon Offset*—vary across REITs, but are relatively constant over time. This data structure poses a challenge when using a firm fixed effects model, as it becomes difficult to disentangle the firm effects captured by our explanatory variables from those captured by the firm fixed effects, leading to multicollinearity issues. We therefore use an Ordinary Least Squares (OLS) model with sector fixed effects and heteroscedasticity-robust standard errors that are adjusted for dependencies within the firm clusters. This approach is preferable to a Generalized Least Squares (GLS) model because, although it also accounts for auto- and cross-sectional correlation, it tends to produce overly confident standard errors, especially for large N and small T samples (Beck & Katz, 1995). Our regression model can be written as

$$Y_{i,t} = \beta_0 + \beta_1 X_{i,t} + \gamma Z_{i,t} + \delta_{s,t} + \epsilon_{i,t} \tag{2.5}$$

where Y_{it} represents our outcome variables alpha, beta, ROA, ROE, M/B ratio, or P/NAV ratio for REIT i in year t. We use the logarithmic form of our dependent variable where the data distribution allows. X_{it} is a vector of our logarithmized independent variables of interest $Solar\ Potential_{i,t}$ and $Carbon\ Offset_{i,t}$. Z_{it} is a vector of controls and $\delta_{s,t}$ represents the sector fixed effects. We include total assets and REIT age in their logarithm. Since year-end financial performance is more likely to be affected by characteristics at the beginning of the reporting period than by contemporaneous effects, we use the lagged versions of real estate investment growth, D/E ratio and ROA as covariates.

To account for cross-sectional residual dependence and a potential unobserved time effect, we also estimate yearly regressions as suggested by Fama and MacBeth (1973). In these situations, the Fama-MacBeth (FM) approach provides unbiased standard errors and correctly sized confidence intervals (Petersen, 2008). Comparing the FM standard errors with the standard errors clustered at the firm level allows us to assess the relative magnitude of the firm and time effects in our models and thus a potential bias in our results. In the first step, we run T cross-sectional regressions for each year in the same setup as our panel models, where β_{0i} and β_{1i} are entity-specific intercepts and coefficients as presented in Eq. (6). From these regressions, we then derive the time series average and the variance of the aggregated coefficients to compute p-values using Fama-MacBeth standard errors.

$$Y_{i,t} = \beta_{0,i} + \beta_{1,i} X_{i,t} + \gamma Z_{i,t} + \epsilon_{i,t} \tag{2.6}$$

2.5 Empirical Results

2.5.1 Stock Performance

Table 2.2 shows the regression results for REITs' stock performance and systematic risk as measured by their alphas and betas. Columns (1) and (3) report the coefficient estimates from the Fama-MacBeth regressions, while Columns (2) and (4) provide the coefficient estimates from the fixed effects models. Although we find a positive coefficient of 0.133 for *Solar Potential* on alphas in the Fama-MacBeth model, suggesting that REITs with more available solar array area may generate excess returns, we do not observe statistical significance in the estimated coefficients for either model. Our results therefore imply that *Solar Potential* has no explanatory power for the excess returns of REITs. Moreover, the models estimate a negative coefficient for *Carbon Offset* (Columns 1 and 2). While a negative coefficient may suggest that REITs with higher exposure to properties operating in areas with carbon-intensive grids have lower excess returns, we again find no significant results. However, if we consider REITs' solar potential as a driver of stock performance, the lack of statistical significance for alphas may indicate that potentially higher cash flows or growth rates are already priced into stock prices and therefore no excess returns are observed (Eichholtz et al., 2012).

According to Columns (3) and (4), we find no significant influence of our primary independent variable, *Solar Potential*, on market betas in either model. This implies that the systematic risk is not affected by a REIT's available solar array area. Bolton and Kacperczyk (2021) point out that the adoption of renewable energy technologies is limited to certain sectors and therefore may not be reflected in the systematic risk. We further suggest that our findings may be due to a lack of progress in solar installations. While the financial viability of these strategies is contingent on a multitude of variables, including the specifics of the individual properties, the ownership structure (owner-operator or third-party ownership), local utility rates, and net metering regulations, the potential economic benefits remain unrealized in many cases. Fleten et al. (2007) and Leskinen et al. (2020) note that the adoption of solar systems lags behind the pace suggested by their financial benefits. They point out that investors are delaying their investments longer than profitability analyses recommend, highlighting a discrepancy between expected and actual investment behavior. In addition, due to their income tax exemption, REITs have historically faced challenges in using renewable energy tax credits to offset taxable income or monetizing the tax credits through sale. Although recent regulations address this issue, REITs' rooftop solar systems have primarily been operated by third-party owners or through taxable REIT subsidiaries.[11] REITs with large solar potential further face the challenge of generating more energy than their properties consume, but are unable to sell excess power to the grid due to local regulatory and technical constraints such as the lack of rooftop community solar projects.

In a similar manner to *Solar Potential*, we do not find a significant relationship between *Carbon Offset* and market betas in the fixed effects model as shown in Column (4). REITs with high offset factors are more dependent on grids that have a carbon-intensive energy mix, primarily due to electricity generation from fossil fuels such as coal, oil or gas. The grid dependency and associated upstream emissions expose REITs to greater carbon risk, as evidenced by the findings of Bolton and Kacperczyk (2023). However, due to relatively stable natural gas prices and historically low carbon prices, any significant impact on REITs' sensitivity to systematic risk has likely been mitigated to date. In addition, the lack of statistical significance is not entirely unexpected, as estimating the impact on alphas and betas derived from daily stock returns using explanatory variables based on annual data can be challenging. Furthermore, alphas and betas are themselves estimates that serve as independent variables. This adds to the complexity of deriving significant estimates for the coefficients, which is evident when compared to the overall statistical significance of our explanatory variables and covariates in other models reported in Tables 2.3 and 2.4.

[11] The Department of the Treasury and the Internal Revenue Service (IRS) have approved final regulations under the Inflation Reduction Act of 2022 for the transfer of tax credits, effective by July 1st, 2024. See the published document: https://www.govinfo.gov/content/pkg/FR-2024-04-30/pdf/2024-08926.pdf.

Table 2.2: Regression Results for Stock Performance

Variables	Alpha		Beta (log)	
	(1)	(2)	(3)	(4)
Solar potential (log)	0.133	−0.043	0.028	−0.001
	(0.378)	(0.074)	(0.020)	(0.017)
Carbon offset (log)	−3.177	−0.641	0.135	0.334
	(2.024)	(2.256)	(0.161)	(0.339)
Total assets (log)	−0.758**	−0.572***	0.046	0.049***
	(0.348)	(0.076)	(0.029)	(0.005)
Firm age (log)	0.560	0.321*	−0.042	0.008
	(0.432)	(0.180)	(0.036)	(0.022)
Real estate Inv. growth$_{t-1}$	0.049**	0.023***	−0.002***	−0.001
	(0.022)	(0.008)	(0.001)	(0.001)
D/E ratio$_{t-1}$	0.437**	0.093	0.012	0.012
	(0.219)	(0.164)	(0.013)	(0.015)
ROA$_{t-1}$	0.331	0.165**	−0.007	−0.004
	(0.283)	(0.080)	(0.016)	(0.009)
Share buildings < 10 years	−0.013	−0.027*	−0.001	−0.003
	(0.026)	(0.015)	(0.002)	(0.002)
GDP growth (MSA)	0.474	0.189	−0.020	−0.027
	(0.441)	(0.212)	(0.042)	(0.035)
Institutional ownership	0.048	0.049	0.018***	0.020***
	(0.085)	(0.056)	(0.006)	(0.007)
HHI region	−3.000	−0.899***	0.220**	0.214
	(2.044)	(0.281)	(0.089)	(0.202)
HHI property type	1.617	1.957***	0.261***	0.112
	(1.289)	(0.161)	(0.081)	(0.112)
Constant	25.157*	5.204	−2.419**	−3.627
	(13.125)	(13.381)	(1.044)	(2.218)
Coefficient estimates	FM	OLS	FM	OLS
Standard errors	FM	CL - F	FM	CL - F
Sector FE	N	Y	N	Y
Year FE	N	Y	N	Y
Observations	1,010	1,010	1,010	1,010
R-squared	0.321	0.238	0.254	0.228

The table reports the results of OLS and Fama-MacBeth regressions for the stock performance of U.S. equity REITs as a function of their solar potential and carbon offset and a vector of control variables. Annualized alphas and betas are derived from a standard Fama-French-Carhart 4-factor model. OLS regressions include both sector and year fixed effects. Heteroskedasticity-robust standard errors clustered at the firm level and Fama-MacBeth standard errors are shown in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

2.5.2 Operating Performance

In the second stage, we examine the impact of *Solar Potential* and the associated proxy for transition risk *Carbon Offset* on the operational performance of REITs. Table 2.3 presents the regression results with the outcome variables ROA and ROE. As reported in Columns (1) and (3), we find a positive and significant relationship between REITs' *Solar Potential* and both ROA (1% significance level) and ROE (10% significance level) in the Fama-MacBeth models, indicating that a greater solar potential leads to better operating performance. The estimated impact on firms' operating performance is economically significant. Considering a 50% increase in *Solar Potential* as depicted in the Fama-MacBeth model in Column (1), we observe a corresponding rise in ROA by 5.92 basis points. Similarly, this increase is even more pronounced for ROE in Column (3), resulting in a rise of 24.9 basis points.

Although the estimated coefficients in the fixed effects models are directionally consistent with those in the Fama-MacBeth models, we find no significant relationship between *Solar Potential* and either ROA or ROE. The results of the Fama-MacBeth model in Column (3) also indicate a significant relationship (at the 1% significance level) between a REIT's potential to offset carbon emissions of grid electricity and operating performance. However, this relationship is again not supported by the other models in Columns (1), (2), and (4). Despite some evidence in line with existing literature that sustainable investments positively influence REITs' operating performance, the overall statistical evidence in our models remains weak.

We attribute this weak evidence to two factors. First, a REIT's return should only be affected if solar panels are installed and generate additional revenue through direct and indirect cash flow channels; therefore, the potential for installation alone should not drive returns. The aforementioned special tax status of REITs and the lack of the ability to feed excess power into the grid are significant barriers to the widespread adoption of solar energy in the REIT sector. Nevertheless, we assume that our results are a reflection of a number of REITs already capitalizing on their potential. As highlighted by Lee and Liang (2024), the implementation of environmental regulations not only raises expectations within the real estate sector, but also forces companies to deepen their commitment to sustainability strategies, which may include exploiting their solar potential. Moreover, as noted by Chmutina et al. (2014), the adoption of solar systems is closely linked to environmental compliance and increasingly reflects the environmental concerns of market participants. Both of these factors may lead to solar installations gaining traction. We therefore expect an increasingly significant and more evident impact on operating performance in the coming years.

Second, notwithstanding the well-documented cash flow effects of sustainable investments at the property level, we do not anticipate a contemporaneous impact on operating performance at the firm level following the adoption of rooftop solar, particularly for REITs with larger property portfolios. As highlighted by Beracha et al. (2019), improving relative operating performance remains challenging for capital-intensive firms such as REITs. In our context, this could be particularly true for REITs implementing owner-operated strategies, as the upfront capital investment for solar systems could offset the operational benefits in the short term. In most cases, the installation of solar panels is part of a holistic approach to a REIT's sustainability strategy and is usually accompanied by other investments in greater technological automation of building operations, such as smart meters and energy efficiency measures. As these investments or the cost of these investments are typically borne solely by the landlord, any positive impact on earnings is further diluted in the short term.

2.5.3 Firm Value

Finally, we regress our firm value variables on *Solar Potential* and *Carbon Offset*. The results are reported in Table 2.4. Controlling for size, leverage, investment strategy, operating performance and diversification, we find strong evidence for a positive relationship between the *Solar Potential* of a REIT and both M/B and P/NAV ratios in our cross-sectional and fixed effects regressions. The coefficients for all of the models are statistically significant at the 5% level, the 1% level, and the 10% level, respectively. We now also observe a considerably higher R-squared across models. Increasing a REIT's average available solar area by 50% results in a 0.49% higher P/NAV ratio in the Fama-MacBeth model and a 0.73% higher P/NAV ratio in the fixed effects model, as presented in Columns (3) and (4). Given coefficients of 0.037 for the Fama-MacBeth model and 0.079 for the fixed effects model in Columns (1) and (2), the same increase in solar potential would result in an even more pronounced improvement in the M/B ratio of 1.51% and 3.26%, respectively. Taking into account the relatively high standard deviation of *Solar Potential* within our sample, the results can be considered economically significant.

Table 2.3: Regression Results for Operating Performance

Variables	ROA		ROE	
	(1)	(2)	(3)	(4)
Solar potential (log)	0.146***	0.001	0.615*	0.166
	(0.044)	(0.074)	(0.353)	(0.346)
Carbon offset (log)	−0.424	0.571	−9.164***	−4.565
	(0.269)	(0.803)	(2.182)	(5.175)
Total assets (log)	0.350***	0.410***	0.810***	0.940*
	(0.018)	(0.050)	(0.284)	(0.496)
Firm age (log)	−0.054*	−0.084	1.454***	1.382***
	(0.031)	(0.056)	(0.407)	(0.170)
Real estate Inv. growth$_{t-1}$	0.001	−0.005***	0.007	−0.022***
	(0.005)	(0.001)	(0.031)	(0.007)
D/E ratio$_{t-1}$	0.345***	0.234***	1.556***	1.510***
	(0.025)	(0.087)	(0.412)	(0.316)
Share buildings < 10 years	0.008**	0.011**	0.080***	0.080*
	(0.004)	(0.006)	(0.030)	(0.044)
Institutional ownership	0.002	0.004	0.240	0.314***
	(0.011)	(0.008)	(0.169)	(0.039)
GDP growth (MSA)	0.104***	0.098	1.597***	0.953
	(0.037)	(0.138)	(0.432)	(0.605)
HHI region	−0.898***	0.046	−2.897*	−1.628*
	(0.146)	(0.541)	(1.499)	(0.893)
HHI property type	−0.205	−0.429	−2.339	−2.726
	(0.141)	(0.736)	(2.407)	(3.087)
Constant	−2.040	−8.181	33.810*	4.498
	(1.648)	(5.196)	(17.585)	(37.277)
Coefficient estimates	FM	OLS	FM	OLS
Standard errors	FM	CL - F	FM	CL - F
Sector FE	N	Y	N	Y
Year FE	N	Y	N	Y
Observations	1,010	1,010	1,010	1,010
R-squared	0.189	0.322	0.151	0.161

The table reports the results of OLS and Fama-MacBeth regressions for the operating performance of U.S. equity REITs as a function of their solar potential and carbon offset and a vector of control variables. OLS regressions include both sector and year fixed effects. Heteroskedasticity-robust standard errors clustered at the firm level and Fama-MacBeth standard errors are shown in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Unlike most listed companies in other sectors, REITs are subject to a double cost of equity capital due to the active trading of their underlying assets in the secondary market (Boudry et al., 2010). As a result, the *Solar Potential* and associated asset-level benefits may also be incorporated into appraisals, potentially reducing the transition risk premium of property yields as suggested by Leskinen et al. (2020). The relationship between property and firm values, as reflected in M/B and P/NAV, provides an understanding of the extent to which markets consider transition risk mitigation potential at both levels. The identified positive impact on both measures suggests that there are channels through which higher levels of *Solar Potential* affect valuation at the firm level beyond the book and market values of the underlying properties. Investors may therefore assume that solar systems enable management to increase a REIT's earnings over its existing property portfolio. As expected, we observe a larger impact on the M/B ratio, primarily due to the fact that depreciated book values, unlike Net Asset Values, are less able to capture the potential benefits of sustainable features in property valuations.

Since corporate valuation is inherently forward-looking, the excess firm values can be seen as a reflection of investors' assumptions on REITs' future earnings. One possible reason for our findings is that investors might reasonably expect the benefits of on-site renewable energy generation to have an increasingly positive impact on property investments through income channels as the aforementioned barriers to REITs' installation of solar systems diminish. In addition to providing higher and more stable cash flows, REITs' *Solar Potential* is expected to influence growth rates of future earnings. This is particularly true for REITs with greater exposure to grids with higher electricity rates, where tenants benefit relatively more from energy cost savings related to the procurement of on-site renewable energy. Therefore, higher effective rental growth rates can be realized. This potential becomes even more important in states where the accelerating rise in electricity prices outpaces rental growth.

Higher firm values may also result from lower discount rates associated with a REIT's future earnings, which reflect lower transition risks due to higher on-site renewable energy potential. As transition risks are increasingly priced into commercial real estate markets, investors may face lower required returns for REITs with higher *Solar Potential*. Differences in return requirements are already evident among investors with a strong focus on environmentally sustainable properties, creating a clientele effect that is driving a valuation premium for greener assets (Fuerst et al., 2017). In addition, debt markets are placing greater emphasis on carbon performance and seeking to reduce their own financed emissions by incentivizing greener properties (An & Pivo, 2020; Devine et al., 2024; Eichholtz, Holtermans, Kok, & Yönder, 2019; Leutner et al., 2024).

As a result, investors expect lower refinancing risk for REITs with greater potential access green financing channels. While investors and lenders are expanding their criteria for evaluating the environmental performance of REITs to include indicators such as on-site renewable energy sources, we do not expect *Solar Potential* to have a significant signaling effect on firm values as it is less prominently communicated externally by most REITs, unlike green building certifications or benchmarking scores.

We also find a significant negative relationship between *Carbon Offset* and our dependent variables in both the Fama-MacBeth and fixed effects models. Assuming a 10% increase in *Carbon Offset*, which implies a higher exposure to carbon-intensive grids, the M/B ratio experiences a notable reduction as can be seen in Columns (1) and (2). In the fixed effects model, with a coefficient of -0.734, the average decrease is 6.76%. The results of the Fama-MacBeth model show an even stronger negative impact. Although less pronounced, the effect on P/NAV lies also in the same direction. Presented in Column (3), a 10% increase in the carbon intensity of the grid leads to a slight reduction in P/NAV of 0.97%. The difference in the size of the change in our two dependent variables could be due to the fact that a rise in electricity rates from grids with a higher share of fossil fuels has already been factored into Net Asset Values.

Our models therefore suggest that REIT values are negatively affected by higher exposure to carbon-intensive grids. This is consistent with the findings of Bolton and Kacperczyk (2021), emphasizing that companies more dependent on fossil fuel-intensive energy sources are more vulnerable to technological transition risks posed by the emergence of cheaper renewable energy. Moreover, our results are in line with previous research showing that upstream emissions for which companies have no direct responsibility pose a significant transition risk for REITs. In the context of our primary explanatory variable, *Solar Potential*, our results underscore the sizable benefits to REITs of adopting solar systems. Importantly, the significant results for our two explanatory variables, *Solar Potential* and *Carbon Offset*, both in models with and without sector fixed effects, imply that transition risks and the ability to mitigate them through solar installations affect firm values across the REIT industry rather than individual sectors.

2.6 Robustness Checks

To validate our empirical results and therefore our research conclusions, we run a number of robustness checks. First, we perform a subsample analysis to assess the consistency of our findings across different firm sizes. As the solar potential of REITs should be unaffected by their financial performance, we do not expect reverse causality issues, widely discussed in the sustainability literature. However, it is conceivable that the market values the solar potential differently depending on the size of the company. It tends to be the larger REITs that publicly commit to net-zero goals and disclose more detailed information about their environmental performance, making them particularly attractive to green investors and funds. In addition, larger REITs may be able to take advantage of economies of scale in the implementation of solar strategies and benefit from more resources, especially in light of the challenges discussed above. To facilitate this analysis, we construct two portfolios by sorting REITs based on whether their total assets are above or below the median observed in our sample.

Second, we assess the reliability of our cross-sectional regressions by calculating coefficients using standard errors clustered by time. This method, as shown by Petersen (2008), yields standard errors that are similar to Fama-MacBeth standard errors when properly specified, as both approaches account for cross-sectionally correlated residuals and the presence of time effects. It allows to detect a potential bias in our cross-sectional regressions due to potential firm effects significantly outweighing time effects.

Third, we examine the robustness of our results for firm values in the fixed effects models to potential unobserved confounders that may lead to omitted variable bias, an issue that has been extensively discussed in the existing literature on environmental-sustainable investments and firm performance. We therefore conduct a sensitivity analysis following Cinelli and Hazlett (2020), which measures how influential a potential confounder needs to be in order to significantly reduce the impact of our independent variables and thus cause a change in our research results.

Table 2.4: Regression Results for Firm Values

Variables	M/B ratio (log)		P/NAV ratio (log)	
	(1)	(2)	(3)	(4)
Solar potential (log)	0.037**	0.079***	0.012*	0.018***
	(0.017)	(0.025)	(0.006)	(0.004)
Carbon offset (log)	−1.037***	−0.734***	−0.102***	−0.062
	(0.077)	(0.159)	(0.034)	(0.048)
Total assets (log)	0.007	0.027*	−0.012**	−0.008***
	(0.009)	(0.015)	(0.005)	(0.002)
Firm age (log)	0.165***	0.128***	0.032***	0.020***
	(0.013)	(0.027)	(0.007)	(0.004)
Real estate Inv. growth$_{t-1}$	0.005***	0.002*	0.001***	0.001***
	(0.001)	(0.001)	(0.0004)	(0.00005)
D/E ratio$_{t-1}$	0.242***	0.222***	−0.010*	−0.017***
	(0.011)	(0.028)	(0.005)	(0.001)
ROA$_{t-1}$	0.128***	0.096***	0.012***	0.007***
	(0.016)	(0.015)	(0.003)	(0.0002)
Share buildings < 10 years	0.002**	0.003	−0.0003	−0.0002
	(0.001)	(0.003)	(0.001)	(0.0002)
Institutional ownership	−0.003	−0.008**	0.004***	0.003***
	(0.006)	(0.004)	(0.001)	(0.001)
GDP growth (MSA)	0.028	0.007	0.0003	−0.004
	(0.018)	(0.047)	(0.005)	(0.011)
HHI region	−0.120	−0.118	−0.094***	−0.059*
	(0.076)	(0.123)	(0.036)	(0.036)
HHI property type	0.073	−0.108	0.018	0.012
	(0.057)	(0.127)	(0.024)	(0.027)
Constant	5.671***	3.748***	5.071***	4.884***
	(0.474)	(1.221)	(0.229)	(0.260)
Coefficient estimates	FM	OLS	FM	OLS
Standard errors	FM	CL - F	FM	CL - F
Sector FE	N	Y	N	Y
Year FE	N	Y	N	Y
Observations	1,010	1,010	1,010	1,010
R-squared	0.435	0.503	0.376	0.354

The table reports the results of OLS and Fama-MacBeth regressions for the firm value variables of U.S. equity REITs as a function of their solar potential and carbon offset and a vector of control variables. OLS regressions include both sector and year fixed effects. Heteroskedasticity-robust standard errors clustered at the firm level and Fama-MacBeth standard errors are shown in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 2.5 shows the results of our subsample analysis. While we lose some statistical significance for both *Solar Potential* and *Carbon Offset* in the models shown in Columns (2) and (7) compared to our initial coefficient estimates, the results do not support the hypothesis that *Solar Potential* may command a larger firm value premium for larger REITs. Instead, since both smaller and larger REITs are affected by the observed changes and the significance of the control variables in the models decreases for both subsamples, we attribute these results to the reduced sample size within our portfolios. In the other models, however, we find significant coefficient estimates for both *Solar Potential* and *Carbon Offset* that are consistent with the size and direction of our initial estimates shown in Table 2.4. Importantly, our overall results remain consistent even when the sample size is reduced by half, indicating that the observed estimates are not driven by heterogeneity and specific portions in the data. This suggests that the market prices in the ability to mitigate transition risks and thus financially benefit from future rooftop solar deployment, irrespective of firm size.

The results of the OLS models with time-clustered standard errors are reported in Table 2.6, showing that both the size and direction of the coefficient estimates and standard errors are consistent with the results of the Fama-MacBeth regressions. As highlighted by Petersen (2008), the estimates obtained from OLS and Fama-MacBeth regressions may be very similar even when the bias due to a firm effect is large. However, the author also notes that in the presence of a bias, the Fama-MacBeth model in particular would significantly underestimate true standard errors, to an even greater extent than the OLS model, which is not the case in our results. Comparing both the Fama-MacBeth and time-clustered standard errors with the firm-clustered standard errors of the fixed effects models in Tables 2.2, 2.3 and 2.4, the latter tend to be higher, although the difference is not large. While the variation in standard errors between the results suggests a somewhat stronger influence of the firm effect, which is to be expected given the structure of our dataset, we consider the estimates of the Fama-MacBeth models to be robust given the size of the effect.

Table 2.5: Subsample Analysis for Firm Values

Variables	M/B ratio (log)		P/NAV ratio (log)		M/B ratio (log)		P/NAV ratio (log)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Solar potential (log)	0.067***	0.032	0.019**	0.041***	0.064*	0.028**	−0.006	0.018**
	(0.022)	(0.020)	(0.009)	(0.006)	(0.033)	(0.017)	(0.014)	(0.007)
Carbon offset (log)	−0.935***	−0.271*	−0.093***	−0.048	−1.016***	−0.627***	−0.034	−0.089
	(0.145)	(0.146)	(0.035)	(0.085)	(0.079)	(0.226)	(0.064)	(0.073)
Total assets (log)	0.079***	0.058***	−0.012	0.0004	0.165***	0.188***	0.015	0.026**
	(0.024)	(0.012)	(0.008)	(0.004)	(0.013)	(0.017)	(0.012)	(0.011)
Firm age (log)	0.166***	0.180***	0.040**	0.016***	0.145***	0.097***	0.028**	0.025***
	(0.020)	(0.033)	(0.021)	(0.006)	(0.018)	(0.019)	(0.011)	(0.003)
Real estate Inv. growth$_{t-1}$	0.006*	0.002*	0.001**	0.0003***	0.005**	0.0004	0.001*	0.0004***
	(0.003)	(0.001)	(0.0003)	(0.0001)	(0.002)	(0.001)	(0.0003)	(0.0001)
D/E ratio$_{t-1}$	0.220***	0.169***	−0.019***	−0.024***	0.177***	0.182***	−0.014*	−0.027***
	(0.026)	(0.034)	(0.006)	(0.002)	(0.036)	(0.027)	(0.008)	(0.002)
ROA$_{t-1}$	0.061**	0.050***	0.002	−0.002	0.218***	0.160***	0.019***	0.009***
	(0.029)	(0.018)	(0.005)	(0.002)	(0.021)	(0.011)	(0.004)	(0.001)
Share buildings < 10 years	0.006*	0.009*	−0.0001	0.0005	−0.001	−0.002	−0.001*	−0.001**
	(0.003)	(0.005)	(0.001)	(0.0004)	(0.002)	(0.001)	(0.0004)	(0.0003)
Institutional ownership	−0.001	−0.001	0.006***	0.002***	−0.001	−0.002	−0.001*	−0.001**
	(0.010)	(0.004)	(0.002)	(0.0004)	(0.012)	(0.009)	(0.003)	(0.001)
GDP growth (MSA)	−0.058**	−0.006	0.003	−0.008	0.155***	0.077*	−0.005	0.013
	(0.029)	(0.041)	(0.006)	(0.012)	(0.016)	(0.040)	(0.008)	(0.014)
HHI region	0.446***	0.735***	−0.026	−0.040	0.046	−0.074	−0.124***	−0.014
	(0.114)	(0.284)	(0.036)	(0.052)	(0.137)	(0.169)	(0.040)	(0.031)
HHI property type	0.157	−0.231	0.017	0.037	0.484***	0.147***	−0.006	0.087***
	(0.110)	(0.143)	(0.025)	(0.038)	(0.169)	(0.048)	(0.057)	(0.033)
Constant	3.948***	0.770	4.916***	4.552***	2.592***	0.799	4.403***	4.390***
	(0.786)	(0.773)	(0.269)	(0.542)	(0.562)	(0.901)	(0.372)	(0.473)
Subsample	< Median	< Median	< Median	< Median	> Median	> Median	> Median	> Median
Coefficient estimates	FM	OLS	FM	OLS	FM	OLS	FM	OLS
Standard errors	FM	CL - F	FM	CL - F	FM	CL - F	FM	CL - F
Sector FE	N	Y	N	Y	N	Y	N	Y
Year FE	N	Y	N	Y	N	Y	N	Y
Observations	505	505	505	505	505	505	505	505
R-squared	0.450	0.487	0.411	0.369	0.610	0.687	0.463	0.445

The table shows the regression results of the subsample analysis for firm values of U.S. equity REITs as a function of their solar potential and carbon offset and vector of control variables. Subsamples are defined according to the median firm size in our sample as measured by total assets. OLS regressions include both sector and year fixed effects. Heteroskedasticity-robust standard errors clustered at the firm level and Fama-MacBeth standard errors are shown in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 2.6: Robustness Checks for Cross-Sectional Regressions

Variables	Alpha	Beta (log)	ROA	ROE	M/B ratio (log)	P/NAV ratio (log)
	(1)	(2)	(3)	(4)	(5)	(6)
Solar potential (log)	0.226	0.022	0.159^{***}	0.723^{**}	0.046^{***}	0.012^{**}
	(0.341)	(0.019)	(0.044)	(0.368)	(0.015)	(0.006)
Carbon offset (log)	-3.522^{*}	0.206	-0.584^{**}	-9.891^{***}	-1.154^{***}	-0.124^{***}
	(2.075)	(0.155)	(0.255)	(1.742)	(0.089)	(0.038)
Total assets (log)	-0.707^{**}	0.043	0.343^{***}	0.851^{***}	-0.004	-0.014^{**}
	(0.321)	(0.028)	(0.022)	(0.242)	(0.012)	(0.006)
Firm age (log)	0.215	0.012	-0.105^{***}	0.974^{***}	0.150^{***}	0.020^{***}
	(0.393)	(0.034)	(0.028)	(0.374)	(0.012)	(0.007)
Real estate Inv. growth$_{t-1}$	0.046^{**}	-0.002	-0.001	0.002	0.004^{***}	0.001^{***}
	(0.019)	(0.001)	(0.003)	(0.022)	(0.001)	(0.0003)
D/E ratio$_{t-1}$	0.302	0.019^{*}	0.310^{***}	1.416^{***}	0.237^{***}	-0.011^{*}
	(0.214)	(0.010)	(0.026)	(0.344)	(0.013)	(0.006)
ROA$_{t-1}$	0.518^{*}	-0.021			0.128^{***}	0.013^{***}
	(0.301)	(0.020)			(0.012)	(0.005)
Share buildings < 10 years	-0.022	-0.001	0.006^{*}	0.068^{**}	0.002^{*}	-0.0004
	(0.026)	(0.002)	(0.004)	(0.030)	(0.001)	(0.001)
GDP growth (MSA)	0.559	-0.036	0.120^{***}	1.756^{***}	0.041^{***}	0.002
	(0.424)	(0.040)	(0.038)	(0.386)	(0.016)	(0.005)
Institutional ownership	0.022	0.018^{***}	0.004	0.311^{**}	-0.0003	0.003^{***}
	(0.095)	(0.006)	(0.010)	(0.132)	(0.006)	(0.001)
HHI region	-2.801	0.219^{***}	-0.958^{***}	-2.790^{**}	-0.166^{**}	-0.110^{***}
	(1.994)	(0.083)	(0.164)	(1.333)	(0.084)	(0.034)
HHI property type	1.539	0.253^{***}	-0.229^{*}	-1.949	0.077	0.005
	(1.252)	(0.085)	(0.138)	(2.336)	(0.051)	(0.023)
Constant	26.724^{*}	-2.864^{***}	-0.838	36.838^{***}	6.482^{***}	5.298^{***}
	(15.468)	(1.004)	(1.322)	(13.618)	(0.657)	(0.268)
Coefficient estimates	OLS	OLS	OLS	OLS	OLS	OLS
Standard errors	CL - T	CL - T	CL - T	CL - T	CL - T	CL - T
Sector FE	N	N	N	N	N	N
Year FE	N	N	N	N	N	N
Observations	1,010	1,010	1,010	1,010	1,010	1,010
R-squared	0.049	0.045	0.181	0.091	0.390	0.105

The table shows the regression results of the robustness checks for stock performance, operating performance and firm value of U.S. equity REITs as a function of their solar potential and carbon offset and vector of control variables. For direct comparison with the Fama-MacBeth regression results, the coefficients are estimated without time and sector fixed effects. Time-clustered robust standard errors are presented in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively

Tables 2.7 and 2.8 present the sensitivity statistics for our independent variables P/NAV and Price to Book (P/B). We estimate the strength of potential confounders using statistically significant covariates with high explanatory power for firm values identified in our fixed effects regressions. The robustness bounds $R^2_{Y\sim Z|\mathbf{X},D}$ and $R^2_{D\sim Z|\mathbf{X}}$ indicate the maximum residual variation in the outcome and treatment variables that could be explained by a omitted variable in an extreme scenario. The strength of a potential omitted variable is assessed by the explanatory power of the controls, firm age, ROA, and real estate investment growth. Comparing the robustness bounds to the robustness values ($RV_{q=1}$) of our key independent variables, we find the robustness values exceeding the robustness bounds in all setups. It can be seen that only the robustness bound of ROA indicates that a possible confounder in an extreme scenario could have a significant impact on the coefficient estimate of *Solar Potential*. However, overall results suggest that potential omitted variables may not be strong enough to fully explain the effects of *Solar Potential* and *Carbon Offset*. Figure 2.1 shows sensitivity plots of the point estimates of our key independent variables with a potential confounder as strong as ROA. For this analysis, we specifically chose ROA because, as noted above, this covariate has the highest confounding power in extreme scenarios. Similar to the sensitivity statistics, the sensitivity plots show high robustness in both the strength and direction of the point estimates of *Solar Potential* and *Carbon Offset* to a potential omitted variable that is k times stronger than ROA. Even in an extreme scenario where a potential confounder is three times stronger than ROA, only the point estimate of *Solar Potential* with the outcome variable P/B ratio would experience a change in coefficient direction. In addition, the vertical axis of the sensitivity plots indicates that the potential confounders and control variables in our models have significantly higher explanatory power for the residual variation of firm values than for the residual variation of our explanatory variables. The results therefore suggest that *Solar Potential* and *Carbon Offset* are robust to potential omitted variables and, importantly, they are not merely proxies for omitted financial firm characteristics.

Table 2.7: Sensitivity Statistics for P/NAV Ratio

Outcome: *P/NAV ratio*						
Treatment:	Est.	S.E.	t-value	$R^2_{Y\sim D\|\mathbf{X}}$	$RV_{q=1}$	$RV_{q=1,\alpha=0.05}$
Solar potential (log)	0.012	0.005	2.238	0.5%	6.8%	0.9%
df = 997	*Bound 1x*	*Firm age (log)*: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 1.3%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.4%				
Carbon offset (log)	-0.124	0.045	-2.767	0.8%	8.4%	2.5%
df = 997	*Bound 1x*	*Firm age (log)*: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 1.3%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0%				
Solar potential (log)	0.012	0.005	2.238	0.5%	6.8%	0.9%
df = 997	*Bound 1x*	ROA_{t-1}: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 1.8%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.9%				
Carbon offset (log)	-0.124	0.045	-2.767	0.8%	8.4%	2.5%
df = 997	*Bound 1x*	ROA_{t-1}: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 1.8%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.2%				
Solar potential (log)	0.012	0.005	2.238	0.5%	6.8%	0.9%
df = 997	*Bound 1x*	*RE inv. growth*$_{t-1}$: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 2.2%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.7%				
Carbon offset (log)	-0.124	0.045	-2.767	0.8%	8.4%	2.5%
df = 997	*Bound 1x*	*RE inv. growth*$_{t-1}$: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 2.1%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0%				

The table shows the robustness report with the sensitivity statistics for the outcome variable P/NAV ratio. It includes the partial R^2 of the treatment with the outcome and the robustness value $RV_{q=1}$. The bounds indicate the strength of an unobserved confounder that has the same effect on the treatment variables (solar potential or carbon offset) as the observed covariates. An *RV* above either bound indicates that a potential confounder is not strong enough to remove the explanatory power of the treatment variables.

Table 2.8: Sensitivity Statistics for P/B Ratio

			Outcome: *P/B ratio*			
Treatment:	Est.	S.E.	t-value	$R^2_{Y\sim D\|\mathbf{X}}$	$RV_{q=1}$	$RV_{q=1,\alpha=0.05}$
Solar potential (log)	0.046	0.019	2.456	0.6%	7.5%	1.5%
df = 997	*Bound 1x*		*Firm age (log)*: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 5.6%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.4%			
Carbon offset (log)	-1.154	0.159	-7.246	5%	20.5%	15.4%
df = 997	*Bound 1x*		*Firm age (log)*: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 5.6%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0%			
Solar potential (log)	0.046	0.019	2.456	0.6%	7.5%	1.5%
df = 997	*Bound 1x*		ROA_{t-1}: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 13.8%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.9%			
Carbon offset (log)	-1.154	0.159	-7.246	5%	20.5%	15.4%
df = 997	*Bound 1x*		ROA_{t-1}: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 13.6%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.2%			
Solar potential (log)	0.046	0.019	2.456	0.6%	7.5%	1.5%
df = 997	*Bound 1x*		*RE inv. growth*$_{t-1}$: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 2.8%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0.7%			
Carbon offset (log)	-1.154	0.159	-7.246	5%	20.5%	15.4%
df = 997	*Bound 1x*		*RE inv. growth*$_{t-1}$: $R^2_{Y\sim Z\|\mathbf{X},D}$ = 2.8%, $R^2_{D\sim Z\|\mathbf{X}}$ = 0%			

The table shows the robustness report with the sensitivity statistics for the outcome variable P/B ratio. It includes the partial R^2 of the treatment with the outcome and the robustness value $RV_{q=1}$. The bounds indicate the strength of an unobserved confounder that has the same effect on the treatment variables (solar potential or carbon offset) as the observed covariates. An *RV* above either bound indicates that a potential confounder is not strong enough to remove the explanatory power of the treatment variables.

2.7 Concluding Remarks

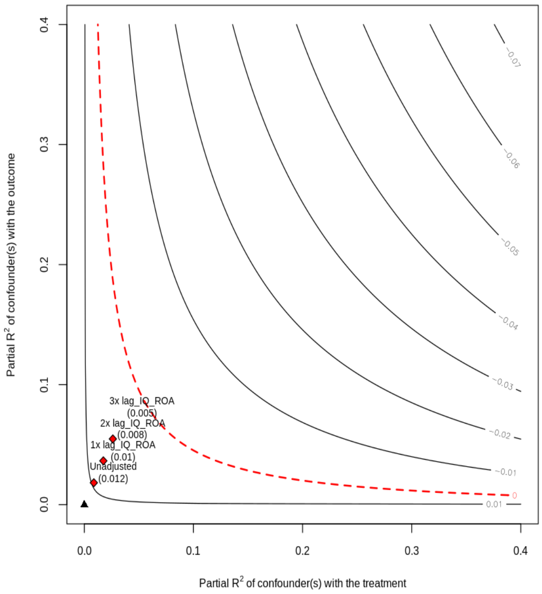
Building on the existing literature showing the profound impact of carbon emissions and mitigation strategies on asset prices, our analysis aims to deepen the understanding of the drivers of (carbon) transition risks and their impact on the financial performance of REITs. Our approach not only adds to the existing REIT literature, but also extends the scope of environmental performance metrics used to date beyond traditional benchmarking and green building certification schemes. To do this, we leverage a unique dataset that combines REITs' available rooftop space for solar installations with a REIT-specific carbon offset factor. This factor captures a company's exposure to carbon-intensive electricity grids, indicating the potential for solar power to reduce carbon emissions by displacing grid electricity. We provide a first look at the extent to which the rooftop solar potential of REITs can be considered as a driver of stock performance, operating performance, and firm values.

Our analysis of firms' stock performance suggests that the market does not currently fully reflect REITs' solar adoption potential, as evidenced by the lack of significant impact on alphas and market betas. We attribute the primary reason for our results to the fact that REITs have not yet adopted rooftop solar on a large scale. This reluctance is largely due to the presence of regulatory hurdles that have impeded progress in this area. Limited adoption also likely contributes to the weak evidence in our models regarding the impact on REITs' operating performance. In addition, the initial capital investment typically required for owner-operated solar strategies and related building technology upgrades may outweigh the immediate increase in revenue, further dampening the short-term cash flow effects.

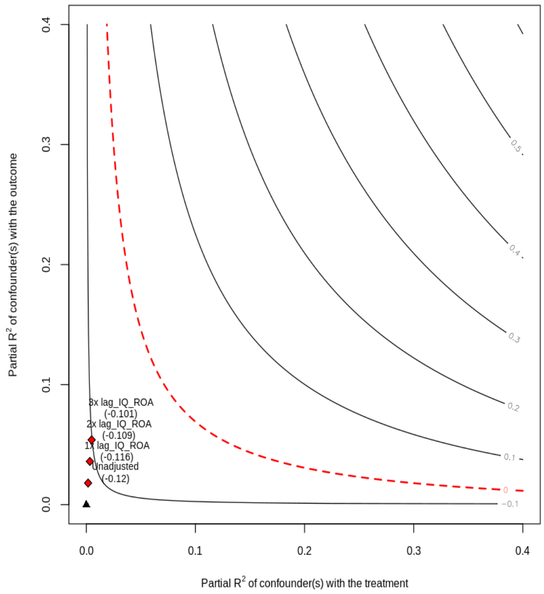
We find strong evidence that the solar potential of REITs has a significant impact on firm values. The results show a discernible valuation effect that extends beyond the market and book values of the underlying property portfolios. Importantly, this effect is not a simple proxy for omitted financial characteristics of firms. It implies that the market places a higher value on REITs' future earnings due to the financial benefits associated with solar. At the same time, higher valuations may result from the market's lower required rate of return in anticipation of a reduction in transition risk. In addition, the significant relationship between carbon offset factors and corporate valuation supports existing evidence that REITs face transition risks, not only from their direct emissions, but also from upstream emissions beyond firms' direct control.

While rooftop solar may be only one potential facet of a REIT's strategy to mitigate transition risks, our findings underscore existing evidence that the market is pricing in management's ability to engage in environmental practices and thereby realize financial benefits. In other words, REITs that fail to integrate these considerations into their business model will face negative financial implications at both asset and firm level. Policymakers could take advantage of these market dynamics to harness the enormous solar potential of the REIT industry and facilitate the decarbonization of the property sector by further reducing barriers and incentivizing the installation of solar systems. Moreover, with the emergence of community rooftop solar projects, there could be mutual benefits for REITs as well as other companies and residents who benefit from greener and more affordable energy.

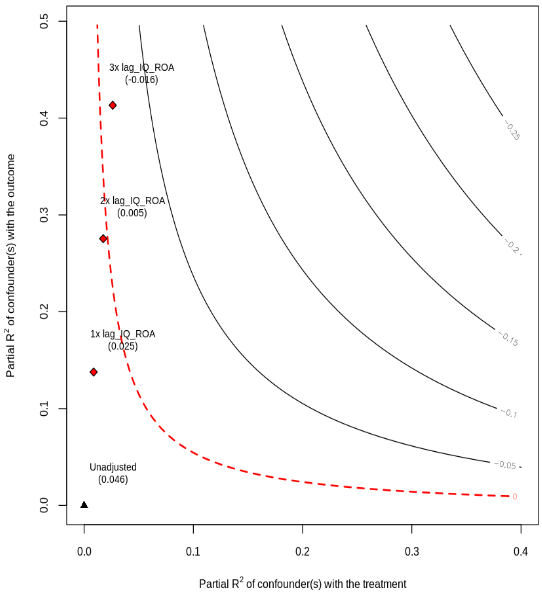
Overall, our results suggest a significant and robust relationship between REITs' solar potential and firm values, although there are some limitations to our data. As more REITs implement solar strategies, data on realized potential could provide deeper insights into the income and capital channels through which firms may derive financial benefits. As the adoption of solar continues to grow, we anticipate that future research will reveal greater impacts on both stock and operating performance of REITs. In addition, further analysis could shed more light on the extent to which REIT sectors with more limited rooftop solar potential could also leverage alternative renewable energy strategies to mitigate transition risks at the firm level, such as the use of solar facades or investing in off-site renewable energy projects.



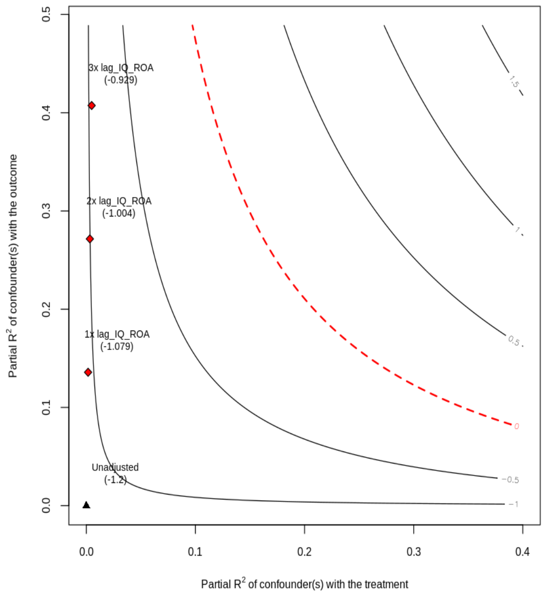
(a) Point estimate of solar potential with the outcome variable P/NAV ratio



(b) Point estimate of carbon offset with the outcome variable P/NAV ratio



(c) Point estimate of solar potential with the outcome variable P/B ratio



(d) Point estimate of carbon offset with the outcome variable P/B ratio

Figure 2.1: Sensitivity contour plots in the partial R^2 scale of unobserved confounders. The figure plots the sensitivity of the point estimates to potential confounders k times stronger than ROA. The horizontal axis shows how strongly the confounder is associated with the treatment variables (solar potential or carbon offset). The vertical axis shows how strongly the confounder is associated with the outcome variables (P/NAV ratio or P/B ratio).

2.8 References

Akhtyrska, Y., & Fuerst, F. (2024). The effectiveness of climate change regulations in the commercial real estate market. *Energy Policy*, *185*, Article 113916.

Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, *65*(10), 4451–4469.

An, X., & Pivo, G. (2020). Green buildings in commercial mortgage-backed securities: The effects of LEED and energy star certification on default risk and loan terms. *Real Estate Economics*, *48*(1), 7-42.

Aydin, E., Eichholtz, P., & Yönder, E. (2018). The economics of residential solar water heaters in emerging economies: The case of Turkey. *Energy Economics*, *75*, 285–299.

Bauer, R., Eichholtz, P., & Kok, N. (2010). Corporate governance and performance: The REIT effect. *Real Estate Economics*, *38*(1), 1–29.

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review*, *89*(3), 634–647.

Beracha, E., Feng, Z., & Hardin, W. G. (2019). REIT operational efficiency: Performance, risk, and return. *The Journal of Real Estate Finance and Economics*, *58*, 408–437.

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, *142*(2), 517–549.

Bolton, P., & Kacperczyk, M. (2023). Global pricing of carbon-transition risk. *The Journal of Finance*, *78*(6), 3677–3754.

Bond, S. A., & Devine, A. (2016). Certification matters: Is green talk cheap talk? *The Journal of Real Estate Finance and Economics*, *52*(2), 117–140.

Boudry, W. I., Kallberg, J. G., & Liu, C. H. (2010). An analysis of REIT security issuance decisions. *Real Estate Economics*, *38*(1), 91–120.

Brown, L. D., & Caylor, M. L. (2009). Corporate governance and firm operating performance. *Review of Quantitative Finance and Accounting*, *32*, 129–144.

Capozza, D. R., & Seguin, P. J. (1999). Focus, transparency and value: The REIT evidence. *Real Estate Economics*, *27*(4), 587–619.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Chacon, R., Feng, Z., & Wu, Z. (2024). Does investing in ESG pay off? Evidence from REITs. *Journal of Real Estate Finance and Economics*.

Chava, S. (2014). Environmental externalities and cost of capital. *Management Science*, *60*(9), 2223–2247.

Chegut, A., Eichholtz, P., & Kok, N. (2014). Supply, demand and the value of green buildings. *Urban studies*, *51*(1), 22–43.

Chmutina, K., Wiersma, B., Goodier, C. I., & Devine-Wright, P. (2014). Concern or compliance? Drivers of urban decentralised energy initiatives. *Sustainable Cities and Society*, *10*, 122–129.

Choi, D., Gao, Z., & Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, *33*(3), 1112–1145.

Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(1), 39–67.

Clayton, J., Devine, A., & Holtermans, R. (2021). Beyond building certification: The impact of environmental interventions on commercial real estate operations. *Energy Economics*, *93*, 105039.

Core, J. E., Guay, W. R., & Rusticus, T. O. (2006). Does weak governance cause weak stock returns? An examination of firm operating performance and investors' expectations. *The Journal of Finance*, *61*(2), 655–687.

Dastrup, S. R., Zivin, J. G., Costa, D. L., & Kahn, M. E. (2012). Understanding the solar home price premium: Electricity generation and "green" social status. *European Economic Review*, *56*(5), 961–973.

Devine, A., & Kok, N. (2015). Green certification and building performance: Implications for tangibles and intangibles. *The Journal of Portfolio Management*, *41*(6), 151–163.

Devine, A., Kok, N., Özgür, H., & Yönder, E. (2024). Green bonds: Environmental performance and capital market response. *RERI Working Paper*.

Devine, A., Sanderford, A., & Wang, C. (2022). Sustainability and private equity real estate returns. *The Journal of Real Estate Finance and Economics*, 1–27.

Devine, A., & Yönder, E. (2023). Impact of environmental investments on corporate financial performance: Decomposing valuation and cash flow effects. *The Journal of Real Estate Finance and Economics*, *66*(4), 778–805.

Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, *131*(3), 693–714.

Eichholtz, P., Holtermans, R., & Kok, N. (2019). Environmental performance of commercial real estate: New insights into energy efficiency improvements. *The Journal of Portfolio Management*, *45*(7), 113–129.

Eichholtz, P., Holtermans, R., Kok, N., & Yönder, E. (2019). Environmental performance and the cost of debt: Evidence from commercial mortgages and REIT bonds. *Journal of Banking & Finance*, *102*, 19–32.

Eichholtz, P., Kok, N., & Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *American Economic Review*, *100*(5), 2492–2509.

Eichholtz, P., Kok, N., & Quigley, J. M. (2013). The economics of green building. *Review of Economics and Statistics*, *95*(1), 50–63.

Eichholtz, P., Kok, N., & Yonder, E. (2012). Portfolio greenness and the financial performance of REITs. *Journal of International Money and Finance*, *31*(7), 1911–1929.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, *50*(1), 131–155.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.

Feng, Z., Hardin III, W. G., & Wang, C. (2022). Rewarding a long-term investment strategy: REITs. *Journal of Real Estate Research*, *44*(1), 56–79.

Feng, Z., & Wu, Z. (2023). ESG disclosure, REIT debt financing and firm value. *The Journal of Real Estate Finance and Economics*, *67*(3), 388–422.

Fleten, S.-E., Maribu, K. M., & Wangensteen, I. (2007). Optimal investment strategies in decentralized renewable power generation under uncertainty. *Energy*, *32*(5), 803–815.

Fuerst, F. (2015). The financial rewards of sustainability: A global performance study of real estate investment trusts. *Available at SSRN 2619434*.

Fuerst, F., Gabrieli, T., & McAllister, P. (2017). A green winner's curse? investor behavior in the market for eco-certified office buildings. *Economic Modelling*, *61*, 137–146.

Fuerst, F., & McAllister, P. (2011). Green noise or green value? Measuring the effects of environmental certification on office values. *Real Estate Economics*, *39*(1), 45–69.

Fuss, S., Johansson, D. J., Szolgayova, J., & Obersteiner, M. (2009). Impact of climate policy uncertainty on the adoption of electricity generating technologies. *Energy Policy*, *37*(2), 733–743.

Hoen, B., Wiser, R., Thayer, M., & Cappers, P. (2013). Residential photovoltaic energy systems in california: The effect on home sales prices. *Contemporary Economic Policy*, *31*(4), 708–718.

Holtermans, R., & Kok, N. (2019). On the value of environmental certification in the commercial real estate market. *Real Estate Economics*, *47*(3), 685–722.

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon tail risk. *The Review of Financial Studies*, *34*(3), 1540–1571.

Jurasz, J., & Campana, P. E. (2019). The potential of photovoltaic systems to reduce energy costs for office buildings in time-dependent and peak-load-dependent tariffs. *Sustainable Cities and Society*, *44*, 871–879.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, *33*(3), 1067–1111.

Lee, C. L., & Liang, J. (2024). The effect of carbon regulation initiatives on corporate ESG performance in real estate sector: International evidence. *Journal of Cleaner Production*, Article 142188.

Leskinen, N., Vimpari, J., & Junnila, S. (2020). Using real estate market fundamentals to determine the correct discount rate for decentralised energy investments. *Sustainable Cities and Society*, *53*, Article 101953.

Leutner, S., Gloria, B., & Bienert, S. (2024). Is there a green discount in commercial real estate lending? *Journal of Property Investment & Finance*.

Matsumura, E. M., Prakash, R., & Vera-Munoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *The Accounting Review*, *89*(2), 695–724.

Newsham, G. R., Mancini, S., & Birt, B. J. (2009). Do LEED-certified buildings save energy? Yes, but.... *Energy and Buildings*, *41*(8), 897–905.

Petersen, M. A. (2008). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, *22*(1), 435–480.

Qiu, Y., Wang, Y. D., & Wang, J. (2017). Soak up the sun: Impact of solar energy systems on residential home values in Arizona. *Energy Economics*, *66*, 328–336.

Reichardt, A. (2014). Operating expenses and the rent premium of energy star and LEED certified buildings in the central and eastern US. *The Journal of Real Estate Finance and Economics*, *49*, 413–433.

Reichardt, A., Fuerst, F., Rottke, N., & Zietz, J. (2012). Sustainable building certification and the rent premium: a panel data approach. *Journal of Real Estate Research*, *34*(1), 99–126.

Sah, V., Miller, N., & Ghosh, B. (2013). Are green REITs valued more? *Journal of Real Estate Portfolio Management*, *19*(2), 169–177.

Salisu, A. A., Ogbonna, A. E., & Vo, X. V. (2024). Climate risks and the REITs market. *International Journal of Finance & Economics*.

Seltzer, L. H., Starks, L., & Zhu, Q. (2022). Climate regulatory risk and corporate bonds. *National Bureau of Economic Research Staff Reports*.

Vincent, L. (1999). The information content of funds from operations (FFO) for real estate investment trusts (REITs). *Journal of Accounting and Economics*, *26*(1-3), 69–104.

Wiley, J. A., Benefield, J. D., & Johnson, K. H. (2010). Green design and the market for commercial office space. *The Journal of Real Estate Finance and Economics*, *41*, 228–243.

3 Picture This: A Deep Learning Model for Operational Real Estate Emissions

3.1 Abstract

We present a deep learning model estimating CO_2e emissions in the real estate sector. The model, which utilizes Convolutional Neural Networks (CNNs) and image classification techniques, is designed to estimate CO_2e emissions based on publicly available images of buildings and their corresponding emissions. Our findings show that the model has the ability to provide reasonably accurate estimations of CO_2e emissions using images as the sole input. Notably, incorporating primary energy sources as additional input further improves the accuracy up to 75%. The creation of such a model is particularly important in the fight against climate change, as it allows for transparency and fast identification of buildings, contributing significantly to CO_2e emissions in the building sector. Currently, information on emission intensity in the real estate sector is scarce, with only a few countries collecting and providing the required data. Our model can help reduce this gap and provide valuable insights into the carbon footprint of the real estate sector.

Keywords: *Deep Learning; Decarbonization; CO_2 Footprint, Image Classification.*

3.2 Introduction

The release of carbon dioxide (CO_2) from the building stock constitutes a significant portion of worldwide GHG emissions (UNEP, 2022). In the European Union, buildings are accountable for approximately 40% of final energy consumption (European Commission, 2019b). The reduction of greenhouse gas emissions in real estate holds significant potential for mitigating global warming and achieving the imperative objective of keeping the global temperature increase below 2 °C (IPCC, 2022). Notably, approximately 85% of the building stock in the European Union predates 2001 (European Commission, 2020), a period characterized by less strict energy performance requirements or more relaxed standards than those in place currently (European Commission, 2019a). Based on the report presented by the European Commission (European Commission, 2020), it becomes clear that the stipulated climate targets aiming at reducing energy consumption by a minimum of 60% shall remain unfulfilled despite the proposed weighted renovation quotas of 1% and 0.2% for deep renovations. There is an apparent discrepancy between actions required and actions taken. This discrepancy, also referred to as decarbonization gap, has increased in the EU in recent years with the only exception being 2021 due to the COVID-19 pandemic (BPIE, 2022).

While reducing the CO_2e emissions of the real estate sector is one of the pillars to limit global warming to below 2 °C (IPCC, 2022; European Commission, 2012), complete nation-wide emission data on the asset level is scarce. To tackle this problem, the EU introduced the EPBD recast in 2010, making it mandatory to provide potential buyers and tenants of a property with an EPC (European Parliament, 2010). After its introduction, EPCs turned into one of the most important sources of information on energy performance of buildings (Arcipowska et al., 2014). However, there is criticism on how EPCs are calculated (Coyne & Denny, 2021; Majcen et al., 2013; Pasichnyi et al., 2019) and also that EPCs are not publicly available in many countries (UNEP, 2021; Arcipowska et al., 2014). The European Union's response to the aforementioned critique is encapsulated within the revised EPBD. Nevertheless, the timely development of relevant databases, serving as a dependable foundation for forthcoming sustainable real estate policies or renovation strategies, remains an open question. Further creation of EPCs is, at the moment, only mandatory in the EU when selling an object or letting it out to a new tenant. Even with the creation of publicly available EPC databases, the energy performance of owner-occupied real estate and long-term rented objects will remain intransparent at the current state.

To reduce this information barrier, a new approach requiring little resource to estimate GHG emissions of buildings has to be developed. As the determination of energy performance is based on many visual features such as the roof, windows, building age, walls, building size, and materials used (BRE, 2022), a deep learning model could possibly provide an approximation of the energy performance and CO_2e emissions. In other fields, deep learning has proven to be a possibility to produce cost effective and efficient tools, using only little data as input (Jenkins & Burton, 2008; Liermann et al., 2019). Nonetheless, despite its already existing utilization within the real estate sector, there is still a considerable scope for continued exploration and investigation into the potential applications of this subject matter (Koch et al., 2019).

In this paper, we provide a to our knowledge novel approach to estimating CO_2e emissions of residential buildings using a pre-trained CNN model. We train our model using EPC data from the United Kingdom and corresponding Google Street View (GSV) images. The remainder of the paper is structured as follows: In Section 3.3, we provide an overview over relevant literature for our study. In Section 3.4, we describe the methodology used to built our model and the database; in Section 3.5, we describe the results of our model, and, in Section 3.6, discuss limitations and further research potential.

3.3 Literature Background and Context

With burgeoning computing capacity, Artificial Intelligence (AI) and especially Machine Learning (ML) systems, due to their potential to automate time consuming processes, have become increasingly popular for practical appliances such as computer vision, speech recognition, natural language processing, robot control, and others (Jordan & Mitchell, 2015). With rising interest, real estate is adapting the new methodologies in various ways.

Deep Learning, a subset of Machine Learning, particularly CNNs, has demonstrated significant utility in the realm of image analysis models and computer vision for real estate applications. Its profound efficacy lies in its ability to accurately forecast a diverse range of variables. Koch et al. (2019) provide a literature overview of the image analysis done in real estate research with CNNs. The authors classify the models into distinct categories, based on the input data employed, which comprise models predicated on aerial and satellite images, exterior front-view images, floor plans, and interior images. Due to the high accessibility of data, many papers utilize satellite images as their training and validation input. Earlier papers in this field on real estate related image analysis predominantly work on land-use segmentation or classification (Yang & Newsam, 2010; Fröhlich et al., 2013; Marmanis et al., 2015) and building footprint detection (Raikar & Hanji, 2016; Cohen et al., 2016; Ok, 2013). Accelerating progress in image classification through improvement of deep learning algorithms, especially the introduction of CNNs, and increasing strength of graphic processor units (Gu et al., 2018) has opened the possibility to take on more complex tasks and better results.

Jean et al. (2016), for example, predict the economic well-being of five different African countries using a CNN, achieving an accuracy of up to 75% explaining the variability of both asset wealth and consumption expenditure. The researchers construct their model by leveraging publicly accessible data, thereby endowing the model with the potential for scalability to different geographical regions. Through the utilization of satellite image data, the presented model allows for the extraction of valuable information, thus contributing to the amelioration of a data transparency gap. We share a similar motivation of extracting information from publicly available data with a CNN, to build a model that could reduce a data gap. Koch et al. (2021) present a CNN to predict where university graduates live in a city at a micro level, using different density classes. The presented model is able to derive the information using only satellite images after training. They find that their model has trouble differentiating neighbouring classes and thus allow a standard deviation of 1, to determine their overall model fit. The accuracy of the model increases by 37.8% when compared to true class prediciton accuracy.

Exterior front-view images provide an alternative basis for image analysis in real estate, presenting features not included in satellite images. With the increase in their availability through web based services such as GSV or Apple's Look Around, new areas of application have opened up (Koch et al., 2019). Paired with the increase in the efficiency of CNNs, several use-cases with high accuracy have been presented. Schmitz and Mayer (2016) use a CNN for the semantic interpretation of facade images. They achieve an accuracy of 85% differentiating facades, doors, windows and other building features. Furthermore, they conclude that even with relatively small training data sets CNNs can be trained sufficiently, if the available data is augmented and an already existing pre-trained model is fine-tuned. Obeso et al. (2017) train their model to distinguish street-view images of buildings in Mexico into four different architectural classes (pre-hispanic, colonial, modern and other). The model is not only capable of identifying buildings correctly, but also of deriving distinct building features from street-view images linked to a particular architectural style. They suggest that CNNs are able to identify architectural characteristics of a building and assign it to the fitting class with relatively high accuracy (88.01%). Further applications utilizing exterior street-view images to classify buildings include the prediction of building age (Zeppelzauer et al., 2018) and building condition (Koch et al., 2018). Considering that both the age and condition of a building contain valuable insights into the emissions of CO_2e from that building, it is reasonable to contemplate the possibility of developing a comprehensive model that can extract pertinent building characteristics for emission prediction through the analysis of street-view images using a CNN.

Despotovic et al. (2019) build on the previous models and simultaneously predict the construction age and the heating demand, assuming that both can be predicted solely based on visual exterior features of a building. The presented model is able to predict both variables with similar precision resulting in an accuracy of 62% differentiating between five demand classes, proving that exterior images provide enough information to approximately determine their energy demand.

Utilizing CNNs in real estate image classification tasks, has been successful in the past, with all the presented models achieving accuracies significantly above the random probability threshold. Even complex matters, such as architectural style or heating demand, usually determined by professionals, can be approximated using CNNs, as the model are able to extract enough of the required information from the input images. The outcome of this literature review indicates that image analysis utilizing CNNs holds promise for mitigating the information gap associated with the estimation of CO_2e emissions attributed to the building stock.

3.4 Methodology

Over the past few years, the use of CNNs as a visual-based algorithm has become widely accepted as the standard in the scientific community (Shi et al., 2018; Zhou et al., 2018; Wu et al., 2019). CNNs are multilayer neural networks, taking their name from the convolutional layers which are part of their architecture (Albawi et al., 2017). The models are able to learn and recognize visual patterns from the input database, making them very effective for image classification tasks (Sarker, 2021). Various models have emerged since the advent of the *AlexNet* model in 2012, originating from different annotated datasets, which vary in architecture, primarily in the number of layers in a network (depth), the number of kernels per layer (width), and the dataset used (Krizhevsky et al., 2017; Wu et al., 2019). Investigations have shown that the depth of the model is particularly important for the success of neural networks, especially looking at accuracy or identifying complex features. However, training networks with a substantial depth is challenging, requiring high computational power and offering limited help after a certain stage (Srivastava et al., 2015). In order to achieve optimization, a variety of architectures have been developed that utilize identity shortcut connections between layers. This allows for improved information flow between different parts of the network, resulting in more efficient and accurate computations. This has led to the general concept of Residual Neural Networks (short: "ResNets") (Targ et al., 2016; He et al., 2016), of which the approach implemented by He et al. (2016) constitutes the fundamental architecture of our model. In our baseline model, we employ a 34-layer *ResNet* model, which processes the image input from Google Street View.

The foundation for the training and validation data consists of two components, EPCs and GSV. While European energy certificates are subject to the EPBD, there are still differences between the individual member states (Semple & Jenkins, 2020). Thus, in order to obtain reliable information regarding the CO_2e emissions of buildings, an extensive and quality-controlled database is necessary. After conducting a thorough comparison of European energy performance certificate databases, it was determined that the Energy Performance of Buildings Register database of the Department for Levelling Up, Housing & Communities England & Wales is particularly suitable as a data source for the required emissions data. Particularly, UK EPCs incorporate three important factors: (1) the database is freely available and, therefore, particularly suitable for research purposes, (2) the CO_2e emission data does not require any conversion since the data is already displayed in the EPCs, and (3) EPCs are only issued by accredited assessors, which should ensure uniform quality. Additionally, there is already extensive literature on the used UK EPCs which can help to consider any potential information asymmetries and specific points in our research as well as identify weaknesses of the EPCs (e.g. as shown by Fuerst et al. (2015); Crawley et al. (2019); Taylor et al. (2019)).

Operational emissions (in CO_2e) of buildings are calculated based on two components. Firstly, the existing energy source, and secondly, the energy intensity of the building. The calculation is then derived by multiplying the conversion coefficient of specific energy sources with the energy intensity of the building. Consequently, in practice, it is possible for older, less modernized buildings, for example, to emit low CO_2e values due to their utilization of renewable energies, resulting in a low emission factor despite their high energy intensity. Thus, the trained model faces a high degree of complexity as the overall energy consumption of the house alone is insufficient to deduce the emissions. Instead, a detailed examination of externally visible characteristics, typical for each energy source, must take place. In this context, indicators such as existing chimneys or renewable energy systems can provide insight into the energy source. If none of these characteristics are externally visible, the model may estimate the energy source based on the observable condition or year of construction. The current Energy Efficiency Report (Office for National Statistics, 2022) suggests that the energy source correlates with the respective construction years, revealing trends specific to each period.

Consequently, a subsequent heating system modernization may go unidentified by the model, potentially leading to inaccuracies.

The calculation and procedures for issuing EPCs for residential buildings in the United Kingdom are generally based on the Standard Assessment Procedure (SAP), which can be applied in a reduced form as Reduced Data SAP (RdSAP) for older buildings constructed before 2008 (BRE, 2022). The latest version of the SAP, the SAP 10.2, has been introduced in England in June 2022 and in Wales in November 2022. The SAP calculates (only) the operational CO_2e emissions based on the emission factors specified in Table 12 of the SAP annex, which are given in emissions of kg CO_2e per Kilowatt-Hour (kWh). These emission factors are based on the energy source present in the building. In addition to the conventional Energy Efficiency Rating of A to G, this construction of the so-called Environmental Impact Rating represents a measure of notional CO_2e emissions. As is the case with the Energy Efficiency Rating: the higher the rating, the lower the expected emissions, with A equating to the lowest emissions and G to the highest emissions (BRE, 2022). For the presentation in our training and validation split, the relative indication of CO_2e emissions per m^2 per year is used.

In December 2022, the England EPC-database held a total of 24,015,615 records. However, considering the ongoing process of EPC updates and their limited validity period, a portion of the available data is outdated. Therefore, it is unsuitable for the purpose of our dataset. If for example an adress has two EPCs in the database, we eliminate the EPCs not depicted in the GSV image. This is accomplished by cross-referencing the date when the GSV image was captured with the date of issuance for the EPC. If the EPC is issued after the image was taken, we eliminate the data point. Further we filter addresses, which have no GSV image availability and are therefore unsuitable for our purpose. In order to maintain balance within our database, we restrict the number of data points for each class to match the lowest available data count among all classes. For a CNN model, a balanced and sizeable quantity of images in each class constitutes an indispensable prerequisite for achieving model precision (Luo et al., 2018). Predominantly, buildings of interest in the database are within the range of 0-140 CO_2e per m^2 per year. EPCs that exceed or fall below this range exist but are lacking in terms of the required frequency and timeliness. The classification of our CO_2e emissions categories is based on the established Environmental Impact Rating, which designates seven classes (BRE, 2022). As depicted in Table 3.1, each classification corresponds to a unique range of 20 kg CO_2e per m^2 per year.

After all expired, duplicate, and outdated EPCs, as well as EPCs for buildings outside the time periods available on Google Street View, were excluded from the database,

Table 3.1: CO_2e Emission Classifications

Classification	1	2	3	4	5	6	7
$CO_2e/m^2/year$	0-20	> 20-40	> 40-60	> 60-80	> 80-100	> 100-120	> 120-140

a random sample of 3,000 images was drawn for each classification. Especially the availability of high emission class data after the filtering process limits us to this number. This results in a dataset of 21,000 EPCs distributed throughout England. The descriptive statistics for the entire dataset are shown in Table 3.2.

Table 3.2: Descriptive Statistics of the EPC Sample

	$CO_2e/m^2/year$	Energy Efficiency Score	Total Floor Area in m^2
Mean	69.51	53.43	84.74
Std	38.43	21.39	46.38
Min	0.40	1.00	5.00
Q25%	36.00	39.00	59.00
Q50%	67.00	54.00	78.00
Q75%	103.00	70.00	100.00
Max	139.00	112.00	1,115.00

N = 21,000.

The descriptive statistics demonstrate a balanced sample across the classifications. The Energy Efficiency Score, which indicates the calculated energy efficiency class, differs from our classification, as it focuses on energy intensity rather than CO_2e emissions (BRE, 2022). The analysis of energy sources' descriptive statistics (refer to Table 3.5 in the Appendix) demonstrates a notable prevalence of gas utilization in low-emission categories, where more than 90% of cases are attributed to gas heating in the initial three classes. In the higher classes, this observation shifts towards electricity as the primary energy source. As previously explained, CO_2e consumption of buildings is fundamentally calculated based on the primary energy source and energy intensity. In this context, a positive correlation coefficient between energy intensity and emission values can be assumed. Therefore, for the validity of the sample, it appears crucial to have sufficient variability between energy sources and energy intensity. Regarding the seven CO_2e classifications, a differentiated picture emerges (see Figure 3.13 in the Appendix). For energy-efficient buildings, there appears to be a strong correlation between energy consumption and CO_2e emissions. However, this strong correlation gradually weakens, until in category 120-140 $CO_2e/m^2/p.a.$, there is only a correlation coefficient of approximately 0.3.

The Google Street View images were retrieved using the address information provided in the EPC database. The dataset comprises numerous images in which buildings are obscured by obstacles such as cars, vegetation, or other buildings, or are not visible at all. To train the model effectively and reduce the error rate, these images must be removed from the dataset. Examples of high and low information content images are provided in Figure 3.1. Images with high information content are characterized by their substantial explanatory value for the model, as they provide a comprehensive view of the building, thus serving as a solid foundation for feature extraction. Conversely, images where buildings are obscured by other objects or captured from unfavorable angles offer limited explanatory value for the model.

Figure 3.1: High (Top) and Low (Bottom) Information Content Images

Following the manual filtration of images with low information content, a total of 17,922 images remained, resulting in the removal of approximately 14.66% of all images from the dataset. Across all classifications, the dataset remains balanced, as illustrated in Figure 3.2. Achieving balance between classes is of paramount importance for machine learning models, as imbalanced training data has the potential to significantly impair the performance of convolutional neural networks (Johnson & Khoshgoftaar, 2019).

As previously mentioned, we use the pretrained *ResNet34* model developed by He et al. (2016) as our baseline model. *ResNet34* is an image classification model with 34 weighted layers. The utilized model undergoes pretraining on the ImageNet dataset, where it initially learns from a large dataset for a different task. The resulting pa-

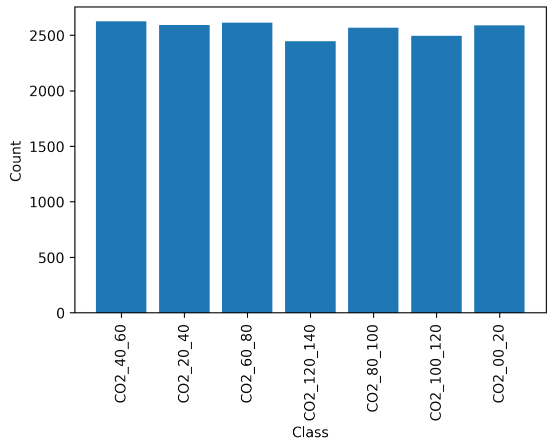


Figure 3.2: Count of Images in Each Classification

rameters from this pretraining are then fine-tuned and adjusted for our specific use case. Pretraining can significantly increase accuracy and is found to be particullary important for use cases with relatively small datasets (H. Li et al., 2019). To validate our results and identify the appropriate architecture and depth for our use case, we test the performance of several other models based on the *ResNet* architecture.

We have selected the following comparative models: a deeper and more complex *ResNet101* (He et al., 2016), a *NoisyStudent EfficientNet* of size B3, which is a less deep and complex CNN that achieves higher efficiency (Tan & Le, 2019), a *ResNext101 model* that incorporates a cardinality building block into the *ResNet* architecture, reducing the number of hyperparameters required (Xie, Girshick, Dollár, Tu, & He, 2017), and finally, a modified *SE ResNet101* that assigns weights to different channels based on their importance for prediction (Hu, Shen, Albanie, Sun, & Wu, 2019). The chosen models either exhibit greater complexity compared to our baseline model (*ResNet101, ResNext101, SE ResNet101*) or are less complex (*NoisyStudent EfficientNet*). We employ this comparison to assess the suitability of our architecture for our specific use case. Deeper models can potentially offer improved performance depending on the task, but also carry the risk of overfitting the data (Ying, 2019). For our training the images are input as 3-channel RGB with a (resized) size of 224 x 224 pixels. The 80%/20% training and validation split is performed using a deterministic learner and a stratified k-fold split to ensure comparability of the splits across all model comparisons. In our baseline model, we use simple image augmentation techniques such as random flipping, rotation, zoom warp, and lighting transform. Additionally, we normalize the images. Hyperparameter tuning of the model is performed using the Adam optimization algorithm (Kingma & Ba, 2017) which optimizes the parameters during training. One of the most important hyperparameters is the learning rate, which determines the step size at which the model adjusts its internal parameters during the training process and varies based on the specific use case (Smith, 2017). We control the learning rate using the Smith (2017) two-stage transfer learning protocol in which the first half of the epochs updates the head of the pre-trained model with a learning rate of *1e-02* and then fine-tunes the entire model with the knowledge gained in the second half. In stage 2, we gradually increase the learning rate, then gradually decrease it, starting from *1e-5* and ending at *1e-4*. This allows the neural network to learn more detailed information from new data and, thereby, improve performance (Ng et al., 2015). To evaluate the results, we employ the conventional accuracy equation shown in Equation 3.1, where the accuracy is defined as the ratio of the number of correctly classified buildings to the total number of buildings in the sample, multiplied by 100 and the F1-Score (Equation 3.2), the harmonic mean of precision and recall. The F1-Score is a popular metric when looking at models with unevenly distributed classification classes, as it accounts for the different class sizes. While our classification classes are evenly balanced we still provide F1-Scores for a better comparability with other models. Furthermore, Training and Validation Loss are compared between the models.

$$\text{Accuracy} = \frac{\text{Number of correctly classified buildings}}{\text{Total number of buildings in the sample}} \times 100 \tag{3.1}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.2}$$

Augmentation and Energy Source: A possible measure to tackle overfitting of CNN models is (further) augmenting the input data or increasing the input sample size (Perez & Wang, 2017). We randomly choose three of the following methods for the augmentation for every photo of the input sample: image rotation (45°, 90°, 135°), image transposition (left to right, top to bottom), image resizing (divide or multiply height or width by factor 2), and image cropping (reduce size by factor 2). With the overarching objective of enhancing the transparency of the market, we also opted to add additional input to our training data.

As the calculated CO_2e emissions are significantly influenced by the primary energy source used (additional elaboration pertaining to the CO_2e calculation is provided in

Section 3), we decided to incorporate an additional RGB-layer into the image matching process, allowing for the inclusion of both the CO_2e emissions and the primary energy source of the building. The necessary information regarding the primary energy source is readily obtainable from the corresponding EPCs. The same methodology utilized in the base model was applied to the new model, which includes heavy image augmentation and the incorporation of primary energy source data.

Grad-CAM Approach: In addition to the numerical metrics, we employ the Gradient-weighted Class Activation Mapping (Grad-CAM) method developed by Selvaraju et al. (2017) to achieve a visual validation of the model's performance. This technique employs a heatmap overlaid on the image to highlight regions of particular importance to the model's predictions. The Grad-CAM method leverages the model's gradients, which represent the direction for adjusting the CNN weights during iterative training to minimize error. By utilizing these gradients, relevant activation regions within the image are identified, providing insights into the weighting of activations corresponding to specific CO_2e classes. This approach significantly improves interpretability and instills confidence. However, it is important to note that the evaluation of the model still relies on established performance metrics such as accuracy, F1-Score, training, and validation loss, which serve as the primary indicators of its overall performance.

3.5 Results

In the following, we first present the results of our baseline model. We then compare the results of the baseline model with alternative CNN models. Finally, we evaluate a modified version of the base model.

Baseline Model: We find that our baseline model achieves an accuracy of 43.38%, already outperforming random probability of 14.29% (Figure 3.3). In Figure 3.4 it is apparent that around epoch 34 the models starts overfitting the training data, causing the validation loss to increase, while the training loss continues to decline.

After this epoch, the algorithm detects characteristics in the training set that are helpful in predicting the training data but, when used in the validation phase, are not improving accuracy due to the inability to generalize the characteristics to the overall population (Mutasa et al., 2020).

Model Comparison: To test the assumption that deeper models might overfit and to benchmark our baseline model, we utilize four additional state-of-the-art models. As

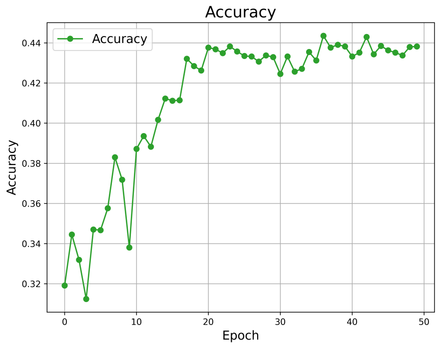


Figure 3.3: Baseline Accuracy

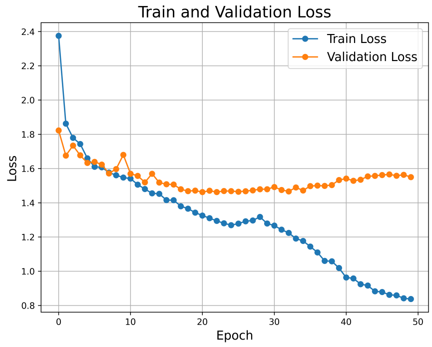


Figure 3.4: Baseline Train/Val. Loss

expected, the results suggest that the more complex models have a lower training loss whereas the validation loss increased, in some cases significantly. Our training set does not seem to be suited to train the deeper and more complex models encouraging overfitting (Table 3.3). While the computing time using the deeper models significantly increases, none of the tested models outperform our baseline model in terms of accuracy. We suspect that due to the relatively small sample size of the data set, the deeper models are not able to generalize the features learned in the training data (Zhang et al., 2021). The only model outperforming the baseline model in computing time is the less complex EfficientNet model. However, it is also achieving a significantly lower accuracy.

Table 3.3: Benchmarking the Baseline Model

Model	Train Loss	Validation Loss	Accuracy	Time
ResNet34 (baseline)	0.8379	1.5500	0.4382	00:44
ResNet101	0.7255	1.6568	0.4201	01:50
NoisyStudent (EfficientNet-B3)	0.7720	1.8317	0.3429	00:26
ResNeXt101	0.5172	1.8328	0.4329	05:06
SE ResNet101	0.1930	2.3859	0.4245	05:50

Augmentation and Energy Source: The results of the image augmentation indicate an increase in accuracy of approximately four percentage points, with the new model achieving an accuracy of approximately 47.59%. Although the new model exhibits superior performance in comparison to the baseline model, the modest increase in accuracy is somewhat unexpected. Upon scrutinizing the dataset, it was observed that a mere 33.4% of the buildings featured a primary energy source other than gas.

If we now assume that the energy source can also be partly determined by features of the house, such as appearance, roof, or chimneys it might become clear why the addition of the primary energy source only had little effect on the prediction results, while being a significant factor in influencing the CO_2e emission intensity. This fact contradicts our initial assumption that the model might only estimate energy intensity without drawing conclusions about the energy source or conversion coefficient. In this case, one would expect a higher increase in accuracy since, in practice, one of the two variables used to calculate CO_2e values is effectively included in the dataset, thereby greatly reducing the complexity of the model. It is important to note that the results do not provide a direct conclusion regarding whether the baseline version of the model derives the energy source directly (through building characteristics) or indirectly (via, for example, typical trends based on construction year). To test if the addition of the primary energy source altered the model comparison, we run a 50 layer (SE ResNet) to once again benchmark the results of the *ResNet34* model. Similar to the previous outcome, the *SE ResNet50* model is achieving a lower accuracy at 46.89%, requiring a computing time roughly six times higher than the *ResNet* model.

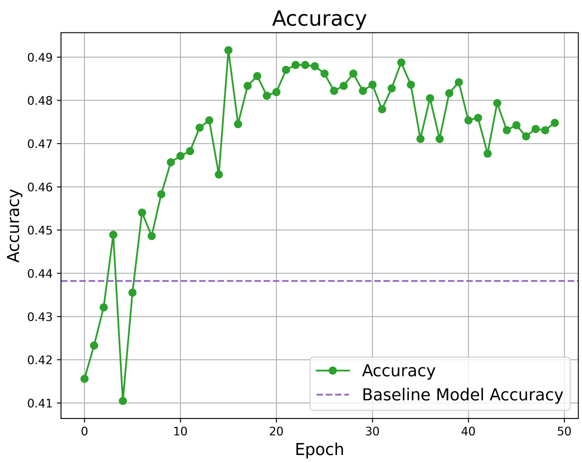


Figure 3.5: Accuracy of ResNet34 Model with Supplemented Primary Energy Sources

Grad-CAM Approach: Although it is apparent that our model significantly outperforms random probability, it remains unclear which specific visual features the model employs to approximate CO_2e emissions. To further validate our model, we use the Grad-CAM approach (Selvaraju et al., 2017). As shown in Figure 3.6, the model seems to identify the relevant buildings in every emission class. With the main focus on the property, important features in determining the CO_2e emission intensity such as the windows, the walls, and the roof among other components might be used for the estimation. Other parts of the image not related to the emission intensity like the sky, the street, or the neighboring buildings are identified as negligible by the model. Solely the front yard of some images seem to be helping the model with the prediction, which is not surprising as the look and size of the yards might help identify the building age of a property and, therefore, its energy efficiency (Aksoezen et al., 2015). In the overall context, however, the front yard does not appear to have any prevailing significance.

0-20 kg/m^2 CO_2e | 40-60 kg/m^2 CO_2e | 80-100 kg/m^2 CO_2e

Figure 3.6: Random Heat Maps Produced by Grad-CAM for Different Emission Classes

Accuracy: Given that our model is demonstrating superior performance compared to random probability and is visually validated through the Grad-CAM approach, it raises the question as to which factors are hindering its ability to accurately predict the emission class. Analyzing the confusion matrix in Figure 3.7 we conclude that the model is able to predict every class above random probability, with the lowest CO_2e emission class being predicted exceptionally well with a F1-Score of 80% (as seen in Table 3.4). With an accuracy rate of close to 50%, the model demonstrates reasonable precision across most of the other emission classes, 20-40 kg CO_2e/m^2 with 37% and 80-100 kg CO_2e/m^2 with 27% being the least accurate. Furthermore, the actual classes often seem to be confused with neighboring classes. The class with the lowest level of predictive accuracy, 80-100 CO_2e/m^2, is frequently misidentified with its two neighboring classes, namely 60-80 CO_2e/m^2 and 80-100 kg CO_2e/m^2.

A similar trend can be observed across all other emission classes, except for the lowest emission class, which was identified with high accuracy from the outset. The model appears to encounter challenges in correctly classifying images falling within the

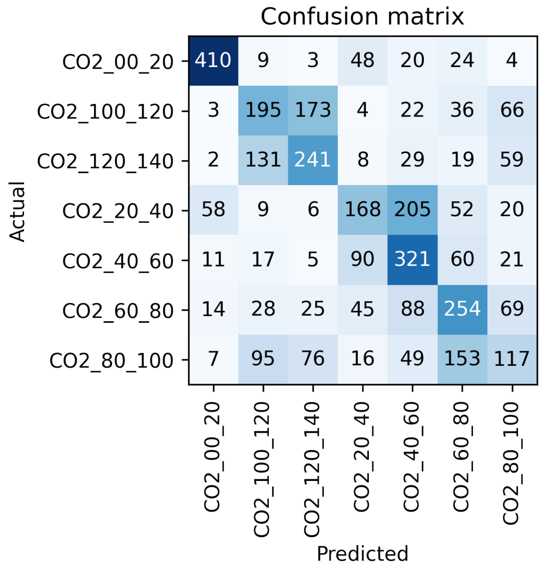


Figure 3.7: Confusion Matrix

mid-range of the emission spectrum into the appropriate category. Given the fixed boundaries of the predetermined emission classes, it is possible for a house with an emission of, for instance, 41 kg CO_2e / m^2, to be erroneously classified as belonging to the lower class, owing to the emission value being in close proximity to the upper boundary of the lower class. We test this assumption by plotting a histogram of the deviation of the model over all classes (Figure 3.8). The histogram reveals that the majority of incorrect predictions are off by a single emission class, with only a minimal number of predictions diverging by more than one class. When accounting for this standard deviation of one emission class, the model demonstrates an accuracy of approximately 75.34%. This noteworthy improvement in accuracy highlights the model's superior performance compared to random probability and underscores its exceptional ability to achieve accurate predictions when inaccuracies stemming from the fixed emission class boundaries are disregarded.

3.6 Discussion and Limitations

To the best of our knowledge, the discussed application of image classification regarding CO_2e emission in real estate represents a novel approach that yields promising results and may play a pivotal role in fostering global transparency in the ecology of real estate. While the simplicity and swiftness of the evaluated model are counterbalanced

Table 3.4: Classification Report

Class	Precision	Recall	F1-Score	Support
CO2_00_20	0.81	0.79	0.80	518
CO2_100_120	0.40	0.39	0.40	499
CO2_120_140	0.46	0.49	0.47	489
CO2_20_40	0.44	0.32	0.37	518
CO2_40_60	0.44	0.61	0.51	525
CO2_60_80	0.42	0.49	0.45	523
CO2_80_100	0.33	0.23	0.27	513
Accuracy			0.48	3585
Macro Average	0.47	0.47	0.47	3585
Weighted Average	0.47	0.48	0.47	3585

by certain limitations, the former predominantly prevails. Subsequently, we discuss the advantages and restrictions to our model, based on a practical application in a suburban street in Manchester, UK. Utilizing the EPC data of the houses and the corresponding Google Street View images, it becomes apparent that, as shown in Figure 3.9, there is significant heterogeneity among the emissions of individual residential buildings in the street.

If the primary objective is to achieve a maximum reduction in emissions, it appears that the buildings located in the middle and lower section of the street would benefit most from retrofits, as they are currently emitting high levels of CO_2e. Based on front view images, the presented CNN model could potentially identify and designate these areas in a city or on a street without the need for EPC data.

As shown in Figure 3.10, a large portion of the buildings are correctly classified by the model. In this example, the emission-intensive street section in the middle-lower area of the image is also correctly predicted. Although this example cannot be equated with an extensive test dataset, the results visualize the potential of the model. When sufficient computing power is available, the model can also be applied at the city or national level. At the same time, the weaknesses and limitations of the model are also apparent, requiring further research. Analyzing the buildings that were not correctly classified reveals the following implications:

The image quality can significantly affect the ability to identify the correct class. Upon examining the faulty classifications, images taken from unfavorable angles or obstructed by obstacles such as cars or greenery in the field of view appear to be a weakness (e.g.,

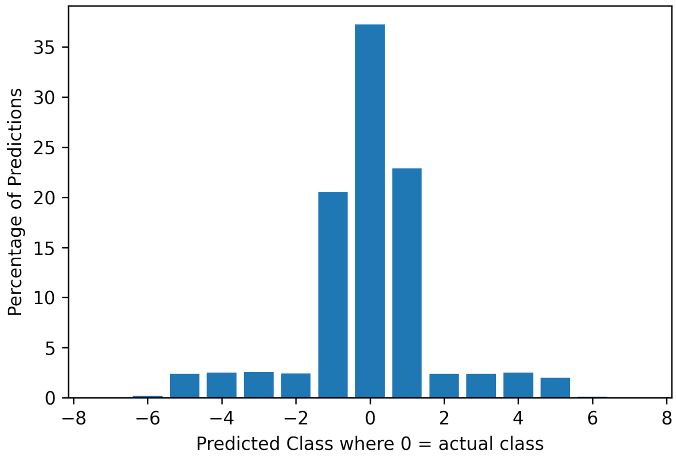


Figure 3.8: Deviation Histogram

as shown in the incorrectly classified object in Figure 3.11). This result is consistent with Koch et al. (2019), who are suggesting that pre-processing may be necessary to reduce noise for further use. Automated procedures (e.g., through prior object detection based on annotated images) for quality assurance of input data might thus further increase accuracy. Furthermore, certain aspects remain unclear, such as the extent to which the influence of land characteristics not directly tied to the main structure, such as the front yard, impacts the accuracy of the assessments. Further research, potentially in conjunction with object detection techniques, could offer solutions to this matter (Zhao et al., 2022).

Google Street View data is not always up-to-date, limiting the possibilities for obtaining information on the current building stock at the macro level or for specifically identifying regular changes in the building stock, such as renovation rates. In the example from Figure 3.10, the images were taken in August 2021 and not updated until the time of publication. The periods of capture might also differ even within one city. Therefore, a comparison of the periods between the training and validation datasets is essential. Figure 3.12 provides an example of an image of a plot of land that was taken in 2008 but is now developed and responsible for emissions as a building.

In addition to technical limitations due to data quality, the data foundation is subject to limitations arising from the interplay between the CO_2e values indicated in the energy certificate, the energy source, and the energy intensity. As previously explained, the

Green = Low CO_2e Emissions, Orange = Average CO_2e Emissions, Red = High CO_2e Emissions
Satellite Image acquired from Google Maps

Figure 3.9: Illustration of Sample EPCs in Manchester, UK

CO_2e emission values correlate strongly and positively with energy intensity, particularly in the more emission-efficient classes. However, this correlation weakens in the less energy-efficient classes, while the F1-Scores remain relatively constant. Despite the overall high accuracy, it remains uncertain to what extent the model may be biased, potentially generalizing the energy source based on external construction quality, rather than identifying it based on specific details such as the presence of a chimney. Further research could explore the combination of two dedicated models. One model could be trained to identify energy intensity—as demonstrated, for example, by Despotovic et al. (2019)—while another model could identify the energy source based on various external details, using techniques such as object detection, to provide additional insights.

One additional factor that potentially limits the data foundation of the model is the quality and meaningfulness of the obtained Energy Performance Certificates. In scientific discourse, the meaningfulness of EPCs has been subject to intense discussion. At its core, the debate centers around the difference between consumption and demand certificates (also referred to as the "Performance Gap") (among others, Cozza et al. (2020); Herrando et al. (2016); Coyne and Denny (2021)), the quality of data acquisition methods (among others, Pasichnyi et al. (2019); Y. Li et al. (2019); Hårsman et al. (2016); Hardy and Glew (2019)), the overall meaningfulness with regard to retrofitting (among others, Gouveia and Palma (2019); Christensen et al. (2014); Cozza et al. (2020)), and potential price premiums (among others, Olaussen et al. (2017); Fuerst and McAllister (2011)).

Satellite Image acquired from Google Maps

Figure 3.10: Predictions Based on the *ResNet34* model [$\sigma = 1$]

Figure 3.11: Wrong Angle Image

Figure 3.12: Empty Plot of Land

In the context of our research, the Performance Gap and potential quality constraints in the data acquisition methodology are particularly relevant. Concerning the Performance Gap, it can be observed that actual consumption values are not incorporated into the model, leading to a divergence between the estimated and actual CO_2e emissions. However, incorporating consumption values would not be advantageous for the model, as individual consumption information is not contained in the front-view images, in contrast to the building's physical properties. As Cozza et al. (2021) pointed out, in addition to imprecise input data and assumptions, individual occupant behavior and varying climate data are factors contributing to the Performance Gap. These individual characteristics of an inhabited property that are difficult to capture through a front-view image play a crucial role in the calculation of EPCs. However, they are not methodologically relevant to the CNN model. On the other hand, quality constraints in the input data and assumptions, similar to inaccurate or incorrectly oriented Google Street View images, can have a direct impact on the accuracy. In this context, quality assurance of the EPCs, as suggested by Hardy and Glew (2019) through machine learning, can further improve the quality of the model.

In conclusion, the input data, as previously elaborated, is one of the most important factors for a successful CNN model. Although EPCs are not perfect, they currently offer the only comprehensive means of analyzing the building's energy-related properties and thus serve as a good foundation for feature extraction in *ResNet* models.

3.7 Concluding Remarks

We presented a novel approach for estimating CO_2e emissions for residential buildings. In our baseline model we achieve an accuracy of 43.38%, using a *ResNet34* architecture. After including image augmentation and primary energy sources to the training data, the accuracy of the model increased to 47.59% (Figure 3.5. Disregarding the rigid borders of the emission classes and allowing a standard deviation of 1 for the prediction results, our accuracy significantly increased to 75.34%.

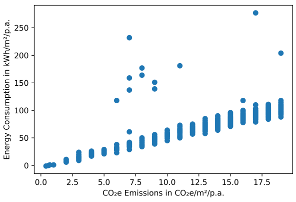
Evaluating the achieved results, the model can help to increase transparency regarding CO_2e emissions in the real estate market and, thereby, to contribute to the containment of global warming. Potential applications include the support of ESG due dilligences, policy making, and identifying need for retrofit. However, there remains a scope of potential for improvement in the model. Especially looking at the database, an increase in quality and sample size of the input data could be beneficial for predicition accuracy. Moreover, the model could be extended by adding additional input information not restricted to visuals, as already demonstrated with the addition of primary energy sources. With increasing progress of computing power, more sophisticated models and images collected by autonomous vehicles, even a model tracking daily progress could be conceivable (Gebru et al., 2017). Owing to the model's capability to enhance emission transparency, it could be correlated with other research domains in the real estate industry, such as price/rent prediction, which often overlooks emission data due to its unavailability. Overall, the presented model delivered promising results and might increase transparency of operational emissions in the residential housing market.

3.8 Appendix

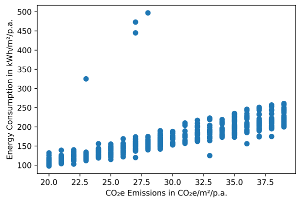
Table 3.5: Main Energy Sources per Classification

$CO_2e/m^2/year$	Electric	Gas	Other
0-20	0.02	0.97	0.02
>20-40	0.02	0.96	0.02
>40-60	0.02	0.97	0.01
>60-80	0.12	0.87	0.01
>80-100	0.42	0.55	0.03
>100-120	0.75	0.15	0.10
>120-140	0.85	0.08	0.07

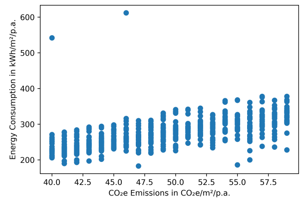
Note: The shares are shown aggregated at the level of the primary energy source. Electric includes e.g. heat pumps and storage heaters, gas includes e.g. gas-fired central heating or gas floor heating, Other includes e.g. coal and oil heating. N = 21,000.



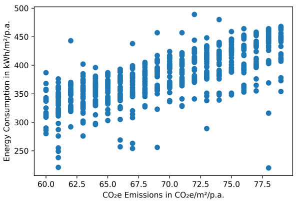
0-20 kg/m^2 CO_2e
$r = 0.89$



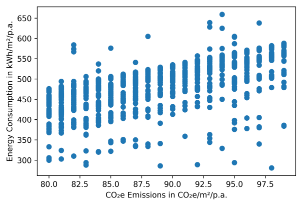
20-40 kg/m^2 CO_2e
$r = 0.91$



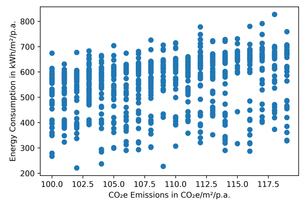
40-60 kg/m^2 CO_2e
$r = 0.86$



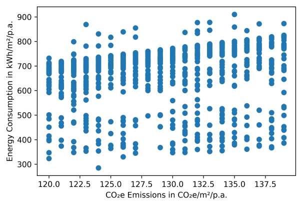
60-80 kg/m^2 CO_2e
$r = 0.85$



80-100 kg/m^2 CO_2e
$r = 0.72$



100-120 kg/m^2 CO_2e
$r = 0.43$



120-140 kg/m^2 CO_2e
$r = 0.30$

Figure 3.13: Scatter Plots for Energy Consumption and CO_2e Emissions

3.9 References

Aksoezen, M., Daniel, M., Hassler, U., & Kohler, N. (2015). Building age as an indicator for energy consumption. *Energy and Buildings*, *87*, 74–86.

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. *2017 International Conference on Engineering and Technology (ICET)*, 1–6.

Arcipowska, A., Anagnostopoulos, F., Mariottini, F., & Kunkel, S. (2014). Energy performance certificates across the EU. *Buildings Performance Institute Europe (BPIE)*.

BPIE. (2022). EU buildings climate tracker: Methodology and introduction of building decarbonisation indicators and their results. *Buildings Performance Institute Europe*. Retrieved from https://www.bpie.eu/publication/eu-buildings-tracker-methodology-and-results-for-building-decarbonisation-indicators/

BRE. (2022). The government's standard assessment procedure for energy rating of dwellings. *Version 10.2 (21-04-2022)*.

Christensen, T. H., Gram-Hanssen, K., de Best-Waldhober, M., & Adjei, A. (2014). Energy retrofits of Danish homes: is the energy performance certificate useful? *Building Research & Information*, *42*(4), 489–500.

Cohen, J. P., Ding, W., Kuhlman, C., Chen, A., & Di, L. (2016). Rapid building detection using machine learning. *Applied Intelligence*, *45*, 443–457.

Coyne, B., & Denny, E. (2021). Mind the energy performance gap: Testing the accuracy of building energy performance certificates in ireland. *Energy Efficiency*, *14*(6), 57.

Cozza, S., Chambers, J., Brambilla, A., & Patel, M. K. (2021). In search of optimal consumption: A review of causes and solutions to the energy performance gap in residential buildings. *Energy and Buildings*, *249*, Article 111253.

Cozza, S., Chambers, J., Deb, C., Scartezzini, J.-L., Schlüter, A., & Patel, M. K. (2020). Do energy performance certificates allow reliable predictions of actual energy consumption and savings? Learning from the Swiss national database. *Energy and Buildings*, *224*, Article 110235.

Crawley, J., Biddulph, P., Northrop, P. J., Wingfield, J., Oreszczyn, T., & Elwell, C. (2019). Quantifying the measurement error on England and Wales EPC ratings. *Energies*, *12*(18), Article 3523.

Despotovic, M., Koch, D., Leiber, S., Doeller, M., Sakeena, M., & Zeppelzauer, M. (2019). Prediction and analysis of heating energy demand for detached houses by computer vision. *Energy and Buildings*, *193*, 29–35.

European Commission. (2012). Energy: Roadmap 2050. *Publications Office*.

European Commission. (2019a). Achieving the cost-effective energy transformation of Europe's buildings energy renovations via combinations of insulation and heating & cooling technologies: Methods and data. *Publications Office*.

European Commission. (2019b). Commission recommendation (EU) 2019/786 of 8 may 2019 on building renovation. *Official Journal of the European Union*, *L 127*, 34-79.

European Commission. (2020). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – a renovation wave for Europe – greening our buildings, creating jobs, improving lives. *Official Publication*.

European Parliament. (2010). Directive 2010/31/EU of the European Parliament and of the Council of 19 may 2010 on the energy performance of buildings. *Official Journal of the European Union*.

Fröhlich, B., Bach, E., Walde, I., Hese, S., Schmullius, C., & Denzler, J. (2013). Land cover classification of satellite images using contextual information. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*, 1–6.

Fuerst, F., & McAllister, P. (2011). The impact of energy performance certificates on the rental and capital values of commercial property assets. *Energy Policy*, *39*(10), 6608–6614.

Fuerst, F., McAllister, P., Nanda, A., & Wyatt, P. (2015). Does energy efficiency matter to home-buyers? an investigation of epc ratings and transaction prices in england. *Energy Economics*, *48*, 145–156.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, *114*(50), 13108–13113.

Gouveia, J. P., & Palma, P. (2019). Harvesting big data from residential building energy performance certificates: Retrofitting and climate change mitigation insights at a regional scale. *Environmental Research Letters*, *14*(9), Article 095007.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., . . . Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354–377.

Hardy, A., & Glew, D. (2019). An analysis of errors in the energy performance certificate database. *Energy Policy*, *129*, 1168–1178.

Hårsman, B., Daghbashyan, Z., & Chaudhary, P. (2016). On the quality and impact of residential energy performance certificates. *Energy and Buildings*, *133*, 711–723.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

Herrando, M., Cambra, D., Navarro, M., de la Cruz, L., Millán, G., & Zabalza, I. (2016). Energy performance certification of faculty buildings in Spain: The gap between estimated and real energy consumption. *Energy Conversion and Management*, *125*, 141–153.

Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2019). Squeeze-and-excitation networks. *arXiv Pre-Print*.

IPCC. (2022). Climate change 2022: Mitigation of climate change. *Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Jenkins, R., & Burton, A. M. (2008). 100% accuracy in automatic face recognition. *Science*, *319*(5862), 435–435.

Johnson, J., & Khoshgoftaar, T. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, *6*(27).

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.

Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *arXiv Pre-Print*.

Koch, D., Despotovic, M., Leiber, S., Sakeena, M., Döller, M., & Zeppelzauer, M. (2019). Real estate image analysis: A literature review. *Journal of Real Estate Literature*, *27*(2), 269–300.

Koch, D., Despotovic, M., Sakeena, M., Döller, M., & Zeppelzauer, M. (2018). Visual estimation of building condition with patch-level convnets. In *Proceedings of the*

2018 ACM workshop on multimedia for real estate tech (pp. 12–17).

Koch, D., Despotovic, M., Thaler, S., & Zeppelzauer, M. (2021). Where do university graduates live? a computer vision approach using satellite images. *Applied Intelligence*, *51*(11), 8088–8105.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Li, H., Singh, B., Najibi, M., Wu, Z., & Davis, L. S. (2019). An analysis of pre-training on object detection. *arXiv Pre-Print*.

Li, Y., Kubicki, S., Guerriero, A., & Rezgui, Y. (2019). Review of building energy performance certification schemes towards future improvement. *Renewable and Sustainable Energy Reviews*, *113*, Article 109244.

Liermann, V., Li, S., & Schaudinnus, N. (2019). Deep learning: An introduction. *The Impact of Digital Transformation and Fintech on the Finance Professional*, 305–340.

Luo, C., Li, X., Wang, L., He, J., Li, D., & Zhou, J. (2018). How does the data set affect CNN-based image classification performance? *2018 5th International Conference on Systems and Informatics (ICSAI)*, 361–366.

Majcen, D., Itard, L., & Visscher, H. (2013). Theoretical vs. actual energy consumption of labelled dwellings in the netherlands: Discrepancies and policy implications. *Energy policy*, *54*, 125–136.

Marmanis, D., Datcu, M., Esch, T., & Stilla, U. (2015). Deep learning earth observation classification using imagenet pretrained networks. *IEEE Geoscience and Remote Sensing Letters*, *13*(1), 105–109.

Mutasa, S., Sun, S., & Ha, R. (2020). Understanding artificial intelligence based radiology studies: What is overfitting? *Clinical imaging*, *65*, 96–99.

Ng, H.-W., Nguyen, V. D., Vonikakis, V., & Winkler, S. (2015). Deep learning for emotion recognition on small datasets using transfer learning. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 443–449.

Obeso, A. M., Benois-Pineau, J., Acosta, A. Á. R., & Vázquez, M. S. G. (2017). Architectural style classification of mexican historical buildings using deep convolutional neural networks and sparse features. *Journal of Electronic Imaging*, *26*(1).

Office for National Statistics. (2022). Energy efficiency of housing in England and Wales: 2022. *Published Online*. Retrieved from https://www.ons.gov.uk/releases/energyefficiencyofhousinginenglandandwales2022

Ok, A. O. (2013). Automated extraction of buildings and roads in a graph partitioning framework. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *Volume II-3/W3*.

Olaussen, J. O., Oust, A., & Solstad, J. T. (2017). Energy performance certificates–informing the informed or the indifferent? *Energy Policy*, *111*, 246–254.

Pasichnyi, O., Wallin, J., Levihn, F., Shahrokni, H., & Kordas, O. (2019). Energy performance certificates—new opportunities for data-enabled urban energy policy instruments? *Energy Policy*, *127*, 486–499.

Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv Pre-Print*.

Raikar, A., & Hanji, G. (2016). Automatic building detection from satellite images using internal gray variance and digital surface model. *International Journal of Computer Applications*, *9*, 25–33.

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, *2*(6), 420.

Schmitz, M., & Mayer, H. (2016). A convolutional network for semantic facade segmentation and interpretation. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *41*, 709.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.

Semple, S., & Jenkins, D. (2020). Variation of energy performance certificate assessments in the European Union. *Energy Policy*, *137*, Article 111127.

Shi, X., Sapkota, M., Xing, F., Liu, F., Cui, L., & Yang, L. (2018). Pairwise based deep ranking hashing for histopathology image classification and retrieval. *Pattern Recognition*, *81*, 14–22.

Smith, L. N. (2017). Cyclical learning rates for training neural networks. *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 464–472.

Srivastava, R. K., Greff, K., & Schmidhuber, J. (2015). Training very deep networks. *Advances in Neural Information Processing Systems*, *28*.

Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*, 6105–6114.

Targ, S., Almeida, D., & Lyman, K. (2016). Resnet in resnet: Generalizing residual architectures. *arXiv Pre-Print*.

Taylor, J., Shrubsole, C., Symonds, P., Mackenzie, I., & Davies, M. (2019). Application of an indoor air pollution metamodel to a spatially-distributed housing stock. *Science of the Total Environment*, *667*, 390–399.

UNEP. (2021). 2021 Global Status Report for Buildings and Construction. Towards a zero-emissions, efficient and resilient buildings and construction sector. *United Nations Environment Programme*.

UNEP. (2022). 2022 Global Status Report for Buildings and Construction. Towards a zero-emissions, efficient and resilient buildings and construction sector. *United Nations Environment Programme*.

Wu, Z., Shen, C., & Van Den Hengel, A. (2019). Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern Recognition*, *90*, 119–133.

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1492–1500.

Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and spatial extensions for land-use classification. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 270–279.

Ying, X. (2019). An overview of overfitting and its solutions. In *Journal of Physics: Conference series* (Vol. 1168).

Zeppelzauer, M., Despotovic, M., Sakeena, M., Koch, D., & Döller, M. (2018). Automatic prediction of building age from photographs. *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, 126–134.

Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, *64*(3), 107–115.

Zhao, K., Liu, Y., Hao, S., Lu, S., Liu, H., & Zhou, L. (2022). Bounding boxes are all we need: Street view image classification via context encoding of detected buildings. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-17.

Zhou, Y., Hu, Q., & Wang, Y. (2018). Deep super-class learning for long-tail distributed image classification. *Pattern Recognition*, *80*, 118–128.

4 Real-GPT: Efficiently Tailoring LLMs for Informed Decision-Making in the Real Estate Industry

4.1 Abstract

Lately, LLMs such as ChatGPT and Llama have gained significant attention. These models demonstrate a remarkable capability in solving complex tasks, drawing knowledge primarily from a generalized database rather than niche subject areas. Consequently, there has been a growing demand for domain-specific LLMs tailored to social and natural sciences, such as BloombergGPT or BioGPT. In this study, we present a domain-specific LLM focused on real estate: "Real-GPT". The model is based on the parameter-efficient fine-tuning technique known as Quantized Low-Rank Adaptation (QLoRA) applied to Mistral 7B. To create a comprehensive fine-tuning dataset, we compiled a curated 21k self-instruction dataset sourced from 670 scientific papers, market research, scholarly articles, and real estate books. To assess the efficacy of Real-GPT, we devised a set of ca. 1,046 multiple-choice questions to gauge the real estate knowledge of the models. Furthermore, we tested Real-GPT in a real investment scenario. Despite its compact size, our model significantly outperforms GPT 3.5 and Mistral 7B. Hence, the model not only showcases superior performance but also illustrates its capacity to facilitate investment decisions or interpret current market data. This development showcases the potential of LLMs to revolutionize real estate analysis and decision-making.

Keywords: *Large Language Models; Real Estate, Fine-tuning.*

4.2 Introduction

Since the launch of ChatGPT in late 2022, LLMs have garnered significant global attention. Alongside ChatGPT, numerous models have emerged, such as Meta's Llama 3 and Google's Gemini. These models can perform a variety of Natural Language Processing (NLP) tasks and are capable of generating human-like texts or solving challenging problems swiftly. Their abilities primarily stem from intensive training on vast text corpora, based on the transformer architecture proposed by Vaswani et al. (2017). The datasets used, for example in the open access Llama model, largely consist of internet-crawled pages or programming databases (Touvron et al., 2023). This leads to a relative underrepresentation in the dataset of knowledge areas less discussed or available on the internet, consequently affecting the models' proficiency in domain-specific knowledge, such as real estate expertise.[1] Subsequently, several models have been developed or fine-tuned based on specific expertise. For instance, S. Wu et al. (2023) developed BloombergGPT, a 50-billion parameter LLM from scratch. Using 363 billion tokens of Bloomberg financial data, it significantly outperformes other LLMs in financial tasks. Similar approaches are seen in areas such as biomedicine (R. Luo et al., 2022), scientific papers (Taylor et al., 2022), and even home renovation (Wen et al., 2023), with a focus on fine-tuning existing open-source models. A model which focuses on real estate does not yet exist.

Fine-tuning LLMs, particularly through "Instruction Tuning", has gained significant prominence in practice and industry for adapting models to domain-specific NLP tasks. In addition to proving helpful in solving highly specialized tasks from various niche areas, fine-tuning is also used to enhance the user-friendliness of LLMs, for example, by adapting to chat functionalities such as ChatGPT (Dettmers et al., 2023; Wei, Bosma, et al., 2022; Ouyang et al., 2022). Data for such fine-tuning includes either human-written instruction data or synthetic datasets mainly generated by existing LLMs (Wang et al., 2023; Wei, Bosma, et al., 2022; Ouyang et al., 2022). The fine-tuning process has evolved considerably since 2022, with many methods focusing on resource-efficient yet effective and competitive techniques. In addition to acquiring specialized skills, it is conceivable that fine-tuned models may also exhibit advantages in fundamental abilities compared to their non-fine-tuned counterparts. This includes, for example, an improved ability

[1]The bulk of the training foundation for models such as Llama (Touvron et al., 2023) originates from the "CommonCrawl" dataset, which compiles text information from the publicly accessible internet without specific content filters. In addition to CommonCrawl, training also utilized domain-specific data from platforms such as "GitHub" or "StackExchange," which contain specialized information on data science and programming skills. Given our current understanding, there appears to be no LLM that incorporates specialized real estate datasets into its training data. Therefore, we conclude that real estate related content is relatively underrepresented in current LLMs.

to solve unseen tasks in a "zero-shot" scenario—that is, a scenario in which the large language model has not been provided with any additional or prior information about the task in the prompt (Wei, Bosma, et al., 2022; Z. Xu et al., 2023; Chung et al., 2024). In the context of real estate, this can be advantageous due to the high heterogeneity of properties, requiring a dynamic understanding of the situation and the resultant tasks. Despite these positive effects of fine-tuning, it requires the creation or use of an appropriate dataset, which can be resource- and time-intensive. Furthermore, fine-tuning with a dataset of low quality could result in diminished capabilities of the trained model (Y. Xu et al., 2023; Wei, Bosma, et al., 2022; Chung et al., 2024; Peng et al., 2023). Newer approaches, such as QLoRA (Dettmers et al., 2023), enable fine-tuning within 24 hours on a single consumer-grade GPU, making the process more affordable and accessible compared to the resource-intensive methods presented by Chung et al. (2024).

In addition to fine-tuning, Retrieval-Augmented Generation (RAG) has become an established method for extending the knowledge boundaries of LLMs. RAG facilitates the dynamic integration of information from external databases into pre-trained LLMs, enabling direct referencing to a text corpus when answering questions (Gao et al., 2024). T. Zhang et al. (2024) have extended the RAG approach by improving models' ability to distinguish between "oracle" documents, which contain the required answers, and "distractor" documents, which lack relevant information for the response. Current research further focuses on expanding the context windows of LLMs. Larger context windows allow for the direct consideration of lengthy conversations, documents, or databases within the initial prompt, thus providing the model with additional information alongside the use of fine-tuning or RAG methods (Chen et al., 2023; Ding et al., 2024; Y. Zhang et al., 2024). Due to the resource-intensive nature of expanding context windows, current scientific discussions are exploring alternative architectures alongside Generative Pre-trained Transformer (GPT) (Gu et al., 2022)[2]. Gu and Dao (2023) have introduced a new approach that uses structured state space models instead of the well-established GPT architecture. According to the authors, this alternative has achieved similar success in benchmarks, yet it is characterized by its ability to scale linearly. This feature could enable the effective processing of large datasets (Gu & Dao, 2023). Another strand of literature is currently exploring the integration of multiple, different models to either improve performance or scale more efficiently, known as the Mixture of Experts approach (e.g., Du et al. (2023) or Jiang et al. (2024)), or to enable multimodality (H. Liu et al., 2023).

[2]Transformer models typically exhibit quadratic scaling in relation to window length, meaning that expanding the context window also exponentially increases resource intensity.

However, these recent developments in LLMs have scarcely penetrated real estate research. In the broader NLP field, there has been significant research on the impact of language in news services or corporate reporting on financial performance. Sentiment analysis reveals a link between language choice and the prediction of REIT returns (Gene Birz & Tsang, 2021; Koelbl, 2020; Ruscheinsky et al., 2018; Paulus et al., 2022) and real estate market trends (Nagl, 2023; Ploessl et al., 2021; Zamani & Schwartz, 2017). Initial research on LLMs in real estate includes exploring ChatGPT's potential for property valuation in accordance with the Red Book (Cheung, 2023). Beyond this, the literature on the topic is limited. Investigations into LLMs have the potential to both broaden the existing research on NLP within the real estate sector and to open a multitude of new research opportunities in terms of potential efficiency gains in the industry. This expansion moves beyond the current reliance on word-matching or dictionary-based approaches for market sentiment analysis. Empirical studies can be conducted on the applications of LLMs to both direct and indirect real estate investments. However, a fundamental question arises regarding the ability of LLMs to produce fact-based responses to real estate economic issues. Pursuing our hypothesis that the training data encompass a relatively minor share of information pertinent to sophisticated real estate economic inquiries, it follows that in their current architecture, the available models are unlikely to exhibit proficiency in addressing real estate economic issues effectively. Given the critical role of the real estate sector within any national economy, it becomes imperative to bridge information asymmetries in order to provide stakeholders with well-grounded bases for decision-making.

Thus, our research aims to integrate SOTA data science methods with real estate economics, developing a specialized LLM model ("Real-GPT") to address the gaps in established LLMs regarding real estate knowledge. We assess whether the final model can enhance typical text-based tasks for scholars and practitioners in the real estate industry, and consider the potential benefits of fine-tuning LLMs for specific applications in this sector. Our approach aims to bridge the gap between academic research and practical application in the field, offering a tool both technically proficient and user-friendly for professionals in this sector. The research further focuses on developing a framework for incorporating additional domain-specific topics into an LLM, using existing and highly specialized literature and data. First, we describe the data sources and the cleansing process, followed by our methodology for processing the data and fine-tuning the open-source Mistral 7B model (Jiang et al., 2023). Finally, we summarize our findings, including evaluation results and comparisons with other established LLMs.

4.3 Data Source and Processing

4.3.1 Dataset Creation

To effectively fine-tune LLMs, it is important to utilize a structured dataset. Such a dataset typically comprises a multitude of natural language tasks, each consisting of input and output components. The model is then expected to predict the corresponding output based on the given input (Wang et al., 2023). Open-source datasets, characterized by their focus on specific tasks or their utility in enhancing understanding of complex topics and imparting specialized knowledge, can serve as an ideal data foundation. Notable examples include the "Alpaca"-dataset, predominantly generated by GPT-3.5, encompassing 52k tasks. This dataset has been shown to enhance the performance of Llama significantly in a diverse array of NLP tasks, including those previously unseen (Taori et al., 2023; Wang et al., 2023; Touvron et al., 2023). Similar methodologies have been employed using more advanced models such as GPT-4 (Peng et al., 2023), with considerably smaller model sizes (M. Wu et al., 2023), or even with multimodal datasets extending beyond textual tasks (Z. Xu et al., 2023). Additionally, there are numerous datasets focused on specialized knowledge areas, such as biomedical research questions (Jin et al., 2019), programming (Z. Luo et al., 2023), and chemical molecules (Nakata et al., 2020). Despite the abundance of freely available datasets from various fields and methods of data creation, certain niche areas remain underexplored. This is partly due to the complex and resource-intensive processes involved in dataset creation (S. Zhang et al., 2023). In the field of real estate, for instance, merely gathering sufficient textual sources poses a significant challenge. The available corpus in this domain largely comprises practical publications, internet sources, textbooks, and scientific papers, which need to be collated from various sources.

To ensure the highest data quality for our sample, we manually compiled the literature (following the approach by Taylor et al. (2022)), drawing on both university and online sources. The specific sub-topics covered comprise a range of subjects such as real estate management, cash flow modeling, valuation, portfolio management, corporate real estate management, finance, sustainability, and many more. An extensive crawl of real estate-related content from selected search engines revealed insufficient quality. Thus we only included manually selected sources. In total, 670 text files were collected of which approximately 251 were from practice-oriented sources (e.g., market reports on real estate markets, valuation guidelines, textbooks, etc.), while the remaining documents and papers originated from real estate journals. The real estate-specific literature freely available in the university library was fully incorporated into the text corpus. Online sources were included only if they met sufficient quality standards for citation, which required reputable authorship or organizational backing and availability in a citable format. Although manually compiling data ensures a certain level of quality, it also introduces the potential for selection bias in the dataset. In the real estate sector, a data foundation reflecting interest-driven opinions (e.g., market reports from brokers who are paid based on transactions or reports from lobby groups with specific agendas) may result in the fine-tuned LLM reflecting similar subjective opinions. Potential biases can also be detected in established LLMs such as ChatGPT (Wan et al., 2023; Huang et al., 2024). Therefore, depending on the dataset compiled, further testing may be necessary to ensure that no bias exists or is introduced, especially if the LLM is used for decision making. As Table 4.1 illustrates that, despite the smaller total number of files, practice-oriented sources contain significantly more words. Assuming an average reading speed of 238 words per minute (Brysbaert, 2019), it would take over 22 days to read the entire text corpus.

During the conversion of PDF files to text documents, it was evident that graphics, tables, and similar non-textual elements diminished the expressiveness of the text segments. Consequently, it was necessary to manually filter the data and remove these sections from the dataset. For better processing of the data, we divided the entirety of the files into individual sections of a maximum of 5,000 characters. Following an automated filtering process aimed at identifying evident text breaks or irrelevant text elements (such as excessively short lines, sequences of numbers, internet or email addresses, and similar components), and a subsequent manual refinement of the sections, approximately 69% of the total characters were retained (final curated data in Table 4.1).

Table 4.1: Dataset Composition

Dataset	Files	Lines	Words	Characters
Practical reports	251	490,351	4,327,210	32,308,295
Academic journals	419	362,581	3,328,727	25,714,534
Final curated data	7,858	508,500	5,840,517	40,231,005

4.3.2 Data Processing

In developing our fine-tuning data, we adopt a modified "Self-QA" approach as proposed by X. Zhang and Yang (2023). Our method (as seen in Figure 4.1) enables the generation of questions (instructions) and corresponding answers from unstructured data, such as the text corpus of our current dataset. Through the Self-QA approach,

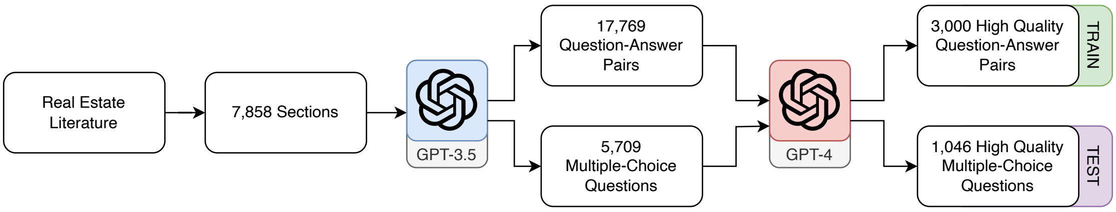


Figure 4.1: Workflow of Text Processing for Model Training and Testing

we can accurately and reliably represent property-specific knowledge within our fine-tuning dataset, reducing our reliance to solely on knowledge generated by LLMs, as seen in the Self-Instruct approach (Wang et al., 2023; S. Zhang et al., 2023). Furthermore, this method minimizes the need for manual audits, hence efficiently handling large volumes of data. The following example illustrates the complete data creation process based on a sample about valuation standards (RICS, 2022):

1. Conversion of the published PDF into individual text files, each containing 5,000 characters.

2. Generation of questions using GPT-3.5 Turbo:

    2.1. **Input prompt:**
    The background knowledge is:
    The cost approach indicates the value of an asset by the cost to create or replace the asset with another similar one, on the premise that a purchaser would not pay more for an asset than the cost to obtain one of equal usefulness. This method is frequently used in valuing investment companies or capital intensive firms. However, it would normally not be used except when the other two approaches have been considered but deemed not applicable, and the report would contain an explanation as to why this was so. [...]

    Please generate ten instruction questions as diverse as possible based on the content of the above article. These questions can be questions about facts or an understanding and evaluation of relevant content. Please assume that there is no corresponding article to refer to when asking questions, so do not use demonstrative pronouns such as "this" or "these" in the question.
    Please generate questions in the following format:

1. Question: ...
2. Question: ...

2.2. **Output:**

1. Question: What is the basic principle behind the cost approach in asset valuation?
2. Question: [...]

3. Answering each generated question:

3.1. **Input prompt:**
The background knowledge is:
The cost approach indicates [...]

Please answer the following question based on the content of the article above:
What is the basic principle behind the cost approach in asset valuation?

Please answer this question as thoroughly as possible, but do not change the key information in the original text, and do not include expressions such as "based on the above article" in the answer. Please generate the corresponding answer in the following format:

- Question: ...
- Answer: ...

3.2. **Output:**

- Question: What is the basic principle behind the cost approach in asset valuation?
- Answer: The cost approach in asset valuation is based on the principle that the value of an asset is determined by the cost to create or replace it with another similar one. [...]

4. Storing the respective questions and answers in a structured format.

As is evident from the example, the real estate-specific content is accurately reflected. From the data used in the curated dataset, a total of 17,769 domain-specific questions and corresponding answers were derived. The number of questions generated per API request did not always precisely reach the full amount, as sometimes the text lacked sufficient information. As previously mentioned, the newly generated dataset serves as the foundation for fine-tuning. However, to evaluate the performance, especially in comparison with other LLMs, a test dataset is also required. This dataset was created based on data which were not included in the fine-tuning dataset. Due to the lack of a standardized and accessible catalogue of questions and answers in the real estate sector, we created our own multiple-choice dataset by using approximately 30% of the text corpus and a slightly modified approach to the existing data creation process (see Appendix I for full description of prompts). In this dataset, out of four answer options, one is marked as correct, and the remaining three are distinct from the correct answer and purely fictional. Consequently, we generated a multiple-choice database comprising 5,709 questions. The utilization of the GPT-3.5 Turbo API resulted in total expenses of $121.35 for the entire dataset, including training and evaluation.

The generated questions and answers, while exclusively derived from manually selected literature, have not yet undergone any quality assurance checks. Quality is a critical criterion for the success of the fine-tuned model and should thus be included in the dataset evaluation (Y. Xu et al., 2023). Previous research has extensively explored how to increase the variability and quantity of instruction datasets (Taori et al., 2023; Köpf et al., 2023; Y. Xu et al., 2023). However, recent studies in this area have found that smaller sample sizes are sufficient, provided the dataset's quality is ensured. Zhou et al. (2023) demonstrated in their study that only 1,000 selected and quality-assured instructions could suffice to perform well in human tests compared to other models. Given this, the authors conclude that nearly all the necessary knowledge must already be present in the pre-training and that it is adequate to fine-tune with small data quantities. Du et al. (2023) developed a novel approach for selecting instruction data based on quality, coverage, and necessity, finding that 4,000 instructions could outperform the original dataset of 214,000 instructions. Y. Xu et al. (2023) also reported similar findings.

In the context of our real estate-specific dataset, to ensure the quality of both the training and test data, we subject each question to further review. We assess the quality of each question by using GPT-4, assigning a binary variable: 0 = not retained in the dataset, and 1 = retained (for complete prompt see Appendix II). Our focus is on verifying whether each question can be answered without additional context, as we have found that questions generated by the LLM often require further explanations, which only become apparent from the broader context. Of the original 17,769 questions in the fine-tuning dataset, only 3,000 quality-assured questions remain after review. The test dataset is reduced by 81.68%, from 5,709 questions to 1,046 multiple-choice questions.

4.4 Model Description and Fine-tuning Methodology

We chose Mistral 7B (Jiang et al., 2023) as our base model for training for two main reasons. First, the comparatively small size of 7B parameters allows for fine-tuning on a single consumer GPU. Second, at the time of writing, Mistral 7B is the most competent model with performance matching the best open 13B model Llama-2 across all evaluated benchmarks. It even surpasses the best released 34B model Llama-1 in reasoning, mathematics, and code generation. In order to improve training efficiency, we use QLoRA (Dettmers et al., 2023) instead of fine-tuning all model parameters. Low-Rank Adaptation (Hu et al., 2021) is a technique used to fine-tune large pre-trained models efficiently. The main idea behind LoRA is to introduce low-rank matrices which capture updates on the weights of a neural network during fine-tuning. Instead of updating all the parameters of the model, which can be computationally expensive for Large Language Models, LoRA focuses on updating these low-rank matrices. This approach reduces the number of parameters which need to be learned, leading to a more efficient fine-tuning process especially on less powerful hardware. Quantized Low-Rank Adaptation (Dettmers et al., 2023) is an extension of the LoRA technique, in which low-rank matrices are quantized. Quantization involves converting continuous or high-precision numerical values into a lower precision or discrete set. In the context of QLoRA, this means that the values in the low-rank matrices are represented with fewer bits. This approach can further reduce the computational cost and memory footprint of the model, making the fine-tuning process even more efficient and feasible.

4.4.1 Training

We format our data according to the template of Taori et al. (2023). A concrete example for a data point in our training dataset is shown below.

Training prompt:
Instruction
How do the inflows of capital in open real estate funds typically relate to the real estate market?
Response
The inflows of capital in open real estate funds typically have an anticyclical relationship with the real estate market. This means that, when the real estate market is performing well, the inflows of capital into these funds tend to decrease. Conversely, during times of economic downturn or financial crisis, the inflows of capital into open real estate funds tend to increase. This suggests that investors see these funds as a safe haven during turbulent times in the real estate market.

We train our model using the Mistral-7B-Instruct-v0.2 checkpoint provided via the Huggingface Model Hub. The training data comprises 80% of the dataset, while the remaining 20% is reserved for evaluation purposes to facilitate hyperparameter optimization. The parameters which yielded the best performance on the validation dataset were chosen for our final model, which is then evaluated on the test set. Our model undergoes fine-tuning for 5 epochs using the AdamW Optimizer (Loshchilov & Hutter, 2019) with a constant learning rate of 0.0002. A batch size of 6 and a gradient accumulation step size of 2 are utilized. Regarding QLoRA, we apply 4-bit quantization and set the matrix rank to eight.

4.4.2 Inference

Based on preceding tokens at each step, autoregressive language models create language by producing a conditional probability distribution over their entire vocabulary. To be precise, the language model outputs a numerical value for each token in the vocabulary. These produces are then fed into a softmax function which projects those values into probabilities between 0 and 1 that add up to one. Two key parameters, top_p and temperature, significantly influence the predictions of the next word. The temperature parameter is used to scale the raw outputs (logits) before feeding them into the softmax function. A higher temperature of l leads to a more diversified output, while a value of 0 makes the model quasi-deterministic. The top_p parameter decides the number of candidate tokens considered for the next token; for instance, a top_p of 0.1 means that only the most probable tokens whose probabilities add up to at least 10% are eligible for sampling for each step. We choose a value of 0.1 for both the top_p and temperature parameters, as this leads to more deterministic results, which is a desired property for question-answering tasks.

4.5 Evaluation

In our study, we assess our dataset by using a quantitative approach. Specifically, we first conduct a performance analysis of various LLMs in the context of real estate knowledge retrieval and then simulate an investment scenario under which the models can maximize their profits by investing in one of four properties.

4.5.1 Evaluation for Real Estate Knowledge

The first analysis is based on our test dataset, enabling us to benchmark the models against one another and quantify their actual capabilities in the realm of real estate.

The evaluation of NLP models using datasets has been a prevalent practice for many years, leading to a diverse range of datasets available for benchmarking. Guo et al. (2023) in their survey on LLM evaluation identified five key areas around which numerous benchmarks currently revolve: 1. knowledge and capability (e.g., assessing reasoning and question-answering capabilities), 2. alignment evaluation (e.g., examining bias and truthfulness), 3. safety (risk and robustness evaluation), 4. specialized LLMs (knowledge retrieval in specific research areas/domain-specific content), and 5. evaluation organization (applicability of benchmarks). In the context of our paper, the evaluation of domain-specific knowledge in real estate economics is paramount. Within the existing literature on evaluation datasets and methods, we found no specific references to this topic. Guo et al. (2023) identified benchmarks in areas primarily including biology, medicine, education, legislation, computer science, and finance. For evaluating (fine-tuned) LLMs in these areas, authors have used methods ranging from manually composed questions to automated procedures and existing examination questions (e.g., from the United States Medical Licensing Exam or the US legal Uniform Bar Examination). Our approach of automatically generating multiple-choice questions based on existing literature aligns more closely with the methodologies of Jin et al. (2019); J. Liu et al. (2023); Araci and Genc (2019).

To facilitate comparative analysis with other models, we also conducted evaluations by using our dataset on non-fine-tuned models. We used GPT-3.5 and GPT-4 as benchmarks, which consistently rank high in various benchmarks and are currently considered state-of-the-art models. Both GPT-3.5 and GPT-4 are closed-source models based on the pre-trained transformer architecture (Yenduri et al., 2024). However, due to their closed-source nature, the exact number of parameters in these models remains unclear. Brown et al. (2020) described GPT-3 as having 175 billion parameters, significantly more than our fine-tuned version. Official details for GPT-3.5 and GPT-4 are not available. Additionally, we evaluated the vanilla version of Mistral 7B, as described in the previous chapter. As we are interested in the language capabilities of the evaluated models, we used an identical prompt for all four models during evaluation (see Appendix III). We include a simple multiple-choice question and correct answer to homogenize the output format for each question. We refrain from advanced prompting techniques like chain-of-thought prompting on purpose to have a fair and thorough comparison. The results of these evaluations are presented in Table 4.2.

Table 4.2: Evaluation Results of Quantitative Approach

Model	Parameters	Fine-tuned on real estate knowledge	Average accuracy on evaluation dataset
Mistral 7B	7 billion	No	71.98%
Real-GPT (ours)	7 billion	Yes	83.46%
GPT-3.5	N / A	No	80.02%
GPT-4	N / A	No	84.51%

When comparing these, it becomes evident that Real-GPT significantly outperforms two of the three models examined. Specifically, compared to the identical but non-fine-tuned Mistral 7B, Real-GPT demonstrates an advantage of 11.48 percentage points, or 120 correctly answered questions. This comparison starkly illustrates the substantial impact of integrating domain-specific and pertinent content. The observed discrepancy aligns with findings from similar literature on other fine-tuned LLMs. In the context of our multiple-choice data collection, Real-GPT also exhibits superior performance compared to GPT-3.5, with a notable margin of 3.44 percentage points. A closer examination of the questions where our model performed accurately, while GPT-3.5 did not, reveals that Real-GPT particularly excels in addressing queries involving highly specialized content. In terms of the current state-of-the-art model from OpenAI, GPT-4.0, the performance of Real-GPT is comparable, albeit GPT-4.0 demonstrates a slight edge. A detailed analysis of the questions which were answered differently by the two models indicates that GPT-4.0 is more adept at handling questions to which the answer choices are closely related or require a more nuanced logical understanding. These observations are consistent with existing literature on GPT-4.0, which emphasizes the remarkable capabilities of this model, especially in comparison to its predecessor, GPT-3.5 (Jones & Bergen, 2024).

4.5.2 Evaluation for an Investment Case

In the practical application of LLMs, a variety of potential uses can be envisioned, and future research may provide insights into how effectively different models perform in specific real estate subfields. Beyond evaluating the capacity of these models for real estate-specific knowledge, we also aim to determine the extent to which fine-tuning existing models can add value in a scenario which closely mirrors practical applications. To this end, we conducted an additional evaluation based on a repeated-sales approach and an investment scenario. For this purpose, we utilized publicly available transaction data from the HM Land Registry (GOV.UK, 2024) and complemented it with detailed building information from the Energy Performance of Buildings Register (Department for Levelling Up, Housing & Communities, 2024). In total, we were able to identify 14,888 residential houses with full information and sold more than once. This dataset was then grouped into sets of four randomly assigned properties, each sold in the same years, resulting in a total of 3,702 groups (or 14.808 properties).

Given the extensive training data on real estate, Real-GPT model may have an informational advantage in investment scenarios over the untrained Mistral 7B model, potentially leading to more accurate investment assessments. Accordingly, we presented both models with the following prompt to evaluate their performance in this context:

Input prompt:
You are a sophisticated real estate investor looking to maximize your profits. Which of the following four UK residential properties would you acquire based on your market and real estate knowledge?

—

Property 1
Address: [. . .]
Acquisition Date: [. . .]
Acquisition Price: [. . .]
Energy Rating: [. . .]
Property Type: [. . .]
Built Form: [. . .]
Floor Area in m^2: [. . .]
Number of Rooms: [. . .]
Window Energy Efficiency: [. . .]
Walls Energy Efficiency: [. . .]
Roof Energy Efficiency: [. . .]
Main Fuel: [. . .]
Planned Selling: [. . .]

—

Property 2 [. . .]

Output:
Property 3

Despite the same data and inference parameters, both models delivered different results (as seen in Table 4.3). Real-GPT invests significantly more than the vanilla version of

Mistral 7B, yet it also has the potential to earn over 10 million pounds more under identical conditions. It is important to note that Real-GPT can only implicitly gain informational advantages through its training data and was not explicitly prepared for the investment scenario. Our findings align with initial research in the finance sector, which suggests that fine-tuned models may also be beneficial in the context of stock market investments (Araci & Genc, 2019). So far, the evaluation process has focused primarily on assessing the capabilities of Real-GPT in a quantitative context. Recognizing that real-world applications in the real estate industry also depend on qualitative characteristics that cannot be directly measured by numerical data, we included a textual comparison. In this comparison, both models were required to thoroughly justify their investment decisions and support those decisions with existing knowledge. This approach to generation quality is an essential component of the evaluation process in comparable literature (Gao et al., 2024). To establish a standardized comparison process, we had GPT-4 review the results—specifically, the textual reasoning of the investment decisions—and output the better answer. Of the 3,702 responses, the Real-GPT responses were preferred in almost 82% of cases. Thus, the qualitative assessment is consistent with the quantitative assessment, demonstrating that the superior performance in the quantitative assessment is reflected in a higher qualitative level of reasoning. Future research on this topic could further explore how models can be prepared more precisely to perform better in investment scenarios, for instance, through methods such as Chain-of-Thoughts, Few-Shot Prompting, or specifically designed datasets.

Table 4.3: Results of Repeated Sales Approach

	Real-GPT	Mistral 7B
Total Investment (in £)	1,057,065,929.00	1,020,196,382.00
Total Earnings (in £)	1,611,607,144.00	1,564,398,209.00
Total Profit (in £)	554,541,215.00	544,201,827.00
Average Investment (in £)	285,539.15	275,579.79
Average Floor Area (in m^2)	104.91	103.66
Average Room Numbers	5.18	5.17

4.6 Results, Outlook and Implications

Lately, there has been a rapid acceleration in advancements regarding NLP and, more broadly, in artificial intelligence technologies. The real estate sector is no exception, with a burgeoning number of PropTech companies harnessing these innovations. It is to be anticipated that AI's role will continue to expand in the future (Siniak et al., 2020). Our research demonstrates that, even in the face of limited textual data in real estate, fine-tuning existing LLMs presents a viable and notably resource-efficient method to enhance knowledge and application possibilities in this field substantially. The potential uses of real estate-specific LLMs are not confined to particular professions or disciplines but can be integrated into almost all segments of the value chain. Ma and Huang (2023) identify a range of applications for LLMs, including but not limited to property listing and search, customer service, marketing, legal support, home staging, home inspection, valuation, and property management. The Chui et al. (2023) forecasts that the adoption of generative AI in the real estate industry could generate up to $180 billion in additional value. This projection is particularly pronounced in the domains of marketing/sales and software engineering, where the greatest value addition is anticipated.[3] The utility of said LLMs in this sector is also recognized by early market players, such as JLL with its proprietary JLL-GPT (JLL, 2023). In addition to company-specific LLMs, there are opportunities to enhance existing closed-source models with additional capabilities and information. For example, OpenAI allows the creation of custom ChatGPTs, which can then be shared with others. These custom GPTs can be equipped with unique instructions, supplementary information, or additional capabilities (OpenAI, 2023). Currently, the OpenAI GPT Store offers a variety of models designed to address real estate-specific use cases.[4] Beyond enhancing domain-specific knowledge, custom-developed and hosted LLMs meet existing data protection standards and can be integrated into domain-specific applications precisely. Processes characterized by repetitive structures stand to gain profoundly from this integration. While the focus has been predominantly on information and text-based tasks, implementation in fields of financial mathematics, such as cash flow modeling or profitability calculations, is also feasible. Thus, fine-tuning LLMs or enhancing information bases represent just two of many areas in which future advancements can be expected.

From an interdisciplinary perspective, our approach has crafted a novel methodology for integrating unstructured niche knowledge into LLMs through fine-tuning methods. This entire process is illustrated in Figure 4.2. In the initial "Preprocessing" phase, we examined the data foundation, illustrating that a diversely constituted dataset, once suitably cleansed and processed, can be employed to create training and testing datasets. It is imperative to note that converted text files often contain erroneous

[3] This encompasses not only LLMs but the entire spectrum of generative AI.

[4] At this time there is no API available for the GPT Store. Therefore, we will not compare Real-GPT with these custom models.

information, such as improperly formatted images or tables, which necessitates manual or automated correction. Additionally, datasets that are created manually may be subject to selection bias. The subsequent “Processing” phase involves the generation of training and testing datasets. Methodologically, we adhere to the approach outlined by X. Zhang and Yang (2023), incorporating unstructured data as background information into an existing state-of-the-art LLM (in our case, GPT-3.5 Turbo) and then generating a set of potential questions from this information (in our instance, 10 questions from 5,000 characters each). A secondary query reintroduces the background information from which the LLM then individually answers each questions. For the test dataset, this query is subtly modified so that, in addition to the actual answer derived from the text, three additional fictitious responses are generated, resulting in the creation of multiple-choice questions. The final “Postprocessing” phase entails the actual fine-tuning—in our case, utilizing the QLoRA approach and the Mistral 7B model—and the evaluation of outcomes based on our test dataset. These results provide insights into the extent to which the methodology has enhanced domain-specific knowledge and other capabilities.

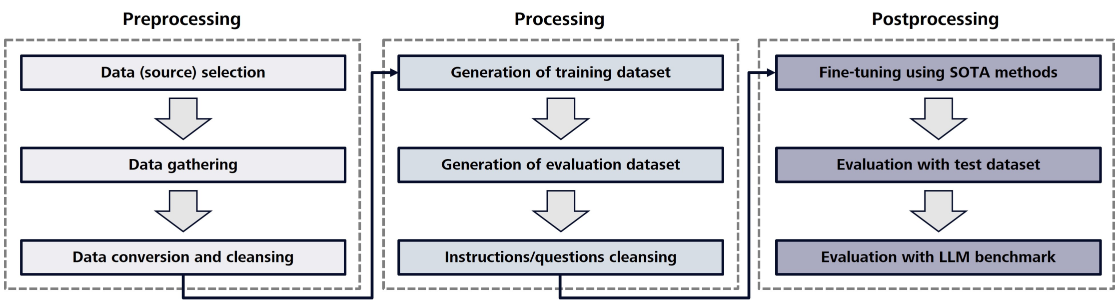


Figure 4.2: Summary of our Fine-tuning Approach Based on Unstructured Text Data

Real-GPT and our paper have demonstrated a clear demand for Large Language Models specifically tailored to the real estate sector. Our method of fine-tuning using unstructured text files emerges as a viable solution, offering benefits not only in real estate but also in other niche markets. The evaluations of our model and comparisons with standard benchmarks underscore the superiority of Real-GPT: it significantly outperforms vanilla versions of models such as Mistral 7B or GPT-3.5 Turbo. Despite our model’s likely smaller size and its operation on consumer-grade hardware, the results suggest that even the state-of-the-art model, GPT-4, does not markedly outperform it in our real estate-focused dataset. Furthermore, preliminary findings related to investment scenarios indicate that our model—likely due to its trained real estate knowledge—is more adept at making profit-maximizing investment decisions. Consequently, our research provides a foundational reference for both academics and practitioners aiming to integrate and leverage LLMs in the real estate domain. However, it is important to note that our approach is merely an initial step towards a comprehensive integration of LLMs in the value creation processes within the real estate industry. Processes in the real estate industry are highly diverse, and thus, it may be advantageous to tailor LLMs to specific sub-segments of the value chain by selectively choosing training data and further refining instructions in this context. It can be anticipated that Real-GPT, in its current form, may not perform equally well across all asset classes and markets.

Further steps in implementing and researching LLMs in the real estate industry must also address the extent to which ethical considerations might play a more significant role. The potential for value creation through the use of LLMs primarily arises from increased efficiency in work processes. Currently, this may mean that jobs are saved as tasks are automated. However, the findings from existing studies on this subject are not yet conclusive. Gmyrek et al. (2023) suggest that clerical workers are particularly at risk from the deployment of LLMs. Furthermore, Noy and Zhang (2023) demonstrate that participants worked more effectively with the aid of ChatGPT, which positively affected inequality among them. In addition to potential job displacement, the issue of privacy is omnipresent in the use of LLMs, starting with the question of what training data may be used. In the context of data relevant to real estate, it is often assumed that this involves highly sensitive information (e.g., lease agreements, tenant correspondence, financial accounting), which requires stringent data management. As previously mentioned, a privately hosted, secure LLM—such as Real-GPT—can help maintain data integrity during both inference and training phases. Market participants should be aware that integrating commercial SOTA models (e.g., through APIs, or employee usage) can result in sensitive data being shared with third parties (X. Wu et al., 2024), exposing an organization to legal risks. When interpreting results, it is important to note that algorithmic biases in LLMs cannot be ruled out. Stakeholders should be aware that results from LLMs may be biased, and if LLMs are to be significantly integrated, results should be partially objectified where necessary. Studies also show that LLMs can be classified across different political spectra (Rutinowski et al., 2024; Motoki et al., 2024) and may be subject to a wide range of biases (Wan et al., 2023; Huang et al., 2024). Due to the lack of literature and guidelines on the ethical use of LLMs in the real estate sector, market participants are currently challenged to create clear and transparent structures to avoid falling behind technologically while also not unduly exposing themselves to ethical and legal risks.

The rapidly evolving technology surrounding LLMs now transcends mere text processing. With the advent of multimodal LLMs such as GPT-4 (V)ision or LLaVA (H. Liu et al., 2023), as well as other approaches for handling domain-specific knowledge (e.g., advancements based on Retrieval Augmented Generation, see Lewis et al. (2020), and research into concepts such as Chain-of-Thought (Wei, Wang, et al., 2022; Cheung, 2023), Tree of Thoughts (Yao et al., 2023), or Communicative Agents (Li et al., 2023), a dynamic approach to integrating potential technological advancements seems beneficial. Moreover, typical professional roles in real estate involve more than just text processing; they also encompass dealing with complex financial mathematics and evaluating highly individualized inquiries at corporate, portfolio, or asset levels. This necessitates a further customization of the models.[5] Hence, Real-GPT should be seen more as a framework for a generalist model which requires specific adjustments to meet the unique demands of the real estate industry.

[5]In addition to customization and expansion, connecting LLMs to existing econometric or machine learning methods can significantly broaden their applications. It is conceivable that necessary calculations or estimations continue to be managed through established methods that are well-researched (e.g., machine learning techniques for determining property values). In a second step, these results can then be further interpreted and processed by the LLM.

4.7 Appendix

Appendix I: Prompting Scheme to Generate Multiple-choice Questions
Based on content from RICS (2022).

1. Generation of questions using GPT-3.5 Turbo:
   1.1. **Input prompt:**
   The background knowledge is:
   An intangible asset is defined as a non-monetary asset that manifests itself by its economic properties. It does not have physical substance but grants rights and/or economic benefits to its owner. It is therefore an asset that is capable of being separated or divided from a business entity and sold, transferred, licensed, rented or exchanged individually or with a related asset, liability or contract. Non-identifiable intangible assets arising from contractual or legal rights that may or may not be separable from the entity, or other rights and obligations, are generally termed 'goodwill'. [...]

   Please generate ten instruction questions as diverse as possible based on the content of the above article. These questions can be questions about facts or an understanding and evaluation of relevant content. The generated questions should be suitable for multiple-choice answers with four options. Do not create the answers. Please assume that there is no corresponding article to refer to when asking questions, so avoiding demonstrative pronouns such as "this" or "these" in the question.
   Please generate questions in the following format and do not include answers:
      1. Question: ...
      2. Question: ...

   1.2. **Output:**
      1. Question: What is the key characteristic which differentiates "goodwill" from other intangible assets?
      2. Question: [...]
2. Creating one true and three false answers for the multiple-choice dataset:
   2.1. **Input prompt:**
   The background knowledge is:

An intangible asset is defined [...]

Please generate four multiple-choice answers on the content of the article above:
What is the key characteristic which differentiates "goodwill" from other intangible assets?

The first answer is the correct answer based on the text. The next three answers are fictional answers that would make sense for someone not knowing the background knowledge. Do not include expressions such as "based on the above article" in the answers. Please generate the corresponding answers in the following format:

- Question: ...
- Correct answer: ...
- Fictional answer 1: ...
- Fictional answer 2: ...
- Fictional answer 3: ...

2.2. **Output:**

- Question: What is the key characteristic that differentiates "goodwill" from other intangible assets?
- Correct answer: Goodwill is a non-identifiable intangible asset arising from contractual or legal rights which may or may not be separable from the entity.
- Fictional answer 1: Goodwill is a tangible asset with physical substance and economic properties.
- Fictional answer 2: Goodwill represents the monetary value of a business entity and can be easily quantified and measured.
- Fictional answer 3: Goodwill is an intangible asset which is always separable from a business entity and sold or exchanged individually.

3. Storing the respective questions and answers in a structured format. Before evaluation, the answers are shuffled.

Appendix II: Quality Assurance Few-shot Prompt for Both Training and Test Dataset
Based on content from RICS (2022).

Input prompt:
You will receive a question generated from a research papers with the help of an LLM. Your task is to decide if the question can be answered without additional context. The additional context could be anything from the article (figures, tables, paragraphs, etc.).

—

Examples:

- Q: What is the purpose of the revaluation without re-inspection mentioned in the article?
- A: 0
- Q: What should a valuer do when information provided by parties other than themselves is involved in a valuation assignment?
- A: 1
- Q: What are the terms of engagement mentioned in PS 2 section 7 about?
- A: 0

—

You should only provide 1 or 0 as an answer, in which 1 stands for "can be answered without additional context" and 0 stands for "cannot be answered without additional context".

Your question is: "How does a detailed description of valuation methods help ensure the proper reporting of assets or liabilities?"

Output:
1

Appendix III: Inference Few-shot Prompt for all Evaluated Models

Input prompt:
Please answer the following multiple-choice question. Only answer with the letter of the correct answer.

—

Example Question:
What is the Capital of Germany?

- A: Hamburg
- B: Cologne
- C: Berlin
- D: Munich

Example Output:
C

—

Question:

What is the key characteristic that differentiates "goodwill" from other intangible assets?

- A: Goodwill represents the monetary value of a business entity and can be easily quantified and measured.
- B: Goodwill is a non-identifiable intangible asset arising from contractual or legal rights which may or may not be separable from the entity.
- C: Goodwill is an intangible asset which is always separable from a business entity and sold or exchanged individually.
- D: Goodwill is a tangible asset with physical substance and economic properties.

Output:
B

4.8 References

Araci, D., & Genc, Z. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv Pre-Print*.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, *109*, Article 104047.

Chen, S., Wong, S., Chen, L., & Tian, Y. (2023). Extending context window of large language models via positional interpolation. *arXiv Pre-Print*.

Cheung, K. S. (2023). Real estate insights unleashing the potential of ChatGPT in property valuation reports: the "red book" compliance chain-of-thought (cot) prompt engineering. *Journal of Property Investment & Finance*, *42*(2), 200–206.

Chui, M., Roberts, R., Yee, L., Hazan, E., Singla, A., Smaje, K., ... Zemmel, R. (2023). The economic potential of generative AI: The next productivity frontier. *McKinsey Global Institute*. Retrieved from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., ... Wei, J. (2024). Scaling instruction-finetuned language models. *Journal of Machine Learning Research*, *25*(70), 1–53.

Department for Levelling Up, Housing & Communities. (2024). Energy performance of buildings data: England and Wales. *Energy Performance Certificates dataset*. Retrieved from https://epc.opendatacommunities.org/

Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). Qlora: Efficient finetuning of quantized llms. *Advances in Neural Information Processing Systems*, *36*, 10088–10115.

Ding, Y., Zhang, L. L., Zhang, C., Xu, Y., Shang, N., Xu, J., ... Yang, M. (2024). Longrope: Extending LLM context window beyond 2 million tokens. *arXiv Pre-Print*.

Du, Q., Zong, C., & Zhang, J. (2023). Mods: Model-oriented data selection for instruction tuning. *arXiv Pre-Print*.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., ... Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. *arXiv Pre-Print*.

Gene Birz, S. D., Erik Devos, & Tsang, D. (2021). Is good news good and bad news bad in the REIT market? *Journal of Real Estate Portfolio Management*, *27*(1), 43-62.

Gmyrek, P., Berg, J., & Bescond, D. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. *ILO Working Paper 96*.

GOV.UK. (2024). HM land registry: Transaction data. *Transaction Dataset*. https://www.gov.uk/guidance/hm-land-registry-transaction-data.

Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. *arXiv Pre-Print*.

Gu, A., Goel, K., & Ré, C. (2022). Efficiently modeling long sequences with structured state spaces. *arXiv Pre-Print*.

Guo, Z., Jin, R., Liu, C., Huang, Y., Shi, D., Supryadi, … Xiong, D. (2023). Evaluating large language models: A comprehensive survey. *arXiv Pre-Print*.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., … Chen, W. (2021). Lora: Low-rank adaptation of large language models. *arXiv Pre-Print*.

Huang, D., Bu, Q., Zhang, J., Xie, X., Chen, J., & Cui, H. (2024). Bias testing and mitigation in LLM-based code generation. *arXiv Pre-Print*.

Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., de las Casas, D., … Sayed, W. E. (2023). Mistral 7b. *arXiv Pre-Print*.

Jiang, A. Q., Sablayrolles, A., Roux, A., Mensch, A., Savary, B., Bamford, C., … Sayed, W. E. (2024). Mixtral of experts. *arXiv Pre-Print*.

Jin, Q., Dhingra, B., Liu, Z., Cohen, W., & Lu, X. (2019). Pubmedqa: A dataset for biomedical research question answering. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 2567–2577.

JLL. (2023). *Jll unveils first GPT model for commercial real estate.* https://www.us.jll.com/en/newsroom/jll-unveils-first-gpt-model-for-commercial-real-estate. Chicago, IL, USA. (Accessed: 2024-01-24)

Jones, C., & Bergen, B. (2024). Does GPT-4 pass the turing test? *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 5183–5210.

Koelbl, M. (2020). Is the MD&A of US REITs informative? A textual sentiment study.

Journal of Property Investment & Finance, *38*(3), 181–201.

Köpf, A., Kilcher, Y., von Rütte, D., Anagnostidis, S., Tam, Z. R., Stevens, K., ... Mattick, A. (2023). Openassistant conversations - democratizing large language model alignment. *Advances in Neural Information Processing Systems*, *36*, 47669–47681.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, *33*, 9459–9474.

Li, G., Hammoud, H., Itani, H., Khizbullin, D., & Ghanem, B. (2023). Camel: Communicative agents for "mind" exploration of large language model society. *Advances in Neural Information Processing Systems*, *36*, 51991–52008.

Liu, H., Li, C., Wu, Q., & Lee, Y. J. (2023). Visual instruction tuning. *Advances in Neural Information Processing Systems*, *36*, 34892–34916.

Liu, J., Xia, C. S., Wang, Y., & Zhang, L. (2023). Is your code generated by ChatGPT really correct? Rigorous evaluation of large language models for code generation. *Advances in Neural Information Processing Systems*, *36*, 21558–21572.

Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. *arXiv Pre-Print*.

Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H., & Liu, T.-Y. (2022). Biogpt: generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, *23*(6).

Luo, Z., Xu, C., Zhao, P., Sun, Q., Geng, X., Hu, W., ... Jiang, D. (2023). Wizardcoder: Empowering code large language models with evol-instruct. *arXiv Pre-Print*.

Ma, J., & Huang, K. (2023). ChatGPT in real estate. In K. Huang, Y. Wang, F. Zhu, X. Chen, & C. Xing (Eds.), *Beyond AI: ChatGPT, web3, and the business landscape of tomorrow* (pp. 219–242). Cham: Springer Nature Switzerland.

Motoki, F., Pinho Neto, V., & Rodrigues, V. (2024). More human than human: Measuring chatgpt political bias. *Public Choice*, *198*(1), 3-23.

Nagl, C. (2023). Sentiment analysis within a deep learning probabilistic framework– new evidence from residential real estate in the United States. *Journal of Housing Research*, 1–25.

Nakata, M., Shimazaki, T., Hashimoto, M., & Maeda, T. (2020). Pubchemqc pm6: Data sets of 221 million molecules with optimized molecular geometries and electronic properties. *Journal of Chemical Information and Modeling*, *60*(12), 5891–5899.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187-192.

OpenAI. (2023). Introducing gpts. *OpenAI Announcements*. Retrieved from https://openai.com/index/introducing-gpts/ (Accessed: 2024-05-21)

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*, 27730–27744.

Paulus, N. M., Koelbl, M., & Schaefers, W. (2022). Can textual analysis solve the underpricing puzzle? A US REIT study. *Journal of Property Investment & Finance*, *40*(6), 548–570.

Peng, B., Li, C., He, P., Galley, M., & Gao, J. (2023). Instruction tuning with GPT-4. *arXiv Pre-Print*.

Ploessl, F., Just, T., & Wehrheim, L. (2021). Cyclicity of real estate-related trends: Topic modelling and sentiment analysis on german real estate news. *Journal of European Real Estate Research*, *14*(3), 381–400.

RICS. (2022). RICS Valuation – Global Standards. *Red Book Online*. Retrieved from https://www.rics.org/profession-standards/rics-standards-and-guidance/sector-standards/valuation-standards/red-book/red-book-global

Ruscheinsky, J. R., Lang, M., & Schäfers, W. (2018). Real estate media sentiment through textual analysis. *Journal of Property Investment & Finance*, *36*(5), 410–428.

Rutinowski, J., Franke, S., Endendyk, J., Dormuth, I., Roidl, M., & Pauly, M. (2024). The self-perception and political biases of ChatGPT. *Human Behavior and Emerging Technologies*, *2024*, 9.

Siniak, N., Kauko, T., Shavrov, S., & Marina, N. (2020). The impact of proptech on real estate industry growth. *IOP Conference Series: Materials Science and Engineering*, *869*(6), Article 062041.

Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., ... Hashimoto, T. B. (2023). Stanford alpaca: An instruction-following llama model. *GitHub repository*. Retrieved from https://github.com/tatsu-lab/stanford_alpaca

Taylor, R., Kardas, M., Cucurull, G., Scialom, T., Hartshorn, A., Saravia, E., ... Stojnic, R. (2022). Galactica: A large language model for science. *arXiv Pre-Print*.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... Lample, G. (2023). Llama: Open and efficient foundation language models. *arXiv Pre-Print*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.

Wan, Y., Pu, G., Sun, J., Garimella, A., Chang, K.-W., & Peng, N. (2023). "kelly is a warm person, joseph is a role model": Gender biases in LLM-generated reference letters. *arXiv Pre-Print*.

Wang, Y., Kordi, Y., Mishra, S., Liu, A., Smith, N. A., Khashabi, D., & Hajishirzi, H. (2023). Self-instruct: Aligning language models with self-generated instructions. *arXiv Pre-Print*.

Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., … Le, Q. V. (2022). Finetuned language models are zero-shot learners. *arXiv Pre-Print*.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., ichter, b., Xia, F., … Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824–24837.

Wen, C., Sun, X., Zhao, S., Fang, X., Chen, L., & Zou, W. (2023). Chathome: Development and evaluation of a domain-specific language model for home renovation. *arXiv Pre-Print*.

Wu, M., Waheed, A., Zhang, C., Abdul-Mageed, M., & Aji, A. F. (2023). Lamini-lm: A diverse herd of distilled models from large-scale instructions. *arXiv Pre-Print*.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., … Mann, G. (2023). Bloomberggpt: A large language model for finance. *arXiv Pre-Print*.

Wu, X., Duan, R., & Ni, J. (2024). Unveiling security, privacy, and ethical concerns of ChatGPT. *Journal of Information and Intelligence*, *2*(2), 102-115.

Xu, Y., Yao, Y., Huang, Y., Qi, M., Wang, M., Gu, B., & Sundaresan, N. (2023). Rethinking the instruction quality: Lift is what you need. *arXiv Pre-Print*.

Xu, Z., Shen, Y., & Huang, L. (2023). MultiInstruct: Improving multi-modal zero-shot learning via instruction tuning. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 11445–11465.

Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T., Cao, Y., & Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, *36*, 11809–11822.

Yenduri, G., Ramalingam, M., Selvi, G. C., Supriya, Y., Srivastava, G., Maddikunta, P. K. R., … Gadekallu, T. R. (2024). Gpt (generative pre-trained transformer) - a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. *IEEE Access*, *12*, 54608-54649.

Zamani, M., & Schwartz, H. A. (2017). Using Twitter language to predict the real estate market. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 28–33.

Zhang, S., Dong, L., Li, X., Zhang, S., Sun, X., Wang, S., ... Wang, G. (2023). Instruction tuning for large language models: A survey. *arXiv Pre-Print*.

Zhang, T., Patil, S. G., Jain, N., Shen, S., Zaharia, M., Stoica, I., & Gonzalez, J. E. (2024). Raft: Adapting language model to domain specific rag. *arXiv Pre-Print*.

Zhang, X., & Yang, Q. (2023). Self-qa: Unsupervised knowledge guided language model alignment. *arXiv Pre-Print*.

Zhang, Y., Li, J., & Liu, P. (2024). Extending LLMs' context window with 100 samples. *arXiv Pre-Print*.

Zhou, C., Liu, P., Xu, P., Iyer, S., Sun, J., Mao, Y., ... Levy, O. (2023). Lima: Less is more for alignment. *Advances in Neural Information Processing Systems*, *36*, 55006–55021.

5 Unveiling the Impact of SFDR on Unlisted Real Estate Funds: A J-Curve and Panel Regression Analysis

5.1 Abstract

Purpose: This paper investigates the relationship between the Sustainable Finance Disclosure Regulation (SFDR) and the performance of unlisted real estate funds.
Design/methodology/approach: While existing literature has primarily focused on the impact of voluntary sustainability disclosure, such as certifications or reporting standards, this study addresses a significant research gap by constructing and analyzing the financial J-Curve of 40 funds under the SFDR. The authors employ a panel regression analysis to examine the effects of different SFDR categories on fund performance.
Findings: The findings reveal that funds categorized under Article 8 of the SFDR do not exhibit significantly poorer performance compared to funds categorized under Article 6 during the initial phase after launch. On average, Article 8 funds even demonstrate positive returns earlier than their peers. However, the panel regression analysis suggests that Article 8 funds slightly underperform when compared to Article 6 funds over time.
Practical implications: While investors may not anticipate lower initial returns when opting for higher SFDR categories, they should nevertheless be aware of the shortcomings inherent in the existing SFDR labeling system, particularly in the context of unlisted real estate.
Originality: To the best of our knowledge, this study represents the first quantitative examination of unlisted real estate fund performance under the SFDR. By providing unique insights into the J-Curves of funds, our research contributes to the existing body of knowledge on the impact of sustainability regulations in the financial sector.

Keywords: *SFDR; Unlisted Real Estate; J-Curve, Performance.*

5.2 Introduction and Background

Building on the objectives set out in Action 9, "Strengthening sustainability disclosure and accounting rule-making" of the action plan on financing sustainable growth, the European Commission introduced the sustainable finance disclosure regulation (SFDR) in Regulation (EU) 2019/2088 in 2019. This regulation became mandatory for all member states in March 2021. The basis for these efforts is the EU's goal of directing private capital flows toward sustainability to meet the Paris Agreement climate goals. The regulation primarily affects European financial market participants and advisors and has the subsidiary goal of harmonizing sustainability requirements and reducing information asymmetries arising from sustainability risks and impacts in principal–agent relationships (such as between fund managers and investors). The SFDR provisions are intended to provide investors with the information they need to make informed decisions and implement their own sustainability efforts and are subject to a uniformly established legislation as opposed to voluntary scorings, benchmarks or certifications (European Parliament and Council, 2019). Aligned with the disclosure requirements of the EU taxonomy, the SFDR serves as a supplementary measure to complement climate policies, such as the EU-ETS or emission standards for buildings, rather than acting as a direct substitute (Schütze & Stede, 2021). Generally, the SFDR provisions require reporting based on three categories, each with different criteria to be considered and disclosed:

- Article 6 products, which have no specific sustainability features;
- Article 8 products, which have environmental or social characteristics; and
- Article 9 products, which pursue specific sustainability objectives.

The categorization into Articles 6, 8 and 9 is derived from the Principles-Based Requirements (PBR) of Level 1 SFDR, supplemented by the more detailed specifications provided in the Regulatory Technical Standards (RTS) of Level 2 SFDR. The RTS serve as guidelines for dealing practically with the scope, methods and disclosure requirements of each category. For Article 6 products, it is mandatory to include ESG risk assessments in the decision-making process or declare if no relevant risks are identified. This allows Article 6 funds to exhibit a spectrum of ESG ambitions, ranging from very low to very high. Article 8 products promote environmental or social characteristics and sustainable investments, without making sustainable investments the primary objective of the product. Additionally, sustainable investments do not significantly harm any objectives outlined in the EU-Taxonomy. Article 9 expands on the foundation set by Article 8 by requiring that properties in the portfolio meet strict sustainability criteria throughout the holding period. These criteria prioritize operational efficiency over the entire life-cycle, resulting in a significant portion of Article 9 funds being allocated to highly efficient new buildings (European Parliament and Council, 2019; INREV, 2023). The associated Principal Adverse Impacts (PAIs) arise from Annex I of the RTS and are divided into mandatory and voluntary indicators, of which, according to the European Association for Investors in Non-Listed Real Estate Vehicles (INREV, 2023), currently only seven are applicable to real estate. These indicators are intended to provide practical guidance on how to report on ESG issues and are currently the subject of intense debate.

Real estate is traditionally an asset class in which investments are predominantly made directly or through unlisted funds. Only about one-tenth of the $36.2 trillion global commercial real estate market is listed real estate (European Public Real Estate Association, 2023). Since approximately 40% of all greenhouse gas emissions can be directly attributed to the real estate sector, the alignment of private capital flows is thus a crucial factor for decarbonizing the property sector (European Commission, 2019). The introduction of the SFDR has resulted in existing investment funds embarking on an intensive rebranding process, seeking to capitalize on the advantages associated with early adoption of a higher SFDR category (Cremasco & Boni, 2022). Although there is no legal obligation for a minimum category for real estate funds, a category of at least Article 8 SFDR has been emerging as a basic market requirement for real estate-related investment products as most institutional investors attach increasing importance to Key Performance Indicators (KPIs) for measuring sustainability. With regard to voluntary standards like GRESB, sizable funds have typically functioned as early adopters, positioning themselves to enjoy the benefits of enhanced financial performance in the form of elevated total returns (Brounen & van der Spek, 2017). Extensive early adoption of higher SFDR categories is underscored by the observation that nearly half of real estate funds are already categorized under Article 8 or Article 9 SFDR (PricewaterhouseCoopers, 2023). This share might increase even further especially if there is evidence that a higher SFDR category is not economically disadvantageous compared to the default category. While Chegut et al. (2019) point out the absence of conclusive evidence concerning disparities in input costs between green and conventional properties, the prevailing sentiment among investors and developers indicates a widely held belief that sustainability is associated with higher upfront expenses. This perception is particularly prominent in the context of renovating existing buildings. From a fund management perspective, it is important to consider whether opting for a higher SFDR category will negatively impact initial fund performance. This assessment becomes even more crucial when factoring in the additional costs of fund reporting and documentation required by SFDR regulations. On the other hand, there might be advantages in the form of increased fund inflows due to positive signaling effects or a conservative risk-return profile resulting from the tendency toward newer and high-quality properties contained in Articles 8 and 9 funds. While voluntary disclosure in the realm of unlisted real estate funds has only gained traction in recent years, the introduction of the SFDR and the associated mandatory disclosure contains uninvestigated implications for investors. In contrast to prior research, predominantly centering on voluntary standards such as benchmarks and certifications, a notable research gap exists regarding the impact of disclosure regulations on fund performance in instances where fund managers do not explicitly and voluntarily opt for such standards. This study seeks to address this gap through a thorough empirical investigation utilizing a dataset comprising unlisted real estate funds categorized under the SFDR. Recognizing the nuanced and phase-specific nature of returns in the early stages of unlisted funds, we adopt a two-part methodology. The first part of our study focuses on the initial fund performance by assessing J-Curve characteristics based on SFDR categorization. Subsequently, employing panel data analysis, the second part seeks to ascertain the significance of the SFDR category's impact on the overall fund performance. At the same time, the study highlights potential shortcomings of the SFDR within the realm of real estate investment. The paper ends by presenting our conclusions.

5.3 Literature Review

An extensive body of existing literature deals with the implications of voluntary green building certifications at property level. In the commercial real estate sector, many studies reveal green premiums for office buildings in terms of rents, values, and prices (Wiley et al., 2010; Eichholtz et al., 2010; Fuerst & McAllister, 2011; Reichardt et al., 2012; Eichholtz et al., 2013; Chegut et al., 2014; Holtermans & Kok, 2019). A green premium in the income and capital component of properties can also be observed for both single- and multi-family homes (Kahn & Kok, 2014; Bond & Devine, 2016; Cadena & Thomson, 2021; Dell'Anna & Bottero, 2021). Moreover, certifications for building sustainability also exert a favorable impact on operational performance, as evidenced by an enhanced occupancy rates (Fuerst & McAllister, 2009; Devine & Kok, 2015). Green premiums are not confined solely to mature and highly liquid commercial real estate markets; they are also evident in emerging real estate markets (Lambourne, 2022; Costa et al., 2017). The green premium observed in both the income and capital components of properties may benefit the financial performance of funds with a higher SFDR category, which typically hold newer and more certified properties.

Another body in the field of sustainable real estate focus on performance and valuation metrics of listed real estate companies at corporate level. The impact of voluntary corporate sustainability reporting standards can be used to infer the potential implications of a mandatory reporting standards such as the SFDR. In both cases, a higher level of (perceived) disclosure leads to greater transparency for lenders and investors. Ansari et al. (2015) find evidence that sustainability reporting based on the GRI standards has a positive impact on both market values and abnormal returns of listed real estate companies. Feng and Wu (2021) provide similar findings for the GRESB reporting standard. Moreover, they show that higher disclosure levels are associated with better credit ratings, higher unsecured debt ratios and lower cost of debt financing. In addition to sustainability reporting, higher ESG ratings also have a positive impact on listed real estate companies. Brounen et al. (2021) indicate that more transparent ESG performance data in the EPRA sBPR database leads to a return premium for listed real estate companies since investors are able to make more informed investment decisions. However, Cajias et al. (2014) attribute differences in valuations of real estate companies to ESG concerns which negatively affect firm value rather than ESG strengths that lead to a sustainability premium. Furthermore, their findings suggest that high ESG ratings negatively impact returns as high ESG ratings may not translate into better operational performance in the short run. While public markets presently incorporate any ESG changes, the real impact on companies' operational performance may exhibit lagging effects.

It is not only the sustainability disclosure and rating of the company itself that has an impact at firm level but also the sustainability of the properties owned by the companies. Sustainability in this context usually refers to properties' environmental performance. Although there is evidence that non-green REITs outperform their green peers, most research highlights the benefits of environmentally friendly property features (Coën et al., 2018; Mariani et al., 2018). Holding greener properties can improve REITs' performance metrics at the operational level, such as return on assets, return on equity, and funds from operations (Eichholtz et al., 2012; Fuerst, 2015; Ooi & Dung, 2019; Morri et al., 2021). Furthermore, REITs with a greener portfolio are valued at a premium and are subject to less systematic risk (Eichholtz et al., 2012; Sah et al., 2013; Fuerst, 2015). Beyond higher firm values, Sah et al. (2013) observe better performance of stock values compared to less green REITs. Greener portfolios are also associated with lower borrowing costs (An & Pivo, 2020; Leutner et al., 2024). Both the bonds issued by REITs and the bonds trading on the secondary market exhibit smaller spreads (Eichholtz et al., 2019). Similarly, the cost of equity tends to be lower for greener REITs (Hsieh et al., 2020). Devine and Yönder (2023) reveal that the benefits experienced by greener REITs can, to some extent, be attributed solely to reputational factors. In the context of the SFDR, setting Article 8 as the minimum category for the market also creates a signaling effect. While directing more investor capital towards Article 8 and 9 funds may enhance financial performance, it does not necessarily imply superior sustainability performance (Cremasco & Boni, 2022).

Given the lack of available data, there are only a limited number of studies in the field of sustainable real estate that specifically address the impact of unlisted fund disclosure on returns. Existing research focuses on voluntary disclosure initiatives, yet there exists a gap in the literature addressing emerging mandatory standards such as SFDR. Brounen and van der Spek (2017) demonstrate that better total GRESB scores not only result in higher excess total returns but can also account for fluctuations in unlisted fund returns in the absence of other explanatory fund variables. This implies that even without considering crucial fund characteristics such as leverage, investors can elucidate variations using GRESB scores. However, it is worth noting that the effects between disclosure and fund performance are lagged, given that fund reporting is typically retrospective. Devine et al. (2022) find that both GRESB participation and GRESB scores have explanatory power for fund returns. Regarding total returns, the impact is discernible solely in the capital component, with no observable effect in the income component. Additionally, it is noteworthy that the relationships between fund returns and both GRESB participation and scores persist unaffected by local economic conditions. Research by Brounen and van der Spek (2023) shows that GRESB participation explains superior performance in unlisted fund returns. However, higher GRESB scores do not necessarily trigger higher returns. Rather, it is the specific environmental sub-scores of individual funds and their overall GRESB performance which contribute to an increase in returns.

Within the broader economic literature, there is evidence supporting the aforementioned signaling effects associated with a higher SFDR category and its impact on mutual funds. Becker et al. (2022) highlight that the implementation of SFDR has a discernible impact on fund inflows within the first four months after its introduction. In particular, Articles 8 and 9 funds experience positive net inflows compared to Article 6 funds. However, Ferriani (2023) notes that the benefits of a higher SFDR category, reflected in higher net inflows and improved returns, are limited to Article 9 funds. In addition to its financial impact, the effectiveness of SFDR from a sustainability perspective is crucial. Cremasco and Boni (2022) indicate that despite the innovations introduced by the SFDR, the European financial market is still characterized by ambiguity and indistinct categorization. Their results show that a distinction between Article 6 and Article 9 funds based on internal incentive structures for managers reveals a cross-category pattern of financial and sustainability performance. The authors express concerns about potential greenwashing practices due to the ambiguity surrounding actual sustainability within the SFDR categories in the financial context. The main section of our paper provides additional information on this matter and its practical implications.

To best of our knowledge, the present work is the first to explore the impact of mandatory regulations (such as the SFDR) on unlisted real estate funds, thereby creating a new research thread. We expect more future research in this field due to increasingly stringent standards for unlisted real estate.

5.4 Data and Descriptive Analysis

The dataset is provided by one of the largest European service capital investment companies specializing in the setup and management of regulated real estate funds on behalf of third parties. Thus, it can be inferred that the funds analyzed in our study represent a significant sample size compared to the market size of funds recently launched around the time of the introduction of the SFDR. The dataset constitutes an unbalanced panel dataset with monthly observations from March 2021 to March 2023. To allow for a rigorous empirical analysis, we cleaned the original dataset of 54 funds by excluding observations that prevent a comprehensive analysis over the entire review period. Our final dataset thus consists of 40 funds, all of which—for the purpose of examining initial periods and the relationship between fund launches and the SFDR—were newly launched starting from March 2021. In the context of our research questions and the scope of our paper, it is particularly important to consider only newly issued funds. This ensures direct comparability of initial performance, while maintaining uniform criteria for reporting requirements or the selection of properties to be acquired. Given the introduction of the SFDR in early 2021, a more extensive time series spanning several years is not yet available. The funds have a predominant investment focus on the German property market and are subdivided into core, core plus, value add, and opportunistic according to INREV fund style classifications. As shown in Figure 5.1, the entire sample exhibits a similar risk structure. Only one fund in the sample pursues a value-add investment strategy while the remaining funds range from core to core+. In general, the funds are also homogeneous in terms of size, property type, duration, and investor class, allowing for inter-fund comparisons.

Internal Rate of Return (IRR) and Bundesverband Investment und Asset Management (BVI) methods serve as performance indicators. The IRR is a money-weighted rate of return that describes the interest rate at which all future or actual cash flows are equal to the initial investments. In the context of real estate funds, the IRR pro-

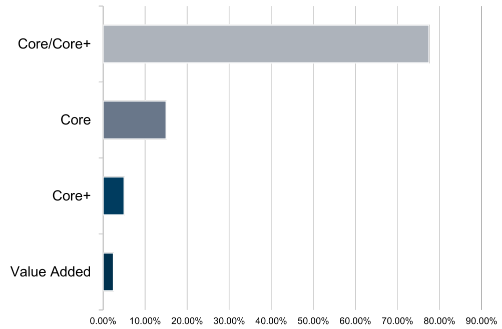


Figure 5.1: Distribution of Investment Styles in % of Total Funds

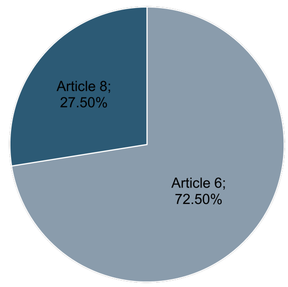


Figure 5.2: Distribution of SFDR Categories in % of Total Funds

vides insight into the annual capital return over the total duration of the investment (Hartman & Schafrick, 2004). The BVI-Method is based on the time-weighted rate of return methodology and is used to calculate the performance of investment funds. The performance is the rate of change between the invested capital at the beginning and end of the investment period. To ensure comparability between accumulating and distributing funds, the BVI-method assumes that distributed earnings are immediately reinvested in new fund shares (BVI, 2015).

Both methods have diametrically opposed explanatory power and objectives due to their different methodologies. Timing is a significant factor in the IRR (Hartman & Schafrick, 2004), whereas the BVI is divided into sub-periods, making time a subordinate factor (BVI, 2015). However, since the examined sample consists exclusively of young funds and new acquisitions of properties are regularly made at the beginning of the fund's term, time is a subordinate factor for the sample. In addition, the IRR reflects the actual performance taking into account individual transactions or fees, whereas the BVI represents an objective measure of value and thus ignores transaction costs. However, since the sample examined is made up of funds with a homogeneous cost structure, it is also conceivable to compare the funds using the IRR. Table 5.1 shows the descriptive statistics of the performance indicators. As expected, both performance indicators exhibit similar extreme points and trends.

Regarding the distribution by SFDR category, it is noticeable that most funds do not have any sustainability features (see Figure 5.2). In this context, our sample corresponds to the figures of the BVI, which still show that the majority of funds are of a conventional nature (BVI, 2022). In addition to the previously highlighted key variables such as BVI return, IRR, investment style and SFDR, our dataset includes insights into fund

Table 5.1: Descriptive Statistics

	Internal Rate of Return (IRR)					
SFDR Category	Mean	Median	Q1	Q3	Kurtosis	Skewness
Art. 6	3.73	1.20	0.00	2.81	30.93	2.75
Art. 8	1.80	0.32	-0.06	2.12	8.19	-0.98
	BVI-Method					
Art. 6	-1.40	0.57	-0.10	1.80	19.98	-3.53
Art. 8	1.81	0.00	-0.79	1.67	0.76	1.23

Note: n equals 547 for both IRR and BVI-Method, distributed across a total of 40 funds.

maturity, distinguishing between open-ended and closed-ended structures. Notably, 22.5% of the funds in our sample are closed-ended, while the majority, accounting for 77.5%, adopt an open-ended structure. It is crucial to note that the examined sample comprises specialized Alternative Investment Funds (AIFs) exclusively dedicated to real estate investments and accessible to (semi-)professional investors. In order to be able to estimate coefficients reliably and to depict the relationship between the variables correctly, the dataset is corrected for outliers (Rousseeuw & Leroy, 2005). In the context of our specific dataset and the targeted panel regression, an adjustment for the extreme values is therefore appropriate. For this purpose, we use the methodology of the interquartile range rule adjusted by a multiplier of 2.2 as recommended by Hoaglin and Iglewicz (1987).

5.5 Methodology

J-Curve analysis

To perform the J-Curve analysis in the first section of the paper, the funds' IRR time series are adjusted to $t = 0$ to depict the months following the launch of the funds. The monthly values are then connected using cubic spline interpolation to obtain a smoother curve that captures the trend of the data points. This yields a cubic spline function $S : [a, b] \to \mathbb{R}$ of piecewise third-degree polynomial functions $si(x)$ (Stoer & Bulirsch, 2002).

$$S(x) = \begin{cases} a_1x^3 + b_1x^2 + c_1x + d_1 & \text{if } x \in [x_1, x_2] \\ \quad\dots & \\ a_ix^3 + b_ix^2 + c_ix + d_i & \text{if } x \in [x_{i-1}, x_i] \\ \quad\dots & \\ a_nx^3 + b_nx^2 + c_nx + d_n & \text{if } x \in [x_{n-1}, x_n] \end{cases} \tag{5.1}$$

The polynomial functions (see Equation 8 in the Appendix) are chosen so that they pass through the data points at the ends of each interval, and that they have continuous first and second derivatives at the points at which the intervals intersect (Stoer & Bulirsch, 2002). Regarding endpoint derivatives, De Boor (2001) suggests that the not-a-knot condition approach is preferable to the free-end condition approach. The not-a-knot condition (see Equation 9 in the Appendix) involves selecting values for the end slopes so that the first and last interior nodes are not active. The result of cubic spline interpolation is a smooth curve that passes through the data points and tends to avoid the excessive oscillations commonly associated with high-degree polynomials (Stoer & Bulirsch, 2002). Hence, the method is particularly useful when the dataset is noisy or when there are gaps in the data, as it can provide a reasonable estimate of the underlying function even in these cases. In finance, cubic spline interpolation is a commonly used method, known for its ability to model yield curves, forward curves and other term structures with remarkable accuracy (Holton, 2013). Piecewise integration is applied to determine the negative area under the curve, the so-called "valley of tears", of the early-stage funds' IRR. This involves calculating the integral of the continuous spline function $S(x)$ between $t = 0$ and the break-even point at which $Si(xi) = 0$.

$$\int_a^b S(x)dx = \int_{x_1}^{x_2} s_1(x)dx + \cdots + \int_{x_{1-1}}^{x_i} s_1(x)dx + \cdots + \int_{x_{n-1}}^{x_n} s_n(x)dx \tag{5.2}$$

Panel regression analysis

To assess whether the SFDR category has a significant influence on the overall return of the individual funds, a panel regression is applied. Beyond the conventional pooled OLS model, panel regression analysis offers the flexibility of employing either random or fixed unobserved effects models, addressing potential bias related to omitted variables. These frameworks are commonly applied in related studies to isolate drivers of real estate fund performance (Farrelly & Matysiak, 2012; Fuerst & Matysiak, 2013; Devine et al., 2022). In contrast to the pooled OLS model, the random effects model incorporates a composite error term v_{it} that encompasses a random error component ε_i that is constant over time but differs across observations. The random error term captures the unobserved heterogeneity across observations that is not accounted for by the explanatory variables in the model and is added to the error term of the observations u_{it}. Thus, in the context of our dataset, ε_i also accounts for unobserved fund characteristics such as property type or gearing. The random effects model uses GLS estimation that takes advantage of the serial correlation present in the composite error term.

$$y_{it} = \alpha + \beta x_{it} + v_{it}, \quad v_{it} = \varepsilon_i + u_{it} \tag{5.3}$$

The fixed effects model also controls for time-invariant unobserved heterogeneity with the individual heterogeneity estimated as a parameter. If a time-invariant exogenous variable, such as the SFDR category, is included in the fixed effects model, it does not vary across time and is collinear with the individual fixed effects. As a result, the effects of the time-invariant observables are indistinguishable from the effects of the time-invariant unobservables. Any time-invariant explanatory variable would thus drop from the fixed effects model. Due to this drawback of the fixed effects model when including binary variables, it is not considered appropriate for the present analysis (Wooldridge, 2002). This is further confirmed by high p-values in the Hausman test (see Table 5.5).

To ensure efficient estimation within the GLS framework, additional assumptions are required that go beyond the OLS model. The Appendix provides a detailed overview of the model specification tests. Specifically, the tests proposed by Baltagi (2021) are presented in more detail to identify potential advantages of a random effects model over a fixed effects model and pooled OLS. This includes the examination of the Hausman test and the T. Breusch and Pagan (1980) Lagrange Multiplier (LM) test. Furthermore, the Appendix encompasses the panel diagnostic tests conducted.

5.6 Empirical Results and Implications

J-Curve analysis

Since the funds in our dataset were launched with the introduction of the SFDR at the beginning of 2021 and are therefore considered immature, we examine their initial return characteristics in more detail. Private equity or unlisted real estate funds typically exhibit high upfront costs related to fund administration, marketing, property acquisition, and management fees, which tend to impact the short-term IRR. These costs accumulate until sufficient fund inflows are generated to begin the investment phase. Moreover, investment managers do not immediately draw down investors' money upon launching the fund but rather gradually as the properties are acquired. Accordingly, investors are not invested in real estate but continue to be invested in cash for the time being. Both factors translate into a negative initial performance of unlisted real estate funds. As soon as the first properties are acquired and the cash flow stabilizes, the initially negative returns are amortized, and hence the so-called J-Curve emerges. Analyzing the funds' J-Curve can provide diverse insight into key figures, such as break-even point or risk of default (Diller et al., 2009). In the context of the SFDR, the question arises as to whether a J-Curve pattern can be discerned in the present dataset and whether this is more pronounced for Article 8 funds. A priori, one may assume that the upfront costs for Article 8 or 9 funds are higher than for Article 6 funds. More extensive due diligence for identifying suitable properties could lead to an increase in funds' upfront costs and further delay the drawdown of investors' capital. In addition, funds categorized under the SFDR must comply with article-specific requirements and disclose how sustainable features, risks and objectives are considered in the respective investment strategy. This implies additional expenses in the elaboration and preparation of the fund documentation, which is also subject to a separate review by the capital investment company. Finally, the reporting continuously monitors compliance with the respective SFDR category. Fundraising, documentation, property acquisition and reporting, therefore, might result in additional costs both before the launch of the fund and after the allocation of investors' money. This might contribute to a more pronounced J-Curve for Article 8 funds.

To investigate this relationship further, the IRR time series of the dataset were plotted as months after fund launch, using cubic spline interpolation. The equation of the cubic spline interpolation for both J-Curves is provided in the Appendix. As Figure 5.3 illustrates, the funds' IRR exhibits a J-Curve pattern as anticipated, regardless of the SFDR category. The analysis reveals that the break-even point of the J-Curve, i.e., the point at which the IRR turns positive is reached approximately 0.53 months earlier for Article 8 funds. This observation becomes especially clear in the direct comparison of the curves in Figure 5.4. To determine the negative area of the J-Curve, which is often referred to as the so-called valley of tears, the piecewise integrals of the spline functions are applied. The break-even point of the IRR and the starting point of the time series are defined as the upper and lower limits of the integrals, respectively. Figure 5.5 highlights the valley of tears for both SFDR categories. The results of the integrals show that Article 8 funds have a less pronounced negative area of the J-Curve, which is at odds with the previously assumed impact of the SFDR category on the initial IRR:

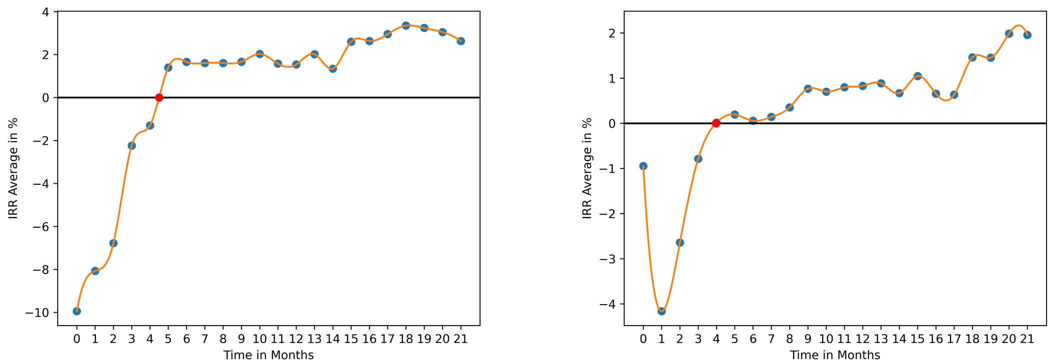
$$\int_0^{4,50} S(x)\,dx = -22.81 \qquad (5.4)$$

Integral Article 6

$$\int_0^{3,97} S(x)\,dx = -8.70 \qquad (5.5)$$

Integral Article 8

Since the dataset does not provide any information on the actual cost structures of the respective funds, no analysis can be carried out in this regard. However, the results allow for the conclusion that either the funds' upfront costs do not depend on the respective SFDR category, or that initially higher costs are amortized more quickly after the funds' cash flow stabilizes. The fund category under Article 8, therefore, does not entail any disadvantages regarding the initial IRR. To conduct a more in-depth analysis of the impact of the SFDR categories on funds' overall returns, a panel regression analysis is carried out.



(a) Article 6 J-Curve

(b) Article 8 J-Curve

Figure 5.3: Comparison of J-Curves

Panel regression

Our analysis comprises a total of four different individual random effects models, consisting of two base models and two extensions thereof. In our base models (Model 1a and 1b), the dependent variable y is defined by the return metrics IRR and BVI. The control variables of the panel regression models are represented by the funds' investment style and whether the funds are open- or closed-ended. Furthermore, a lagged return variable $y(t-1)$ is included to improve model accuracy (Farrelly & Matysiak, 2012; Fuerst & Matysiak, 2013). Relevant for answering the research question is the independent dummy variable "SFDR", in which $0 = Article$ 6 and $1 = Article$ 8. Investment style is a factor variable in which $0 = Core$ and the Open-/Closed-Ended variable is denoted as $Closed - Ended = 0$. In models 2a) and 2b), model accuracy is to be improved by adding external economic factors which may correlate with returns. GDP growth is typically considered to be a significant factor contributing to rental and value growth in the real estate market. However, GDP data is updated on a quarterly basis and can thus not be included as an explanatory variable in the present monthly

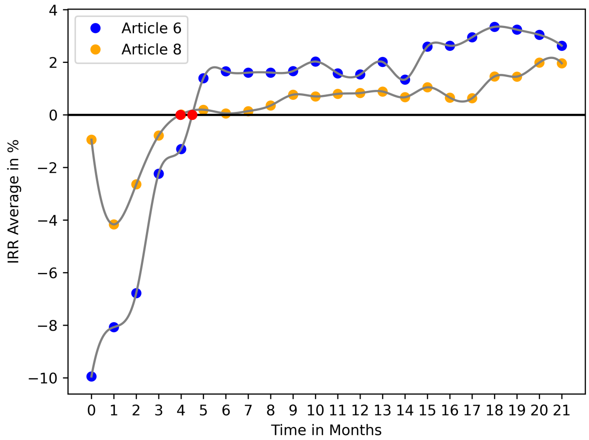
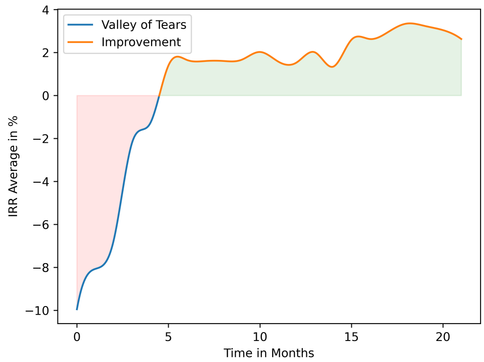
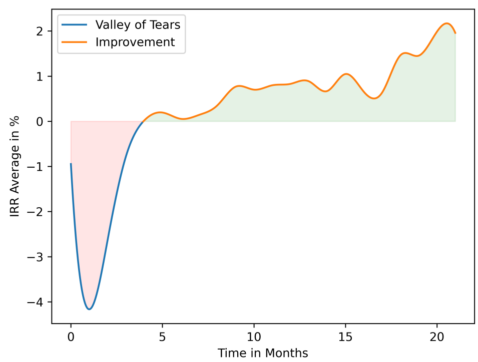


Figure 5.4: Superimposed J-Curves



(a) Article 6 J-Curve



(b) Article 8 J-Curve

Figure 5.5: Comparison of J-Curves with Valley of Tears

panel data analysis. Instead, the term spread between the German 10-year government bond yield and the 3-month yield is used (Federal Reserve Bank of St. Louis, 2023). The yield curve has predictive capabilities that go beyond the growth in economic activity and short-term interest rates, making it a valuable leading indicator for economic output and recessions (Estrella & Hardouvelis, 1991; Estrella et al., 2000; Chinn & Kucko, 2015). This sensitivity to the term structure of interest rates can also observed within the real estate context (Akimov et al., 2015). In addition to interest rates, which are endogenously determined by monetary policy, the exogenous availability and cost of debt capital can have an impact on the performance of real estate investments. For this reason, the relative change in credit margins for average loans to businesses is included as an independent variable (Deutsche Bundesbank, 2023). The movements in banks' terms and conditions for loans indicate the perception of risk in the lending sector.

The results of the model specification tests strongly suggest the use of individual random effects in both IRR and BVI models. The LM test yields p-values of 0.0000 in both cases, providing substantial evidence against the null hypothesis of the pooled OLS model. This implies that the random effects model is a more suitable choice. Additionally, the Hausman test with p-values of 0.9343 (IRR) and 0.7731 (BVI) fails to reject the null hypothesis of systematic differences between random and fixed effects, further reinforcing the choice of random effects. Subsequently, we examine whether there is a systematic change in the statistical properties of the IRR and BVI series. The low p-values of the unit root tests in Table 5.2 provide strong evidence against the presence of a unit root and hence support stationarity for both series. Regardless of the assumption of cross-sectional dependence, the results are unambiguous for both types of test. Only in the inverse normal Fisher-type test, the p-value for BVI is slightly above the significance level of 0.05 yet still under the 0.10 significance level. Regarding serial correlation, the results of the Wooldridge test align with expectations based on the relatively short time series. With p-values of 0.6345 (IRR) and 0.1599 (BVI), the null hypothesis of no serial correlation is not rejected. As the Breusch–Pagan test indicates the presence of heteroskedasticity in the models, it is corrected for by using a heteroskedasticity-consistent covariance estimation as suggested by Zeileis (2004). Given the absence of serial correlation in the models, robust standard errors are employed instead of heteroskedasticity- and autocorrelation-consistent clustered standard errors. While in time series data the information criteria AIC and BIC are applied to determine the optimal lag length, panel data analysis commonly relies on economic theory and alternative measures such as adjusted R^2 to determine the appropriate lag length. Accordingly, the lag length for both term and loan spread in models 2a) and 2b) is set at three months.

Table 5.2: Panel Unit Root Tests

	IRR	**BVI**
Cross-sectional independence		
Fisher-type: Inverse Chi-square	155.25	147.52
	(0.0000)	(0.0000)
Fisher-type: Inverse Normal	-2.9103	-1.2823
	(0.0018)	(0.0999)
Fisher-type: Logit	-4.4705	-2.0939
	(0.0000)	(0.0188)
Cross-sectional dependencies		
Pesaran Panel Unit Root Test	-2.0167	-2.0586
	(0.0100)	(0.0100)

Note: p-values are presented in parentheses.
Sample sizes: 437 (38 funds) for IRR and 434 (36 funds) for BVI.

The results of models 1a) and 1b) can be seen in Table 5.3. The SFDR Article 8 variable has a significant influence on both the IRR and the BVI return, with p-values smaller than 0.05 and 0.01, respectively. In contrast to the initial J-Curve analysis which suggests a performance premium for Article 8 funds up to the funds' break-even point, the results of the regression analysis indicate that Article 8 funds perform slightly worse than their Article 6 peers when the overall time series is considered. When controlling for differences in fund characteristics, the inclusion of the Article 8 dummy leads to a decrease in returns of about 0.5% points in both models. We attribute the minimal difference in results between IRR and BVI to the fact that the distinct methodologies of money-weighted and time-weighted returns, as explained earlier, have little impact due to the homogeneity and relative young vintage year of the funds under consideration. This pattern of returns for Article 8 funds is further supported by an eyeball test of the figures in the J-Curve analysis. Toward the end of the time series, the monthly IRR values for Article 8 funds consistently fall below the values of Article 6 funds.

One possible explanation for this observation is that Article 8 funds tend to offer a more conservative risk-return profile. The enhanced sustainability standards applied to Article 8 funds, along with the availability of more comprehensive data for reporting, can contribute to improved risk management capabilities. At the same time, Article 8 funds tend to acquire newer properties which meet the sustainability standards and thus exhibit less property-related risks. However, it should be noted that the analyzed

Table 5.3: Random Effects Model 1a) and Model 1b)

	IRR					BVI				
Coefficients	Estimate	Std. Error	Z-Value	Prob.		Estimate	Std. Error	Z-Value	Prob.	
(Intercept)	1.4146	0.4710	3.0032	0.0028	**	0.6140	0.2391	2.5683	0.0106	*
Core/Core+	-0.8140	0.3864	-2.1064	0.0357	*	-0.4444	0.2148	-2.0683	0.0392	*
Core+	1.1710	0.5996	1.9531	0.0515	.	1.1918	0.4892	2.4362	0.0153	*
Value Added	-0.8201	1.7453	-0.4699	0.6387		0.0371	0.6481	0.0573	0.9543	
Art. 8	-0.5026	0.2218	-2.2658	0.0240	*	-0.4873	0.1425	-3.4192	0.0007	***
Open Ended	0.1767	0.2334	0.7569	0.4495		0.4509	0.1491	3.0239	0.0026	**
y (t-1)	0.4475	0.1415	3.1626	0.0017	**	0.5338	0.0660	8.0818	0.0000	***
TSS	974.46					730.81				
RSS	379.31					203.39				
R-Squared	0.6108					0.7217				
Adj. R-Squared	0.6054					0.7178				
Chi-squared	668.41					1095.44				
Funds	38					36				
T-Max	21					21				
Sample size (n)	437					434				

Note: The significance levels are indicated as follows: *** $= p < 0.001$, ** $= p < 0.01$, * $= p < 0.05$, $\cdot = p < 0.1$.

time series represents only the initial period of the funds, i.e., the time when the first assets were being acquired. Hence, we assume that (default) risks are not realized to the extent that they would be perceptible in calculations of IRR or BVI. Although the dataset is corrected for outliers, the return metrics are also strongly influenced by initial property acquisitions and irregular cash flows. Consequently, both factors imply that Article 8 funds may be subject to a temporal bias, which could adversely affect the reported returns. Despite the absence of existing literature on SFDR and unlisted real estate returns, which limits our ability to interpret our findings, we can still draw parallels with relevant studies.

Our results can be aligned with the observations made by Devine et al. (2022), who found that voluntary sustainability reporting through GRESB does not confer advantages on funds' income component of total returns but rather affects the capital component. Given the brevity of our time series, it is unlikely that realized capital gains are already reflected in our return series. In a more comprehensive perspective, our observations also correspond to the findings presented by Cajias et al. (2014), which suggest that, whereas share prices incorporate ESG changes contemporaneously, the tangible effects on business operations exhibit a time lag. The authors note that negative ESG scores are associated with higher short-term returns whereas a positive impact of

ESG factors might tend to manifest more prominently in the long run. As such, both in the context of the implied risk pricing and the capital growth of the funds' underlying properties, our study lacks a longer time series to make conclusive statements about the long-term impact of the SFDR category on fund returns. When considering the results in the context of the regulatory requirements outlined in the SFDR, it is essential to address the potential influence of evolving reporting requirements on performance. The underlying dataset does not provide a breakdown of individual indicators (such as the classification of properties within the framework of the minimum "Do No Harm" indicators) but rather assumes a holistic approach encompassing all indicators. Brounen and van der Spek (2023) demonstrate that the voluntary GRESB Total Score does not have a significant impact on value, whereas an examination of individual scores can indeed influence value. Hence, it is conceivable that a more nuanced analysis within the context of the SFDR and unlisted real estate funds could offer further insights.

Regarding the remaining independent variables, it is noteworthy that, contrary to initial expectations, there is no observed performance advantage for core/core+ funds compared to pure core funds. Instead, core/core+ funds appear to exert a significant negative impact on performance. This outcome can be attributed to the nuanced categorization of these funds, which seems to blur the distinction between core and core+ strategies. On the other hand, a distinct premium of approximately 1.2% points is observed for pure core+ funds in comparison to core funds. With regard to the BVI, open-ended funds demonstrate a premium of 0.45% points compared to closed-ended funds. One possible explanation for this observation could be that closed-ended funds inherently place a stronger emphasis on the capital component, as their investment horizon is predetermined from the outset, in contrast to open-ended funds. However, as previously explained, our time series is expected to demonstrate minimal to nonexistent positive influence from the capital component, with the emphasis primarily placed on the income component. The ambiguous influences of the fund characteristics across models can also be attributed to the fact, as already noted by Stevenson (2006), that, in the case of very young funds, the specified fund objectives such as leverage, diversification, or investment volume may not have been fully realized yet. The inclusion of the lagged return variable $y(t-1)$ proves to be applicable, as indicated by the positive and statistically significant results.

The outcomes of models 2a) and 2b) presented in Table 5.4 provide complete affirmation of the findings obtained from the previous models. While the market variables we have highlighted can potentially serve as leading indicators for predicting returns of unlisted real estate funds, the inclusion of external variables does not offer any new revelations. This observation can be attributed to the relatively new launch of the

Table 5.4: Random Effects Model 2a) and Model 2b)

	IRR					BVI				
Coefficients	Estimate	Std. Error	Z-Value	Prob.		Estimate	Std. Error	Z-Value	Prob.	
(Intercept)	1.3050	0.4821	2.7070	0.0071	**	0.5143	0.2898	1.7749	0.0766	.
Core/Core+	-0.8286	0.3988	-2.0780	0.0383	*	-0.4516	0.2276	-1.9839	0.0479	*
Core+	1.0831	0.6051	1.7900	0.0742	.	1.1406	0.5132	2.2225	0.0268	*
Value Added	-1.0369	1.7592	-0.5894	0.5559		-0.0977	0.6414	-0.1524	0.8790	
Art. 8	-0.5374	0.2354	-2.2832	0.0229	*	-0.5232	0.1485	-3.5242	0.0005	***
Open Ended	0.1659	0.2333	0.7112	0.4773		0.4666	0.1531	3.0476	0.0025	**
y (t-1)	0.4252	0.1471	2.8910	0.0040	**	0.5072	0.0662	7.6670	0.0000	***
Term Spread (t-3)	9.0415	14.6195	0.6185	0.5366		6.8916	13.6005	0.5067	0.6126	
Loan Spread (t-3)	8.6171	6.0201	1.4314	0.1531		6.3647	4.0149	1.5853	0.1136	
TSS	950.65					691.71				
RSS	366.58					194.33				
R-Squared	0.6145					0.7191				
Adj. R-Squared	0.6073					0.7138				
Chi-squared	674.61					1074.23				
Funds	38					36				
T-Max	21					21				
Sample size (n)	437					434				

Note: The significance levels are indicated as follows: $^{***} = p < 0.001$, $^{**} = p < 0.01$, $^{*} = p < 0.05$, $\cdot = p < 0.1$.

funds, with returns primarily influenced by property acquisition and volatile cash flows rather than the overall economy or financing conditions. This supposition is also reflected in the adjusted R^2, which shows no significant increase. The fact that external economic variables do not appear significant aligns with the findings of Devine et al. (2022), which suggest that the impact of GRESB scores is independent of local economic conditions. In our case, we attribute this phenomenon firstly to the reaction time of market participants, as indicated by our comparatively short time series. Secondly, we ascribe this to the long-term investment horizon of our sample, which is less dependent on local economic conditions in the short term. Regarding the SFDR category, Article 8 funds exhibit a slightly greater adverse impact on performance. Overall, the results obtained remain consistent.

Beyond the purely financial perspective, it is valuable to examine the tangible effectiveness of the SFDR for sustainable investment strategies. While there is some evidence that increased sustainability transparency is encouraging investment funds to raise their ESG ambitions, the SFDR is falling somewhat short of its intended purpose, particularly in the area of non-listed real estate investments (Becker et al., 2022; Ferriani, 2023).

INREV not only published a comprehensive document outlining the shortcomings of the SFDR but also submitted a response to the European Commission's SFDR consultation in December 2023, highlighting the key weaknesses of the SFDR in the real estate asset class (INREV, 2023; INREV, 2023). A key concern in the documents is the difficulty of applying SFDR to non-static asset classes such as real estate. Real estate investments involve a business strategy that is executed over an extended investment horizon. This is particularly evident in sustainable real estate investment products, which focus on transition strategies of property portfolios. Transition strategies involve retrofitting older buildings to enhance energy efficiency, phasing out on-site fossil fuel combustion, and installing on-site renewable energy systems. The implementation of these measures requires integration into the funds' Capital Expenditure (CapEx) planning to ensure an optimal balance between environmental payback (trade-off between embodied and operational carbon) and economic payback (trade-off between CapEx and operational savings). Of particular importance in supporting funds' transition strategies is the reallocation of capital. While the SFDR is intended to shift capital to more sustainable investments, a large proportion of older and inefficient properties is held in Article 6 funds. With Article 8 funds as a minimum requirement for many investors and banks, there is a risk that capital will not be made available for the transformation of the existing, inefficient property stock. Furthermore, as noted by Cremasco and Boni (2022) and INREV (2023), there is a risk of greenwashing when the categories of the SFDR are rigidly perceived as green labels, with Article 8 being categorized as light green and Article 9 as deep green. However, these categories do not guarantee actual sustainability performance. Hence, to ensure effective risk management, it is necessary to evaluate not only the SFDR category of a fund but also its business plan and sustainability objectives. It is important to note that the SFDR is a dynamic framework that is subject to continuous changes, particularly in response to pressures from various associations and initiatives. Investment decisions based solely on the SFDR category can turn into risk positions in case of possible changes in the SFDR labeling regime.

5.7 Conclusions

The results of the present study indicate that Article 8 funds do not significantly underperform Article 6 funds in the initial phase after launch. On the contrary, Article 8 funds achieve positive returns even before their Article 6 peers. As typical for unlisted funds, the development of the funds' IRR in our dataset follows a J-Curve pattern over time. A priori, we assumed that Article 8 funds, especially in the initial phase, would exhibit a more pronounced J-Curve with longer periods of negative returns. However, the results of the J-Curve analysis refute this hypothesis. This may reflect either that the initial costs of the funds are not affected by the SFDR category or that the higher initial costs are recovered more quickly, for example by attracting more investors and hence generating more fund inflows. The results of the panel regression analysis confirm, across models, that the SFDR category has a significant influence on both the IRR and BVI returns. In contrast to the examination of funds' J-Curve, the models in the panel regression analysis show that Article 8 funds perform slightly worse than their Article 6 peers when considered over the entire period. This can be attributed to several potential factors. We assume that properties eligible for Article 8 SFDR have a more conservative risk-return profile as they tend to be newer and more sustainable properties. The availability of comprehensive data required for SFDR reporting will, in any case, enable better risk management at the fund level. Article 8 funds may exhibit a more conservative risk-return profile, resulting in lower returns. Additionally, any incremental capital gains may materialize with a time lag and become apparent in the long run. Finally, it is possible that some risks related to the properties have not been fully captured in the IRR or BVI return figures yet, and the returns themselves may be biased due to property acquisitions and volatile cash flows at the initial stage of the funds. As real estate funds typically have a long-term investment horizon, this paper serves as a foundation for further research. Future analysis could provide valuable insights into the impact of SFDR disclosure on non-listed real estate funds, in particular whether the SFDR has a systematic impact on financial markets within the EU compared to other markets where no regulatory action has been taken. Current evidence from the industry suggests that the SFDR falls short of its objective with regard to unlisted real estate. It remains to be seen to what extent policymakers will accelerate the decarbonization of the property sector through targeted regulation. Real estate stakeholders should prepare for dynamically changing regulatory conditions and continue to monitor a broad range of sustainability KPIs to meet future market requirements.

5.8 Appendix

The following equation reflects the cubic spline for Article 6 funds:

$$
s(x) = \begin{cases}
1.2970 \cdot x^3 - 8.0726 \cdot x^2 + 1.7016 \cdot 10^1 \cdot x - 2.0192 \cdot 10^1, & \text{if } x \in [1,2], \\
1.2970 \cdot x^3 - 8.0726 \cdot x^2 + 1.7016 \cdot 10^1 \cdot x - 2.0192 \cdot 10^1, & \text{if } x \in (2,3], \\
-2.6593 \cdot x^3 + 2.7534 \cdot 10^1 \cdot x^2 - 8.9805 \cdot 10^1 \cdot x + 8.6629 \cdot 10^1, & \text{if } x \in (3,4], \\
2.4899 \cdot x^3 - 3.4256 \cdot 10^1 \cdot x^2 + 1.5736 \cdot 10^2 \cdot x - 2.4292 \cdot 10^2, & \text{if } x \in (4,5], \\
-1.9344 \cdot x^3 + 3.2108 \cdot 10^1 \cdot x^2 - 1.7446 \cdot 10^2 \cdot x + 3.1011 \cdot 10^2, & \text{if } x \in (5,6], \\
1.0549 \cdot x^3 - 2.1699 \cdot 10^1 \cdot x^2 + 1.4838 \cdot 10^2 \cdot x - 3.3557 \cdot 10^2, & \text{if } x \in (6,7], \\
-1.7048 \cdot 10^{-1} \cdot x^3 + 4.0341 \cdot x^2 - 3.1754 \cdot 10^1 \cdot x + 8.4739 \cdot 10^1, & \text{if } x \in (7,8], \\
6.1301 \cdot 10^{-4} \cdot x^3 - 7.2122 \cdot 10^{-2} \cdot x^2 + 1.0957 \cdot x - 2.8596, & \text{if } x \in (8,9], \\
1.6596 \cdot 10^{-1} \cdot x^3 - 4.5366 \cdot x^2 + 4.1276 \cdot 10^1 \cdot x - 1.2340 \cdot 10^2, & \text{if } x \in (9,10], \\
-4.1380 \cdot 10^{-1} \cdot x^3 + 1.2856 \cdot 10^1 \cdot x^2 - 1.3265 \cdot 10^2 \cdot x + 4.5636 \cdot 10^2, & \text{if } x \in (10,11], \\
3.7522 \cdot 10^{-1} \cdot x^3 - 1.3181 \cdot 10^1 \cdot x^2 + 1.5376 \cdot 10^2 \cdot x - 5.9382 \cdot 10^2, & \text{if } x \in (11,12], \\
1.2648 \cdot 10^{-1} \cdot x^3 - 4.2267 \cdot x^2 + 4.6306 \cdot 10^1 \cdot x - 1.6400 \cdot 10^2, & \text{if } x \in (12,13], \\
-7.6730 \cdot 10^{-1} \cdot x^3 + 3.0631 \cdot 10^1 \cdot x^2 - 4.0684 \cdot 10^2 \cdot x + 1.7996 \cdot 10^3, & \text{if } x \in (13,14], \\
1.2809 \cdot x^3 - 5.5394 \cdot 10^1 \cdot x^2 + 7.9751 \cdot 10^2 \cdot x - 3.8207 \cdot 10^3, & \text{if } x \in (14,15], \\
-1.2926 \cdot x^3 + 6.0415 \cdot 10^1 \cdot x^2 - 9.3963 \cdot 10^2 \cdot x + 4.8650 \cdot 10^3, & \text{if } x \in (15,16], \\
7.5240 \cdot 10^{-1} \cdot x^3 - 3.7746 \cdot 10^1 \cdot x^2 + 6.3095 \cdot 10^2 \cdot x - 3.5114 \cdot 10^3, & \text{if } x \in (16,17], \\
-2.1142 \cdot 10^{-1} \cdot x^3 + 1.1409 \cdot 10^1 \cdot x^2 - 2.0468 \cdot 10^2 \cdot x + 1.2239 \cdot 10^3, & \text{if } x \in (17,18], \\
-1.2627 \cdot 10^{-1} \cdot x^3 + 6.8103 \cdot x^2 - 1.2191 \cdot 10^2 \cdot x + 7.2723 \cdot 10^2, & \text{if } x \in (18,19], \\
1.4802 \cdot 10^{-1} \cdot x^3 - 8.8241 \cdot x^2 + 1.7514 \cdot 10^2 \cdot x - 1.1541 \cdot 10^3, & \text{if } x \in (19,20], \\
-5.5360 \cdot 10^{-2} \cdot x^3 + 3.3788 \cdot x^2 - 6.8917 \cdot 10^1 \cdot x + 4.7295 \cdot 10^2, & \text{if } x \in (20,21], \\
-5.5360 \cdot 10^{-2} \cdot x^3 + 3.3788 \cdot x^2 - 6.8917 \cdot 10^1 \cdot x + 4.7295 \cdot 10^2, & \text{if } x \in (21,22].
\end{cases} \tag{5.6}
$$

Table 5.5: Hausman Tests

Hausman tests	IRR	BVI
Chi-Square	0.8304	1.7966
	(0.9343)	(0.7731)
Degrees of Freedom	4	4

Note: p-values are presented in parentheses.
Sample sizes: 437 (38 funds) for IRR and 434 (36 funds) for BVI.

The following equation reflects the cubic spline for Article 8 funds:

$$s(x) = \begin{cases} -8.4637 \cdot 10^{-1} \cdot x^3 + 7.4483 \cdot x^2 - 1.9639 \cdot 10^1 \cdot x + 1.2091 \cdot 10^1, & \text{if } x \in [1,2], \\ -8.4637 \cdot 10^{-1} \cdot x^3 + 7.4483 \cdot x^2 - 1.9639 \cdot 10^1 \cdot x + 1.2091 \cdot 10^1, & \text{if } x \in (2,3], \\ -1.7337 \cdot 10^{-1} \cdot x^3 + 1.3913 \cdot x^2 - 1.4681 \cdot x - 6.0802, & \text{if } x \in (3,4], \\ 1.5093 \cdot 10^{-1} \cdot x^3 - 2.5003 \cdot x^2 + 1.4098 \cdot 10^1 \cdot x - 2.6835 \cdot 10^1, & \text{if } x \in (4,5], \\ -7.4049 \cdot 10^{-4} \cdot x^3 - 2.2521 \cdot 10^{-1} \cdot x^2 + 2.7230 \cdot x - 7.8765, & \text{if } x \in (5,6], \\ 1.5888 \cdot 10^{-1} \cdot x^3 - 3.0983 \cdot x^2 + 1.9962 \cdot 10^1 \cdot x - 4.2354 \cdot 10^1, & \text{if } x \in (6,7], \\ -9.1780 \cdot 10^{-2} \cdot x^3 + 2.1655 \cdot x^2 - 1.6885 \cdot 10^1 \cdot x + 4.3622 \cdot 10^1, & \text{if } x \in (7,8], \\ 1.0800 \cdot 10^{-1} \cdot x^3 - 2.6292 \cdot x^2 + 2.1473 \cdot 10^1 \cdot x - 5.8665 \cdot 10^1, & \text{if } x \in (8,9], \\ -2.6341 \cdot 10^{-1} \cdot x^3 + 7.3988 \cdot x^2 - 6.8780 \cdot 10^1 \cdot x + 2.1209 \cdot 10^2, & \text{if } x \in (9,10], \\ 2.6395 \cdot 10^{-1} \cdot x^3 - 8.4219 \cdot x^2 + 8.9427 \cdot 10^1 \cdot x - 3.1526 \cdot 10^2, & \text{if } x \in (10,11], \\ -1.5044 \cdot 10^{-1} \cdot x^3 + 5.2529 \cdot x^2 - 6.0995 \cdot 10^1 \cdot x + 2.3628 \cdot 10^2, & \text{if } x \in (11,12], \\ 1.1076 \cdot 10^{-1} \cdot x^3 - 4.1502 \cdot x^2 + 5.1842 \cdot 10^1 \cdot x - 2.1506 \cdot 10^2, & \text{if } x \in (12,13], \\ -2.0266 \cdot 10^{-1} \cdot x^3 + 8.0732 \cdot x^2 - 1.0706 \cdot 10^2 \cdot x + 4.7352 \cdot 10^2, & \text{if } x \in (13,14], \\ 4.0108 \cdot 10^{-1} \cdot x^3 - 1.7284 \cdot 10^1 \cdot x^2 + 2.4794 \cdot 10^2 \cdot x - 1.1832 \cdot 10^3, & \text{if } x \in (14,15], \\ -5.3434 \cdot 10^{-1} \cdot x^3 + 2.4810 \cdot 10^1 \cdot x^2 - 3.8347 \cdot 10^2 \cdot x + 1.9739 \cdot 10^3, & \text{if } x \in (15,16], \\ 3.7377 \cdot 10^{-1} \cdot x^3 - 1.8779 \cdot 10^1 \cdot x^2 + 3.1396 \cdot 10^2 \cdot x - 1.7457 \cdot 10^3, & \text{if } x \in (16,17], \\ 1.7844 \cdot 10^{-1} \cdot x^3 - 8.8178 \cdot x^2 + 1.4461 \cdot 10^2 \cdot x - 7.8609 \cdot 10^2, & \text{if } x \in (17,18], \\ -6.1010 \cdot 10^{-1} \cdot x^3 + 3.3764 \cdot 10^1 \cdot x^2 - 6.2186 \cdot 10^2 \cdot x + 3.8127 \cdot 10^3, & \text{if } x \in (18,19], \\ 5.8007 \cdot 10^{-1} \cdot x^3 - 3.4076 \cdot 10^1 \cdot x^2 + 6.6710 \cdot 10^2 \cdot x - 4.3506 \cdot 10^3, & \text{if } x \in (19,20], \\ -3.3612 \cdot 10^{-1} \cdot x^3 + 2.0896 \cdot 10^1 \cdot x^2 - 4.3233 \cdot 10^2 \cdot x + 2.9789 \cdot 10^3, & \text{if } x \in (20,21], \\ -3.3612 \cdot 10^{-1} \cdot x^3 + 2.0896 \cdot 10^1 \cdot x^2 - 4.3233 \cdot 10^2 \cdot x + 2.9789 \cdot 10^3, & \text{if } x \in (21,22]. \end{cases} \tag{5.7}$$

Table 5.6: Breusch-Pagan Tests for Lagrange Multiplier and Heteroscedasticity

Test Type	**IRR**	**BVI**
Lagrange Multiplier Test		
Chi-Square	142.00 (0.0000)	51.99 (0.0000)
Degrees of Freedom	1	1
Heteroscedasticity Test		
BP	13.29 (0.0041)	472.46 (0.0000)
Degrees of Freedom	3	3

Note: *p*-values are presented in parentheses.
Sample sizes: 437 (38 funds) for IRR and 434 (36 funds) for BVI.

Table 5.7: Wooldridge Test

Wooldridge tests	**IRR**	**BVI**
F	0.2263 (0.6345)	1.9815 (0.1599)
Degrees of Freedom	1	1

Note: *p*-values are presented in parentheses.
Sample sizes: 437 (38 funds) for IRR and 434 (36 funds) for BVI.

Cubic Spline Interpolation: Intersection Condition.

$$\begin{gathered} s_i\left(x_{i-1}\right)=y_{i-1}, s_i\left(x_i\right)=y_i \\ \frac{d^1}{d x^1} s_{i-1}\left(x_i\right)=\frac{d^1}{d x^1} s_i\left(x_i\right), \quad \frac{d^2}{d x^2} s_{i-1}\left(x_i\right)=\frac{d^2}{d x^2} s_i\left(x_i\right) \end{gathered} \tag{5.8}$$

Cubic Spline Interpolation: Not-a-Knot Condition.

$$\begin{gathered} \frac{d^3}{d x^3} s_1\left(x_1\right)=\frac{d^3}{d x^3} s_2\left(x_2\right) \quad \mid x_1=x_2 \\ \frac{d^3}{d x^3} s_{n-1}\left(x_{n-1}\right)=\frac{d^3}{d x^3} s_n\left(x_n\right) \mid x_{n-1}=x_n \end{gathered} \tag{5.9}$$

Model selection and panel diagnostics:
Under the strict exogeneity assumption of the GLS framework the composite error term is random and uncorrelated with the explanatory variables $E(v_{it}/x_{it}) = 0$. Furthermore, it is assumed that the idiosyncratic errors u_{it} have a constant and unconditional variance $E(u_{it}^2) = \sigma_u^2$ and exhibit no autocorrelation among different observations $E(u_{it}/u_{is}) = 0$ (Wooldridge, 2002). Baltagi (2021) suggests employing the Hausman test statistic, which incorporates the assumption $E(v_{it}/x_{it}) = 0$, to ascertain whether the GLS estimator $\hat{\beta}_{GLS}$ of the random effects model is consistent and efficient compared to the estimator of the within transformation of the fixed effects model $\hat{\beta}_{Within}$. It has the null hypothesis of no systematic difference between the estimates obtained from both random and fixed effects model. The test statistic follows an asymptotical distribution χ_K^2 with K being a measure for the dimension of the slope vector β.

$$m_1=\hat{q}_1^{\prime}\left[\operatorname{var}\left(\hat{q}_1\right)\right]^{-1} \hat{q}_1, \quad \hat{q}_1=\hat{\beta}_{GLS}-\hat{\beta}_{within} \tag{5.10}$$

The T. Breusch and Pagan (1980) LM test as proposed by Baltagi (2021) is applied to determine the presence of random effects in the model establishing the suitability of either a random effects model or a pooled ordinary least squares model. Under the null hypothesis of no effects, the variance of the random effect σ_ε^2 is zero. The LM test statistic is asymptotically distributed as χ^2 with $\hat{u}_{it}$ as the residual of the pooled ordinary least squares model.

$$LM=\frac{NT}{2(T-1)}\left[\frac{\sum_{i=1}^N\left(\sum_{t=1}^T \hat{u}_{it}\right)^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2}-1\right]^2 \tag{5.11}$$

T. S. Breusch and Pagan (1979) developed another test based on the Lagrange multiplier which is used to detect heteroscedasticity in the model and, therefore, a change in the variance of the residuals. It has the null hypothesis of homoskedasticity and is asymptotically distributed as χ^2. According to T. S. Breusch and Pagan (1979), using an auxiliary regression with $g_t = \hat{\sigma}^2 \hat{u}_t^2$ where $\hat{\sigma}^2 = N^{-1} \sum \hat{u}_t^2$, the LM test statistic can be determined as $2^{-1}ESS$.

Serial correlation is another issue that may cause the regression results to be biased. Although panel datasets with short time series are less prone to serial correlation, the Wooldridge (2002) first-difference test as suggested by Drukker (2003) is applied. It is a straight-forward test for unbalanced panels in both fixed and random effects one-way models with the null hypothesis of no serial correlation in the idiosyncratic errors.

A fundamental assumption in panel regression analysis is the stationarity of the underlying data over time, characterized by the absence of a unit root. As a result, the statistical properties of the data, including the mean, variance and autocorrelation, are presumed to remain constant throughout the observation period. According to Maddala and Wu (1999) and Choi (2001), the Fisher-type test is a superior alternative to the panel unit root tests by Levin-Lin (LL) and Im-Pesaran-Shin (IPS). Unlike the IPS test, the Fisher-type test can be applied to unbalanced panels. The Fisher-type test combines the p-values p_i of N independent unit root tests. It assumes that all series under the null are non-stationary based on the p-values of the individual unit root tests, while the alternative is that at least one series in the panel is stationary. As proposed by Choi (2001), three versions of the Fisher-type test based on the inverse chi-square, inverse normal and logit test are performed in this paper. While the inverse chi-square and inverse normal test statistics have an exact asymptotic distribution, the logit test statistic follows an approximated asymptotic distribution.

$$\begin{aligned} &P = -2\sum_{i=1}^{N} \ln\left(p_i\right) \sim \chi^2_N, \quad Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}\left(p_i\right) \sim N(0,1), \\ &L = \sum_{i=1}^{N} \ln\left(\frac{p_i}{1-p_i}\right) \text{ with } L^* = \sqrt{k}L \Rightarrow t_{5N+4} \text{ where } k = \frac{3(5N+4)}{\pi^2 N(5N+2)} \end{aligned} \tag{5.12}$$

The Fisher-type tests assume cross sectional independency, which requires that the individual cross-sections in the dataset do not correlate with each other (Choi, 2001). Based on the data used, this assumption may be unrealistic and, therefore, violated. For this reason, we additionally apply Pesaran's unit root test to avoid spurious inferences due to possibly skewed test results. The test employs a cross-sectional augmented Dickey-Fuller (ADF) approach, which involves augmenting the regressions with the cross-sectional averages of lagged levels and first-differences of the individual series (Pesaran, 2007).

5.9 References

Akimov, A., Stevenson, S., & Zagonov, M. (2015). Public real estate and the term structure of interest rates: a cross-country study. *The Journal of Real Estate Finance and Economics*, *51*, 503–540.

An, X., & Pivo, G. (2020). Green buildings in commercial mortgage-backed securities: The effects of LEED and energy star certification on default risk and loan terms. *Real Estate Economics*, *48*(1), 7–42.

Ansari, N., Cajias, M., & Bienert, S. (2015). The value contribution of sustainability reporting-an empirical evidence for real estate companies. *Journal of Finance and Risk Perspectives*, *4*(4), 190–205.

Baltagi, B. H. (2021). *Econometric analysis of panel data* (Vol. 6). Springer.

Becker, M. G., Martin, F., & Walter, A. (2022). The power of ESG transparency: The effect of the new SFDR sustainability labels on mutual funds and individual investors. *Finance Research Letters*, *47*.

Bond, S. A., & Devine, A. (2016). Certification matters: Is green talk cheap talk? *The Journal of Real Estate Finance and Economics*, *52*(2), 117–140.

Breusch, T., & Pagan, A. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, *47*(1), 239–253.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, 1287–1294.

Brounen, D., Marcato, G., & Op't Veld, H. (2021). Pricing ESG equity ratings and underlying data in listed real estate securities. *Sustainability*, *13*(4), 2037.

Brounen, D., & van der Spek, M. (2017). Sustainable insights in private equity performance. *Advance online publication*. Retrieved from www.tias.edu/docs/default-source/Kennisartikelen/gresb-lessons_db07172017.pdf

Brounen, D., & van der Spek, M. (2023). ESG insights in non-listed fund performance. *INREV-GRESB Online*. Retrieved from https://www.inrev.org/system/files/2023-05/INREV-GRESB-ESG-Insights-in-Non-listed-Fund-Performance_0.pdf

BVI. (2015). Die BVI Methode: Wertentwicklungsberechnung von Investmentfonds [the BVI method: calculating the performance of investment funds]. *Published Online*.

BVI. (2022). Der nachhaltige Fondsmarkt im zweiten Quartal 2022 [the sustainable fund market in the second quarter of 2022]. *BVI Fokus Nachhaltigkeit*.

Cadena, A., & Thomson, T. A. (2021). The value of "green" in resale residential real estate: Premium by neighborhood value quintile, homestead status and year. *Journal of Sustainable Real Estate*, *13*(1), 12–29.

Cajias, M., Fuerst, F., McAllister, P., & Nanda, A. (2014). Do responsible real estate companies outperform their peers? *International Journal of Strategic Property Management*, *18*(1), 11–27.

Chegut, A., Eichholtz, P., & Kok, N. (2014). Supply, demand and the value of green buildings. *Urban studies*, *51*(1), 22–43.

Chegut, A., Eichholtz, P., & Kok, N. (2019). The price of innovation: An analysis of the marginal cost of green buildings. *Journal of Environmental Economics and Management*, *98*.

Chinn, M., & Kucko, K. (2015). The predictive power of the yield curve across countries and time. *International Finance*, *18*(2), 129–156.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, *20*(2), 249–272.

Coën, A., Lecomte, P., & Abdelmoula, D. (2018). The financial performance of green REITs revisited. *Journal of Real Estate Portfolio Management*, *24*(1), 95–105.

Costa, O., Fuerst, F., Robinson, S. J., & Mendes-Da-Silva, W. (2017). Are green labels more valuable in emerging real estate markets? *Available at SSRN 2982381*.

Cremasco, C., & Boni, L. (2022). Is the European Union (EU) Sustainable Finance Disclosure Regulation (SFDR) effective in shaping sustainability objectives? An analysis of investment funds' behaviour. *Journal of Sustainable Finance & Investment*, 1–19.

De Boor, C. (2001). *A practical guide to splines* (Vol. 27). Springer-Verlag New York.

Dell'Anna, F., & Bottero, M. (2021). Green premium in buildings: Evidence from the real estate market of Singapore. *Journal of Cleaner Production*, *286*, 125327.

Deutsche Bundesbank. (2023). Bank Lending Survey for Germany. *Published Online*. Retrieved from https://www.bundesbank.de/resource/blob/626216/a0900b10382636c34c0f56b6d7d02c27/mL/bank-lending-survey-nettosaldo-data.pdf

Devine, A., & Kok, N. (2015). Green certification and building performance: Implications for tangibles and intangibles. *The Journal of Portfolio Management*, *41*(6), 151–163.

Devine, A., Sanderford, A., & Wang, C. (2022). Sustainability and private equity real estate returns. *The Journal of Real Estate Finance and Economics*, 1–27.

Devine, A., & Yönder, E. (2023). Impact of environmental investments on corporate financial performance: Decomposing valuation and cash flow effects. *The Journal of Real Estate Finance and Economics*, *66*(4), 778–805.

Diller, C., Herger, I., & Wulff, M. (2009). The private equity j-curve: cash flow considerations from primary and secondary points of view. *Investing in private equity*.

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *The Stata Journal*, *3*(2), 168–177.

Eichholtz, P., Holtermans, R., Kok, N., & Yönder, E. (2019). Environmental performance and the cost of debt: Evidence from commercial mortgages and reit bonds. *Journal of Banking & Finance*, *102*, 19–32.

Eichholtz, P., Kok, N., & Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *American Economic Review*, *100*(5), 2492–2509.

Eichholtz, P., Kok, N., & Quigley, J. M. (2013). The economics of green building. *Review of Economics and Statistics*, *95*(1), 50–63.

Eichholtz, P., Kok, N., & Yonder, E. (2012). Portfolio greenness and the financial performance of REITs. *Journal of International Money and Finance*, *31*(7), 1911–1929.

Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, *46*(2), 555–576.

Estrella, A., Rodrigues, A. P., & Schich, S. (2000). How stable is the predictive power of the yield curve. Evidence from Germany and the United States. *Federal Reserve Bank of New York Staff Report*, *113*.

European Commission. (2019). *Commission Recommendation (EU) 2019/786 of 8 May 2019 on building renovation.*

European Parliament and Council. (2019). Regulation (EU) 2019/2088 of the European Parliament and of the council of 27 november 2019 on sustainability-related disclosures in the financial services sector. *Official Journal of the European Union*.

European Public Real Estate Association. (2023). Total markets table Q1-2023. *Published*

Online. Retrieved from https://prodapp.epra.com/media/EPRA_Total_Markets_Table_-_Q1-2023_1681460716641.pdf

Farrelly, K., & Matysiak, G. (2012). Performance drivers of UK unlisted property funds. *Working Paper*.

Federal Reserve Bank of St. Louis. (2023). Federal Reserve Economic Data. *FRED Database Online*.

Feng, Z., & Wu, Z. (2021). ESG disclosure, REIT debt financing and firm value. *The Journal of Real Estate Finance and Economics*, 1–35.

Ferriani, F. (2023). The importance of labels for sustainable investments: SFDR versus Morningstar globes. *Applied Economics Letters*, 1–7.

Fuerst, F. (2015). The financial rewards of sustainability: A global performance study of real estate investment trusts. *Available at SSRN 2619434*.

Fuerst, F., & Matysiak, G. (2013). Analysing the performance of nonlisted real estate funds: a panel data analysis. *Applied Economics*, *45*(14), 1777-1788.

Fuerst, F., & McAllister, P. (2009). An investigation of the effect of eco-labeling on office occupancy rates. *Journal of Sustainable Real Estate*, *1*(1), 49–64.

Fuerst, F., & McAllister, P. (2011). Green noise or green value? Measuring the effects of environmental certification on office values. *Real Estate Economics*, *39*(1), 45–69.

Hartman, J. C., & Schafrick, I. C. (2004). The relevant internal rate of return. *The Engineering Economist*, *49*(2), 139–158.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association*, *82*(400), 1147–1149.

Holtermans, R., & Kok, N. (2019). On the value of environmental certification in the commercial real estate market. *Real Estate Economics*, *47*(3), 685–722.

Holton, G. A. (2013). *Value-at-risk: Theory and practice* (2nd ed.). Published Online. Retrieved from https://www.value-at-risk.net

Hsieh, H.-C., Claresta, V., & Bui, T. M. N. (2020). Green building, cost of equity capital and corporate governance: Evidence from US real estate investment trusts. *Sustainability*, *12*(9), 3680.

INREV. (2023). INREV response to EC SFDR Targeted-consultation document. *INREV Publication*. Retrieved from https://www.inrev.org/system/files/2023-12/INREV%20response%20to%20EC%20SFDR%20Targeted-consultation%

20document_6%20December.pdf

INREV. (2023). SFDR's Impact on Real Estate Investment. *Published Online*. Retrieved from https://www.inrev.org/system/files/2023-03/SFDR%27s-impact-on-real-estate-investment.pdf

Kahn, M. E., & Kok, N. (2014). The capitalization of green labels in the california housing market. *Regional Science and Urban Economics*, *47*, 25-34.

Lambourne, T. (2022). Valuing sustainability in real estate: a case study of the United Arab Emirates. *Journal of Property Investment & Finance*, *40*(4), 335–361.

Leutner, S., Gloria, B., & Bienert, S. (2024). Is there a green discount in commercial real estate lending? *Journal of Property Investment & Finance*.

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and statistics*, *61*(S1), 631–652.

Mariani, M., Amoruso, P., Caragnano, A., & Zito, M. (2018). Green real estate: does it create value? Financial and sustainability analysis on european green REITs. *International Journal of Business and Management*, *13*, 80–92.

Morri, G., Anconetani, R., & Benfari, L. (2021). Greenness and financial performance of european REITs. *Journal of European Real Estate Research*, *14*(1), 40–61.

Ooi, J. T., & Dung, D. D. (2019). Finding superior returns in green portfolios: Evidence from Singapore REITs. *Journal of Sustainable Real Estate*, *11*(1), 191–215.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265–312.

PricewaterhouseCoopers. (2023). The ESG challenge for real estate. *Published Online*. Retrieved from https://www.pwc.lu/en/sustainable-finance/a-year-of-esg/the-esg-challenge.html

Reichardt, A., Fuerst, F., Rottke, N., & Zietz, J. (2012). Sustainable building certification and the rent premium: a panel data approach. *Journal of Real Estate Research*, *34*(1), 99–126.

Rousseeuw, P. J., & Leroy, A. M. (2005). *Robust regression and outlier detection*. John Wiley & Sons.

Sah, V., Miller, N., & Ghosh, B. (2013). Are green REITs valued more? *Journal of Real Estate Portfolio Management*, *19*(2), 169–177.

Schütze, F., & Stede, J. (2021). The EU sustainable finance taxonomy and its contribution to climate neutrality. *Journal of Sustainable Finance & Investment*, 1–33.

Stevenson, S. (2006). Driving forces in the performance attributes of real estate funds. *INREV Conference, Rome*.

Stoer, J., & Bulirsch, R. (2002). *Introduction to numerical analysis* (3rd ed.). Springer New York, NY.

Wiley, J. A., Benefield, J. D., & Johnson, K. H. (2010). Green design and the market for commercial office space. *The Journal of Real Estate Finance and Economics*, *41*, 228–243.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT press.

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, *11*(10), 1–17.

6 Is there a Green Discount in Commercial Real Estate Lending?

6.1 Abstract

Purpose: This study examines whether green buildings enjoy more favorable financing terms compared to their non-green counterparts, exploring the presence of a green discount in commercial real estate lending. Despite the extensive research on green premiums on the equity side, lending has received limited attention in the existing literature, even as regulations have increased and ambitious net-zero targets have been set in the banking sector.
Design/methodology/approach: In this study, we leverage a unique dataset comprising European commercial loan data spanning from 2018 to 2023, with a total loan value exceeding €30 billion. We employ hedonic regression models to analyze the data. Specifically, we rely on property assessments conducted by lenders to investigate whether green properties exhibit lower interest rate spreads and higher Loan to Value (LTV) ratios.
Findings: Our findings reveal the existence of a green discount in European commercial real estate lending, with green buildings enjoying a 5.35% lower contracted loan spread and a 3.92% lower target spread compared to their non-green counterparts. However, our analysis does not indicate any distinct advantage in terms of LTV ratios for green buildings.
Originality/value: This study, to the best of our knowledge, represents the first of its kind in a European context and provides empirical evidence for the presence of a green discount.
Practical implications: Our research contributes to a deeper understanding of the interaction between green properties and commercial real estate lending, offering valuable insights for both lenders and investors.

Keywords: *Commercial Real Estate Lending; Sustainable Finance, Green Discount.*

6.2 Introduction and Background Literature

Building on the UN Sustainable Development Goals (SDGs) and the Paris Agreement, sustainable finance and investment have emerged as a transformative approach that aims to integrate ESG considerations into financial decision-making processes. This paradigm shift has taken hold in various industries, and the property sector is no exception. Most market participants have become aware of the real estate industry's major responsibility in the transition to a carbon-neutral economy and to the importance of realigning capital flows accordingly. In this context, it is evident that sustainable property investments are not solely driven by philanthropic motives. Extensive research conducted over the last decade has revealed that sustainable properties are linked to a green premium on the equity side. In literature, green and non-green properties are predominantly differentiated based on either common certification schemes such as LEED, BREEAM, and ENERGY STAR or national energy performance certificates. The primary focus of studies within the commercial real estate sector revolves around office buildings, revealing a connection between property certification and higher rents, prices, as well as enhanced operational performance (Wiley et al., 2010; Eichholtz et al., 2010; Fuerst & McAllister, 2011a; Reichardt et al., 2012; Eichholtz et al., 2013; Chegut et al., 2014; Holtermans & Kok, 2019). Similar green premiums have been observed in real estate markets across the United States, Asia, and Europe for both single- and multi-family residential properties (Deng et al., 2012; Kahn & Kok, 2014; Fuerst et al., 2015; Bond & Devine, 2016; Cajias et al., 2019; Cadena & Thomson, 2021; Dell'Anna & Bottero, 2021). In addition to studies that focus on the property level, another body of literature highlights the positive impact of holding green properties on the corporate level of REITs (Eichholtz et al., 2012; Sah et al., 2013; Fuerst, 2015; Ooi & Dung, 2019; Devine & Yönder, 2021). Despite the evident long-term demand-side benefits leading to a green premium for property owners, the adoption of new green construction and retrofitting has not surged as one would expect. Chegut et al. (2019) attribute this inefficiency in the market to the lack of short-term incentives for the supply side and further point out that longer construction times and higher upfront costs for green construction are an obstacle for equity-constrained developers. Lenders in particular have the capability to remove potential barriers to entry for real estate developers by offering lower rates and LTV ratios, effectively addressing potential short-term capital shortages (Chegut et al., 2019; Mathew et al., 2021).

The necessity of debt capital in financing a carbon-neutral economy is becoming increasingly apparent to banks. For example, ambitions at the global level are reflected in the United Nations Net Zero Banking Alliance (NZBA), which assembles 126 banks from 41 different countries with a total of over 40% of global banking assets (UNEP

FI, 2021b). To this end, NZBA has released the Guidelines for Climate Change Target Setting (UNEP FI, 2021a). Furthermore, the UN Principles for Responsible Banking serve as a voluntary framework for banks to align their strategies with the UN Sustainable Development Goals (UNEP FI, 2019). Other initiatives with global reach, such as Partnership for Carbon Accounting Financials (PCAF), Network for Greening the Financial System (NGFS), and Task Force on Climate-related Financial Disclosures (TCFD), complement legally binding (EU-)frameworks such as the EU taxonomy or the Corporate Sustainability Reporting Directive (CSRD). As a component of the EU Taxonomy, a novel financial metric emerges: the Green Asset Ratio (GAR), slated for mandatory reporting by banks starting in 2024. Although there are currently no legally binding benchmarks for achieving a specific GAR, financial institutions could increase their efforts to promote green lending and actively participate in the discernible shift toward quantifying green assets in their portfolios. The rising global requirements for (sustainability) reporting can continue to result in banks raising their assessment criteria for the properties eligible for financing. In the event that banks are committed to Net-Zero goals—whether voluntarily or through regulatory requirements—the portfolio of properties financed may be influenced by this commitment. Consequently, an increase in the GAR as well as potential incentivization is anticipated for numerous lenders.

Although some residential lenders already provide incentives for green properties, such as additional loan proceeds and lower spreads, a green discount is not yet openly promoted in commercial real estate lending (Mathew et al., 2021). Despite usually not being explicitly advertised during loan origination, the question arises as to whether a green discount for commercial property loans is still observable, particularly when considering green real estate from a risk perspective and the imperative for lenders to enhance the proportion of green real estate in their new business. To this date, only a limited number of studies in the realm of sustainable real estate focus on the debt side of properties. Devine and McCollum (2021) analyzed mortgage-backed securities data from Fannie Mae and discovered that loans for green multifamily properties, backed by green bonds, exhibit lower interest rates and lower debt service coverage ratios. Their findings reveal that loan packages committed to reducing carbon emissions are linked to higher loan-to-value ratios and greater debt levels per building unit, even after accounting for loan and property characteristics. Eichholtz et al. (2019) found that commercial mortgages on buildings owned by REITs and certified with LEED or ENERGY STAR experience notably lower spreads. Attributed to reduced bond spreads, REITs holding a higher proportion of certified buildings also benefit from more advantageous corporate-level financing in the secondary market. For REITs investing in office and retail properties, Devine and Yönder (2021) also observed that, regardless of the differences in quality characteristics, lenders attribute lower risk to pure participation in green certification schemes, which leads to lower interest costs at the corporate level. Based on CMBS loan data secured by individual properties from 2000 to 2013, research by An and Pivo (2020) revealed that LEED and ENERGY STAR certified buildings exhibit notably lower default risk. This reduction can be attributed not only to the certification itself but also to the actual level of green achievement. Although the decrease in default risk economically outweighs the impact on spreads at loan origination, their study indeed highlights lower spreads for green properties during the origination process. Upon examination of sub-periods, they observed that the green discount in spreads began to emerge gradually from 2006 onwards, and became even more pronounced following the financial crisis.

We aim to extend existing research by analyzing the influence of a property's categorization as either green or non-green on spreads and LTV ratios in commercial real estate lending. This allows us to determine the potential of lenders to both mitigate entry barriers for the supply side of green properties and to enhance the own portfolio's green asset ratio. Within our analysis, a distinction is established between the target spread, denoting the minimum spread derived from lender's internal pre-calculations, and the spread that is contractually agreed. A differentiated view further enables us to identify a potential pricing effect on green properties between internal calculation and the spreads effectively conveyed during the loan origination process. Given that prior studies predominantly rely on data of listed real estate, focus on owner-occupied residential lending exclusively, or examine periods when lenders were only beginning to distinguish between green and non-green properties, the empirical evidence for benefits on the debt side of green properties is very limited. To the best of our knowledge, our study is the first to examine the green discount at the loan-level within a European context, using the most recent data available. Moreover, our focus diverges from external commercial certification schemes, which are based on voluntary participation, to the actual energy intensity of buildings. This aspect is increasingly garnering attention from lenders due to mounting regulatory pressures. Our empirical analysis is based on a unique dataset consisting of 728 commercial real estate loans across a range of property types, spanning from 2018 through 2023. In the following part of this paper, besides providing a more comprehensive outline of the dataset and descriptive statistics, we will describe the loan criteria applied to categorize properties as either green or non-green. Subsequently, we will delineate our methodological approach and present the results along with their implications. The paper ends with our final conclusions.

6.3 Data and Descriptive Statistics

The cross-sectional dataset underlying our analysis was provided by one of Germany's largest commercial real estate lenders. It encompasses European new commercial transactions consisting of 728 loans, amounting to a total volume of over €30 billion, spanning from January 2018 to June 2023. The frequency of loan approvals remains constant over time, with the highest number of transactions occurring in 2019, accounting for 21% of all observations in our sample. The loan data is solely related to investment properties. Furthermore, in the context of residential real estate, we specifically focused on acquiring loans designated for investment purposes. This focus is essential, as it significantly influences risk assessment and the corresponding terms of the loan, depending on whether the property is owner-occupied or intended for income generation. We employ the loan spreads and LTV ratios in our sample to identify the presence of a potential green discount.

In financial market theory, the nominal interest rate is composed of both time and risk components for which a lender must be compensated in order to lend money. Fisher (1930) describes the nominal interest rate as the sum of real interest rate and expected inflation. In addition to the real rate and expected inflation, the liquidity premium which incorporates the loan maturity is considered as a further time component. The risk component of the nominal rate consists of the specific risks and the systematic risks. Typically, specific risks are related to loan and property characteristics, such as customer-specific risk classification (including investor profile), LTV, loan volume, location, and property type. The relevant descriptive statistics can be found in Table 6.1. In our analysis, it is crucial to note that the actual component of time risk is explicitly and independently quantified as a separate funding spread by the lender. To mitigate heterogeneity concerns across different loans in relation to time risks and fluctuations in the real rate, we explicitly requested an adjustment for both the funding spread and the base rate. Therefore, these spreads have been adjusted for all maturity-based risk factors, such as the loan maturity, the type of loan (whether amortization or bullet loan), and the nature of the base rate (fixed or variable). The spreads observed in our dataset predominantly reflect the specific and systematic risk components, internal costs, and loan origination fees.

Table 6.1: Summary Statistics of the Sample

Variable	N	Mean	Std. dev.	Min	Q25	Q75	Max
Spread	728	1.30	0.55	0.39	0.94	1.70	4.30
Target spread	728	1.30	0.49	0	0.93	1.60	4.40
LTV (in %)	728	61	17	11	50	70	163
Volume (in t. €)	728	40,770	46,775	0.01	11,000	54,000	550,000
Maturity (in months)	728	73	45	3	36	120	360
Green building	728	33%					
Investor profile	728						
Developer	150	21%					
Fund	35	5%					
Foreign investor	213	29%					
Domestic investor	202	28%					
Investment management	76	10%					
Housing company	52	7%					
Customer rating	728						
Customer rating 1	212	29%					
Customer rating 2	178	24%					
Customer rating 3	229	31%					
Customer rating 4	100	14%					
Customer rating 5	9	1%					
Property type	728						
Office	305	42%					
Retail	85	12%					
Specialty retail center	10	1%					
Hotel	5	1%					
Department store	5	1%					
Logistics	80	11%					
Management property	15	2%					
Parking garage	4	1%					
Mall	3	0%					
Residential investment	205	28%					
Social property	2	0%					
Other	9	1%					
Region	728						
DE: old states A-cities	185	25%					
DE: old states B-cities	90	12%					
DE: old States (other)	108	15%					
DE: new States A-cities	25	3%					
DE: new States (other)	23	3%					
DE: Berlin	115	16%					

Continued on next page

Table 6.1: Summary Statistics of the Sample (Continued)

Belgium	4	1%
France	38	5%
Great Britain	1	0%
Luxembourg	1	0%
Netherlands	99	14%
Poland	38	5%
Other countries	1	0%
Year	728	
2018	124	17%
2019	156	21%
2020	142	20%
2021	135	19%
2022	138	19%
2023	33	5%

As is evident from Figure 6.1, the spread data does not follow a normal distribution. This observation is confirmed by the results of the Jarque and Bera (1980) test, with the null hypothesis of normally distributed data (p-value = 0, Kurtosis: 5.53, Skewness: 1.11). By applying a logarithmic transformation to the data, as depicted in Figure 6.2, we attain a normal distribution, which is a fundamental assumption of the models employed in our analysis. This observation is supported by the Jarque-Bera test (p-value > 0.18, Kurtosis: 2.82, Skewness: -0.14). Similar observations are made for the target spread data.

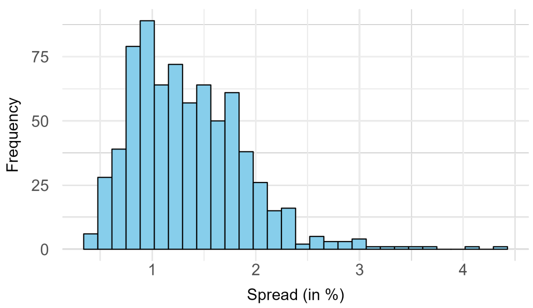


Figure 6.1: Spread Histogram

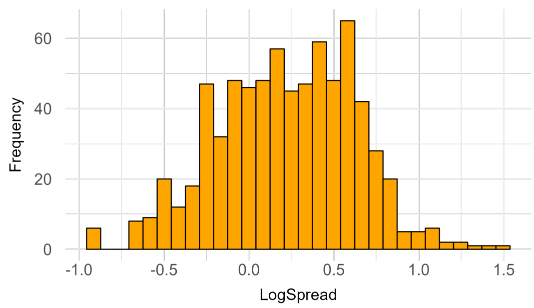


Figure 6.2: LogSpread Histogram

The lender's customer rating system comprises risk classes numbered from 1 (representing the lowest risk) to 5 (indicating the highest risk). Within this framework, a notable segment (31%) of customers is categorized as class 3, while the second-largest class is class 1, comprising 29%. Class 5 (representing the highest level of risk) constitutes less than 2% of the loans. Figure 6.3 illustrates that a greater risk appears to necessitate a higher spread. Figure 6.4 displays the allocation of loan frequency based on the respective investor profiles. It indicates that a significant portion of loans was extended to both foreign and domestic investors, comprising 57% of the total. Notably, a slightly greater proportion of these loans was allocated to foreign investors. The smallest portion pertains to funds (5%). On average, the bank establishes a target spread of 1.30 percentage points above the given interest rate for these transactions. Correspondingly, the actual spread realized in the loan contracts aligns with this target, averaging at the same level. In our sample, properties are financed at an LTV of 61%, accompanied by an estimated standard deviation of 16.57. The average loan size is just over €40 million, which could correspond to a single asset or multiple similar properties financed together. The average term to maturity in our sample is 6 years. While foreign investors contribute a significant absolute volume of loans, approximately three quarters of the total funded amount remains within Germany. The Netherlands rank second (13%), followed by France (7%) and Poland (5%). Office and residential properties make up a clear focus in the portfolio, accounting for a total of 70% of loans. The other property types, except for retail (12%) and logistics (11%), are only sporadically represented in our sample.

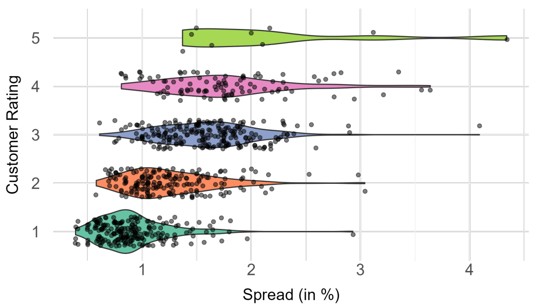


Figure 6.3: Violin Plot of Customer Ratings vs. Spread in %

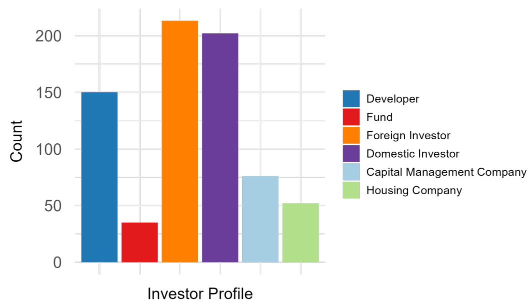


Figure 6.4: Bar Plot of Loan Counts by Investor Profiles

In the analysis of green properties, existing energy performance certificates are frequently used as a benchmark for ecological performance (Cajias et al., 2019; Fuerst et al., 2015; Fuerst & McAllister, 2011b). Energy certificates allow for a standardized and clear-cut benchmarking process. However, in light of the prevailing regulatory demands and the intensifying pressures of the market, this relatively straightforward classification is progressively proving insufficient. Within the European context relevant to our dataset, the EU taxonomy regulation stipulates that existing buildings can be classified as sustainable if they belong to the top 15% in terms of energy efficiency

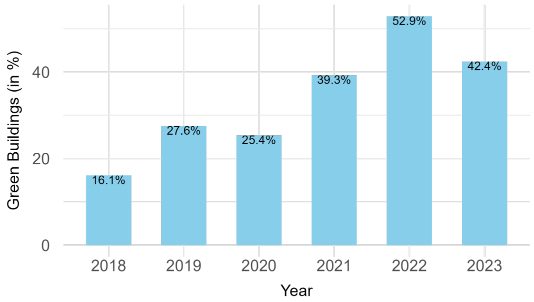


Figure 6.5: Green Asset Ratio in % of Loans per Year

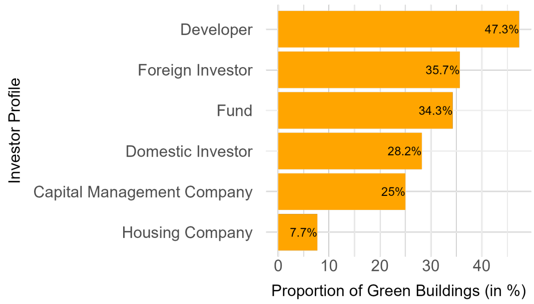


Figure 6.6: Green Asset Ratio in % of Total Loans per Investor Profile

within the national or regional building stock. Therefore, to derive a binary variable that categorizes properties as non-green or green, a comprehensive examination of current regulatory guidelines, available data, and relevant key performance indicators is necessary. In the sample of buildings we analyzed, the green building variable classifies properties as green if they either (1) exhibit exceptional energy efficiency or (2) adhere to the technical criteria set forth in the EU taxonomy regulation for buildings. The first set of criteria was introduced by the financial institution prior to the publication of the EU taxonomy criteria and was subsequently complemented by the second set to align with the EU taxonomy requirements. Generally, a building qualifies as a green building if it meets at least energy performance certificate class A, does not exceed specific energy demand thresholds per square meter, holds a high level of sustainability certifications, and/ or complies with the EU Taxonomy criteria for existing buildings, new construction, or renovations. Hence, the principal criterion for a green loan in our dataset is the aggregate best-in-class energy requirement for heating and electricity, aimed at precluding buildings with a suboptimal energy performance in terms of building envelope or those with a disproportionately high energy demand from qualifying for a green loan. Moreover, older buildings are eligible for a green loan either by aligning with the best-in-class of the top 15 percent of the national or regional building standards, or through energy modernization and a reduction of primary energy demand by 30%. The focus on energy intensities is in line with the 'efficiency first principle' as outlined in the European Parliament (2023). This approach is pivotal, as energy savings are not only a feasible method for reducing carbon emissions but also, even under the most optimistic scenarios, there is an anticipated limitation in the availability of renewable energy for building operations by 2050, as projected by the IEA (International Energy Agency, 2021).

Within our dataset, the proportion of green buildings is approximately 33%, with the share showing a tendency to increase over time (see Figure 6.5). As shown in Figure 6.6, developers constitute the largest share at 47%, while housing companies represent the smallest share at 8% in terms of green buildings within our sample. This observation may be attributed to the distinction between new construction and existing stock. Developers typically engage in new construction or extensive revitalization, whereas housing companies often acquire substantial portions of their portfolios from existing stock. With regard to real estate values, which can be estimated by extrapolating average LTVs and loan volumes, our findings align with those of Eichholtz et al. (2019) in that green properties, on average, command a higher value of approximately €75 million compared to non-green properties at €63 million.

6.4 Empirical Framework

In order to identify and isolate the potential impact of property categorization as green or non-green on loan spreads and LTV, while also accounting for other explanatory loan characteristics, we employ a hedonic regression model based on the framework first introduced by Rosen (1974). According to Fuerst and McAllister (2011a) this methodology is used as the primary framework for analyzing the factors that determine pricing effects in real estate research and is therefore commonly applied in empirical studies examining the green premium in property markets (Eichholtz et al., 2010; Fuerst & McAllister, 2011a; Bond & Devine, 2016; Holtermans & Kok, 2019). Within the ambit of this methodology, individual characteristics of each loan, encompassing elements like property type, region, or year, are assigned specific hedonic weights. These weights represent the distinct value contributions of each characteristic to the overall loan spread. Employing this empirical framework facilitates a targeted isolation and precise quantification of the value contribution ascribed to the classification of a property as a green property.

$$\log S_i = \alpha + \beta_i X_i + \delta g_i + \sum_{j=1}^{J} \gamma_j c_{ij} + \sum_{t=1}^{T} \lambda_t h_{it} + \epsilon_i \tag{6.1}$$

$$\log S_i = \alpha + \beta_i X_i + \delta g_i + \sum_{j=1}^{J} \gamma_j c_{ij} + \sum_{t=1}^{T} \lambda_t h_{it} + \sum_{t=1}^{T} \theta_t h_{it} g_i + \epsilon_i \tag{6.2}$$

We define the functional form of the hedonic regression function as semi-log with the (target) spread as the logarithmized dependent variable S_i. According to Sirmans et al. (2005), the semi-log model offers two decisive advantages over the linear form. On one hand, the value of a specific characteristic can vary proportionally with that of another characteristic. On the other hand, logarithmization can help mitigate heteroscedasticity, thus reducing changes in the variance of the residuals. Moreover, if the dependent variable is skewed, applying logarithmization can help to obtain a distribution which follows a normal distribution more closely (see Figure 6.1 and 6.2). X_i denotes a vector of the hedonic loan characteristics described in the initial section of the paper, including factors such as property type, initial loan size, or maturity. g_i defines the dummy variable that is crucial for our research question. It takes a value of 1 if the lender categorizes the property as green, and zero otherwise. As highlighted by An and Pivo (2020), addressing endogeneity bias, which often stems from the unobserved heterogeneity among different properties, is crucial in research focused on accurately isolating the green premium. Ignoring vital influencing factors in the model can lead to spurious regression results, thereby distorting interpretations and compromising the validity of conclusions. To counteract this issue effectively, our models integrate both location-specific (c_{ij}) and time-specific (h_{it}) fixed effects. This approach ensures a nuanced capture of the distinct characteristics and temporal dynamics inherent to each cluster, thus fostering a more accurate and reliable analytical outcome. β_i, δ, γ_j, and λ_t are estimated coefficients of the explanatory variables. ϵ_i is an error term. The inclusion of an additional interaction term between the two dummy variables h_{it} and g_i in Equation 2 allows us to detect a potential change in the impact of property categorization on spreads over time, reflecting the pattern of green discount found by An and Pivo (2020).

$$L_i = \alpha + \beta_i X_i + \delta g_i + \sum_{j=1}^{J} \gamma_j c_{ij} + \sum_{t=1}^{T} \lambda_t h_{it} + \sum_{t=1}^{T} \theta_t h_{it} g_i + \epsilon_i \tag{6.3}$$

The same methodological framework is applied in Equation 3 to determine a potential influence of the building categorization on LTVs, resulting in additional loan proceeds for green properties. As we did not identify any benefits from applying a logarithmic transformation to the dependent variable L_i, we opted for a linear form for these models. Although the semi-log model contributes to mitigating heteroscedasticity, the Breusch and Pagan (1979) test is applied to detect significant changes in the variance of the residuals across all models. Persisting heteroscedasticity is corrected for by using a heteroskedasticity-consistent covariance matrix estimator (White, 1980). Another issue in hedonic regression analysis that may cause the regression results to be biased is multicollinearity, a situation where two or more independent variables are highly correlated. Beyond examining the correlations among individual predictors, the Variance Inflation Factor (VIF) and the Generalized Variance Inflation Factor (GVIF) for categorical variables enable us to identify which specific variables are affected by multicollinearity. When comparing GVIFs, Fox and Monette (1992) recommend using the modification GVIF$^{(1/(2*Df))}$ with Df degress of freedom, as this ensures comparability across dimension. A value of 10 or higher is considered a common rule of thumb for a high degree of multicollinearity, suggesting a need to correct the independent variable accordingly. As outlined by O'brien (2007), we consider additional factors such as t-values and confidence intervals that affect the stability of the estimates, so that reliable statistical inferences might be drawn even with elevated VIF or GVIF values. Our analysis encompasses a total of twelve models. There are four models each for target spread, spread, and LTV as dependent variables. In each case, Model 1 provides the baseline model, while Models 2 and 3 introduce time and regional effects. Model 4 expands Model 3 by including an interaction term between the green building dummy and the time dummy.

6.5 Empirical Results and Implications

Our analysis comprises a total of 12 hedonic models evenly distributed across the three predefined dependent variables ($LogSpread$, $LogTargetSpread$, and LTV). As described, the respective Model 1 serves as the base model, which is augmented with regional and time effects in Models 2 and 3, respectively. In terms of our research question, the independent green building dummy is central to our analysis. In this context, a value of 0 denotes non-green buildings, while a value of 1 signifies green buildings. Model 4 further incorporates an interaction term between the year and green building variables to account for potential time effects with regard to green building. In addition to continuous variables such as loan volume, term to maturity, and (target) spread or LTV, we address the heterogeneity in loan data by incorporating investor profile, customer rating (ranging from 1 for low risk to 5 for high risk), and property type.

The metrics adjusted R^2, AIC, BIC, and RMSE facilitate the identification of the models that have the best goodness of fit for each corresponding dependent variable. In Table 6.2, regarding the dependent variable $LogSpread$, Model 3 emerges as the preferred choice. Consequently, coefficients are interpreted with reference to Model 3. Notably, the adjusted R^2 is close to 0.7, and we observe the lowest values for AIC and BIC. In addition, there is consistency in the RMSE, which is 0.22 for both Model 3 and Model 4. An examination of the confidence intervals as well as the VIF and GVIF values shown in Table 6.5 in the Appendix does not reveal any unexpected multicollinearity. We observe analogous results in the models with $LogTargetSpread$ and LTV as the dependent variables.

Table 6.2: Hedonic Model Results where $y = LogSpread$

Model	1	2	3	4
(Intercept)	0.210*** (0.059)	0.206*** (0.059)	0.091 (0.061)	0.086 (0.062)
Green building	0.006 (0.021)	-0.010 (0.021)	-0.055** (0.020)	-0.032 (0.058)
LTV (in %)	0.002*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)
Volume (in t. €)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Maturity (in months)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Investor profile (ref: developer)				
Fund	-0.197*** (0.053)	-0.223*** (0.053)	-0.242*** (0.050)	-0.243*** (0.051)
Foreign investor	-0.131*** (0.031)	-0.252*** (0.037)	-0.276*** (0.035)	-0.274*** (0.035)
Domestic investor	-0.140*** (0.031)	-0.151*** (0.030)	-0.164*** (0.029)	-0.163*** (0.029)
Investment management	-0.360*** (0.043)	-0.380*** (0.043)	-0.384*** (0.041)	-0.380*** (0.041)
Housing company	-0.389*** (0.050)	-0.438*** (0.051)	-0.469*** (0.049)	-0.467*** (0.049)
Customer rating (ref: 1)				
Customer rating 2	0.163*** (0.028)	0.156*** (0.027)	0.171*** (0.026)	0.172*** (0.026)
Customer rating 3	0.334*** (0.028)	0.307*** (0.028)	0.302*** (0.027)	0.300*** (0.027)
Customer rating 4	0.397*** (0.034)	0.368*** (0.034)	0.358*** (0.033)	0.358*** (0.033)
Customer rating 5	0.526*** (0.086)	0.528*** (0.085)	0.476*** (0.081)	0.471*** (0.081)
Property type (ref: office)				
Retail	-0.064+ (0.033)	-0.058+ (0.034)	-0.053+ (0.032)	-0.051 (0.032)
Specialty retail center	0.126 (0.079)	0.096 (0.079)	0.002 (0.076)	-0.003 (0.077)
Hotel	0.271* (0.111)	0.270* (0.108)	0.218* (0.104)	0.222* (0.104)
Department store	0.146 (0.111)	0.198+ (0.110)	0.114 (0.106)	0.106 (0.106)
Logistics	-0.042 (0.031)	-0.032 (0.033)	-0.044 (0.031)	-0.043 (0.032)
Management property	-0.057 (0.065)	-0.046 (0.064)	-0.012 (0.061)	-0.006 (0.061)
Parking garage	0.408** (0.125)	0.278* (0.125)	0.144 (0.121)	0.136 (0.122)
Mall	0.262+ (0.143)	0.252+ (0.140)	0.186 (0.134)	0.184 (0.135)
Residential investment	-0.131*** (0.025)	-0.127*** (0.024)	-0.153*** (0.024)	-0.154*** (0.024)
Social property	0.187 (0.176)	0.204 (0.171)	0.176 (0.164)	0.187 (0.164)
Other	0.010 (0.083)	0.022 (0.081)	0.029 (0.078)	0.024 (0.078)
Region (ref: DE: old states A-cities)				
DE: old states B-cities		0.001 (0.033)	-0.005 (0.031)	-0.005 (0.032)
DE: old states (other)		0.005 (0.032)	0.020 (0.031)	0.017 (0.031)
DE: new states B-cities		0.114* (0.053)	0.120* (0.050)	0.120* (0.050)
DE: new states (other)		0.043 (0.056)	0.024 (0.054)	0.021 (0.054)
DE: Berlin		0.021 (0.029)	0.027 (0.028)	0.027 (0.028)
Belgium		-0.052 (0.122)	-0.020 (0.117)	-0.027 (0.119)
France		0.021 (0.050)	0.041 (0.048)	0.043 (0.048)
Great Britain		0.147 (0.240)	0.056 (0.229)	0.044 (0.230)
Luxembourg		0.111 (0.240)	0.215 (0.229)	0.218 (0.232)
Netherlands		0.181*** (0.039)	0.194*** (0.037)	0.189*** (0.037)

Continued on next page

Table 6.2: Hedonic Model Results where $y = LogSpread$ (Continued)

Poland		0.303*** (0.051)	0.312*** (0.049)	0.311*** (0.049)
Other countries		-0.051 (0.240)	-0.086 (0.229)	-0.098 (0.230)
Year (ref: 2018)				
Year 2019			0.023 (0.028)	0.027 (0.032)
Year 2020			0.093** (0.029)	0.083* (0.033)
Year 2021			0.131*** (0.030)	0.146*** (0.036)
Year 2022			0.203*** (0.030)	0.217*** (0.038)
Year 2023			0.273*** (0.046)	0.292*** (0.059)
Green building x 2019				-0.025 (0.070)
Green building x 2020				0.031 (0.072)
Green building x 2021				-0.049 (0.070)
Green building x 2022				-0.040 (0.069)
Green building x 2023				-0.056 (0.100)
Num.obs.	728	728	728	728
R^2	0.650	0.676	0.708	0.709
R^2 adj.	0.638	0.659	0.691	0.690
AIC	35.6	3.6	-63.4	-55.9
BIC	155.0	178.1	134.0	164.4
RMSE	0.24	0.23	0.22	0.22
Regional effects	N	Y	Y	Y
Time effects	N	N	Y	Y

Note: The standard errors are indicated in parentheses. The levels of statistical significance are represented as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Model 3 and 4 are estimated using heteroskedasticity-robust covariance matrix estimators. The variable green building is a dummy variable that is one if the respective asset is classified as a green building by the lender, according to the classification described in the text. The categorical variables for investor profile describe the classification of the type of investor made by the lender and is independent of the financed asset. The variable customer rating describes the risk classification of the customer made by the lender, where 1 = least risky and 5 = most risky, and is also independent of the asset. The year refers to the year of financing issuance. The interaction term between year and green building is one if a financed asset was classified as green in the respective years.

The combined influence of these variables provides insights into how different loan and borrower characteristics affect the lender's spread requirements. It becomes apparent that loan-related factors, specifically LTV, loan volume (which appears to have no significant impact), and term to maturity, exert only minimal influence on the spread. The limited impact of maturity on spreads aligns with our expectations. As outlined in the descriptive statistics, this can be attributed to the fact that maturity-related effects are predominantly accounted for by the lender's distinct funding spread, resulting in only a marginal effect on the dependent (target) spread variables. The categorical variable investor profile, in conjunction with customer rating, represents the customer profile. With developer serving as the reference category, negative coefficients of the remaining variables indicate that developers pay the highest spreads. In addition to the inherently higher risk for developers compared to investors pursuing buy and hold strategies, the recourse options for lenders in these cases are often limited due to special purpose entities established for developments. For joint ventures or general and limited partner situations, underwriting for lenders, especially with recourse options, is not straightforward and often entails higher complexity. That is because other assets that are used as collateral need also to be assessed thoroughly. Although there are notable differences in underwriting requirements and pricing between an investment loan and a development loan, we can effectively control for these variations. Firstly, the maturity-specific risks, as previously mentioned, are not included in the spreads under examination. Secondly, as shown in Tables 6.2 through 6.4, the highly significant hedonic weights of all investor types indicate that our statistical models are capable of isolating the influences of various loan types. The lowest spread is paid by housing companies (-0.469 or -37.44%). Coefficients in Model 3 for customer rating align with intuitive expectations - higher risk categories correspond to higher spread levels. The values of the coefficients increase from the reference category Customer Rating 1 to the highest risk category Customer Rating 5, amounting to 0.476, equivalent to an increase in spreads of nearly 71%.

Property type and location-specific parameters (regional effects), along with the green building variable, collectively account for the building-specific characteristics within our sample. Unlike the investor profile variables, only a small subset of the building-specific characteristics significantly affects $LogSpread$. Among property types (with office as the reference category), three out of eleven coefficients lie within the defined significance level of 0.1. Retail properties experience a decrease of -0.053 or -5.16%, whereas hotels demonstrate an increase of 0.218 or 24.38% compared to offices. Similar to the reduction observed for housing companies, there is a decrease evident for residential properties, which might be attributed to the availability of promotional loans for residential buildings. As indicated by a strongly significant coefficient of -0.153, the lender realizes 14.19% less spread. The location-specific parameters encompass 13 locations (country, German federal state, or city level), with Germany: old states A-cities serving as the reference category. However, only three of these parameters exhibit a significant influence on the dependent variable. The most pronounced increase in spreads is observed for Poland, amounting to 0.312 or 36.62% compared to Germany: old states A-cities. This observation is consistent with Poland's higher property yield environment compared to other European countries covered in our analysis. The aforementioned adjustments for customer and building characteristics mirror the bank's individual assessment of risk-return profiles.

The time effects, using the reference year 2018 to account for year-specific characteristics, are statistically significant for all years except for 2019. It can be observed that the annual effects on the *LogSpread* steadily increase over time. The explanatory power can be further enhanced by considering the time dimension when comparing the goodness of fit between Model 2 and Model 3, two models that differ only in the inclusion of time effects. The introduction of the interaction term between the year and the green building variable in Model 4 does not improve the model's performance any further and also lacks statistical significance. Hence, contrary to the findings of An and Pivo (2020), we cannot observe a change in the impact of green building classification on credit spreads over the years. Moreover, the inclusion of the interaction term leads to an approximate tripling of the modified GVIF factor for the green building variable, indicating a substantial increase in multicollinearity.

In Model 3, the green building variable exhibits a highly significant coefficient of -0.055, resulting in a reduction in spreads of approximately 5.35%. Given the average loan size in our sample of €40.77 million and an average spread of 130 bps, this translates into a green discount of approximately 7 bps. This green discount corresponds to an interest saving of approximately €28,356 for borrowers in the first year. In contrast to the coefficients of other independent variables in the model, the green discount may seem relatively modest. However, in the context of our research question, the finding of a green discount in European commercial lending is novel. The reduction in spreads we observe is somewhat less pronounced when compared to the findings of Eichholtz et al. (2019). Their research shows a green discount ranging from 24 to 29 bps, depending on the specifications of the green-certified property, with a relative spread reduction of 7.95 to 9.60%. In contrast to our dataset and the methodology employed by An and Pivo (2020), Eichholtz et al. (2019) derive the spread by utilizing mortgage rates sourced from U.S. REIT data and subtracting the corresponding treasury rate based on mortgage rate maturity. Our results are more in line with those of An and Pivo (2020), indicating a green discount of 5.85%. Based on the limited empirical evidence available, the findings cautiously support the conclusion that a comparable green discount exists in both the European and U.S. commercial real estate markets.

Table 6.3 presents the results of the models with the *LogTargetSpread* as the dependent variable. These findings largely align with the results of the models outlined above. Model 3 once again exhibits the highest goodness of fit. Notable deviations are observed in certain coefficients. For example, in the case of customer ratings, which reflect the same reference variable, the coefficients now have a higher average value. Specifically, the riskiest customer rating category 5 is now associated with a significantly higher coefficient of 0.841 or 131.87% compared to the reference category of customer rating

1. Regarding property types, it is noteworthy that the coefficient for the mall variable becomes negative, implying a lower target spread as for offices, all else being equal. This finding is rather surprising in light of the increasing pressures from e-commerce and the post-COVID environment. In terms of cities, the coefficient for Luxembourg turns statistically significant and positive (0.533 or 70.40% compared to the reference Germany: old states A-cities).

Table 6.3: Hedonic Model Results where $y = LogTargetSpread$

Model	1	2	3	4
(Intercept)	-0.020 (0.059)	-0.012 (0.060)	-0.206*** (0.059)	-0.220*** (0.060)
Green building	0.038+ (0.021)	0.019 (0.021)	-0.040* (0.020)	0.017 (0.056)
LTV (in %)	0.001* (0.001)	0.002* (0.001)	0.002*** (0.001)	0.002*** (0.001)
Volume (in t. €)	0.000 (0.000)	0.000 (0.000)	0.000+ (0.000)	0.000 (0.000)
Maturity (in months)	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	0.000+ (0.000)
Investor profile (ref: developer)				
Fund	-0.150** (0.053)	-0.191*** (0.053)	-0.230*** (0.049)	-0.231*** (0.049)
Foreign investor	-0.136*** (0.031)	-0.262*** (0.037)	-0.286*** (0.034)	-0.286*** (0.034)
Domestic investor	-0.149*** (0.031)	-0.158*** (0.031)	-0.171*** (0.028)	-0.170*** (0.028)
Investment management	-0.204*** (0.043)	-0.236*** (0.044)	-0.249*** (0.040)	-0.247*** (0.040)
Housing company	-0.297*** (0.050)	-0.348*** (0.051)	-0.377*** (0.047)	-0.369*** (0.048)
Customer rating (ref: 1)				
Customer rating 2	0.141*** (0.028)	0.124*** (0.027)	0.146*** (0.025)	0.146*** (0.025)
Customer rating 3	0.445*** (0.028)	0.421*** (0.028)	0.411*** (0.026)	0.411*** (0.026)
Customer rating 4	0.555*** (0.034)	0.526*** (0.035)	0.510*** (0.032)	0.509*** (0.032)
Customer rating 5	0.962*** (0.087)	0.921*** (0.086)	0.841*** (0.078)	0.831*** (0.079)
Property type (ref: office)				
Retail	-0.026 (0.033)	-0.003 (0.034)	-0.001 (0.031)	-0.001 (0.031)
Specialty retail centers	0.150+ (0.080)	0.153+ (0.080)	0.030 (0.074)	0.034 (0.074)
Hotel	0.321** (0.112)	0.337** (0.110)	0.246* (0.100)	0.256* (0.101)
Department store	0.170 (0.111)	0.124 (0.111)	-0.015 (0.102)	-0.030 (0.103)
Logistics	0.033 (0.031)	0.049 (0.033)	0.027 (0.030)	0.023 (0.030)
Management properties	-0.107 (0.065)	-0.093 (0.065)	-0.049 (0.059)	-0.046 (0.059)
Parking garage	0.497*** (0.125)	0.414** (0.127)	0.245* (0.117)	0.240* (0.118)
Mall	-0.348* (0.143)	-0.385** (0.142)	-0.489*** (0.129)	-0.483*** (0.130)
Residential investment	-0.124*** (0.025)	-0.106*** (0.025)	-0.142*** (0.023)	-0.144*** (0.023)
Social property	0.498** (0.176)	0.496** (0.173)	0.445** (0.158)	0.448** (0.158)
Other	-0.338*** (0.084)	-0.313*** (0.082)	-0.316*** (0.075)	-0.321*** (0.076)
Region (ref: DE: old states A-cities)				
DE: old states B-cities		0.053 (0.033)	0.054+ (0.030)	0.059+ (0.031)

Continued on next page

Table 6.3: Hedonic Model Results where $y = LogTargetSpread$ (Continued)

DE: old states (other)		-0.037 (0.033)	0.007 (0.030)	0.011 (0.030)
DE: new states B-cities		0.023 (0.053)	0.037 (0.049)	0.041 (0.049)
DE: new states (other)		-0.038 (0.057)	-0.045 (0.052)	-0.039 (0.052)
DE: Berlin		0.017 (0.030)	0.026 (0.027)	0.027 (0.027)
Belgium		0.147 (0.124)	0.248* (0.113)	0.255* (0.115)
France		0.173*** (0.050)	0.209*** (0.046)	0.215*** (0.046)
Great Britain		0.442+ (0.243)	0.333 (0.222)	0.333 (0.223)
Luxembourg		0.395 (0.243)	0.533* (0.222)	0.539* (0.224)
Netherlands		0.145*** (0.039)	0.163*** (0.036)	0.162*** (0.036)
Poland		0.242*** (0.052)	0.251*** (0.047)	0.250*** (0.047)
Other countries		0.047 (0.243)	-0.045 (0.222)	-0.061 (0.222)
Year (ref: 2018)				
Year 2019			0.062* (0.027)	0.075* (0.031)
Year 2020			0.188*** (0.028)	0.195*** (0.032)
Year 2021			0.264*** (0.029)	0.295*** (0.034)
Year 2022			0.288*** (0.029)	0.300*** (0.037)
Year 2023			0.243*** (0.044)	0.207*** (0.057)
Green building x 2019				-0.066 (0.068)
Green building x 2020				-0.047 (0.070)
Green building x 2021				-0.107 (0.068)
Green building x 2022				-0.060 (0.067)
Green building x 2023				0.049 (0.097)
Num.obs.	728	728	728	728
R^2	0.609	0.630	0.696	0.698
R^2 adj.	0.595	0.611	0.678	0.678
AIC	38.4	21.7	-111.9	-106.8
BIC	157.7	196.1	85.5	113.5
RMSE	0.24	0.23	0.21	0.21
Regional effects	N	Y	Y	Y
Time effects	N	N	Y	Y

Note: The standard errors are indicated in parentheses. The levels of statistical significance are represented as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models are estimated using heteroskedasticity-robust covariance matrix estimators. The variable green building is a dummy variable that is one if the respective asset is classified as a green building by the lender, according to the classification described in the text. The categorical variables for investor profile describe the classification of the type of investor made by the lender and is independent of the financed asset. The variable customer rating describes the risk classification of the customer made by the lender, where 1 = least risky and 5 = most risky, and is also independent of the asset. The year refers to the year of financing issuance. The interaction term between year and green building is one if a financed asset was classified as green in the respective years.

Central to our research question is a significant green building dummy variable in Model 3, which indicates a green discount in the lender's target spreads. With a coefficient of -0.040 or -3.92%, and taking into account the average loan size and the average target spreads in our sample, we can observe a discount of approximately 5 bps, which translates into interest savings of €20,776 in the first year. In this context, it is crucial to note that the green discount is comparatively lower in the lender's internal calculations than in the contracted spreads. This distinction promotes a nuanced understanding of the green discount from both the lender's and the borrower's perspective. As for the *LogSpread*, it can be observed that the interaction term between the year and the green building variable in Model 4 does not exert a significant influence on the *LogTargetSpread* and suggests an increase in multicollinearity.

We attribute the green discount in loan spreads to two potential drivers. One perspective is that lenders perceive a reduced risk associated with green properties, as suggested by previous research from An and Pivo (2020) and Devine and Yönder (2021). The lower default risk of green properties identified by An and Pivo (2020) is primarily reflected in the specific risk component of the spread. A green premium in the income component of properties has been extensively documented in the literature. It translates into improved operational performance and higher Net Operating Income (NOI). This risk consideration holds significant importance for lenders, given that a higher NOI correspondingly improves the property's Debt Service Coverage Ratio (DSCR), ensuring the property's capacity to meet fixed debt obligations encompassing both principal and interest payments. The presence of a green premium in property values also result in less collateral risk for lenders. A lower exposure to transition risks could further reduce the collateral risk of green properties. Unlike voluntary commercial certifications, the actual energy intensity of properties not only mirrors a current market standard that positively affects the property's marketability but also potentially poses an ongoing and future risk to the property's cash flow. This risk emanates from two principal factors: firstly, the costs for fossil energy sources are expected to continue rising; secondly, consumption levels exceeding regulatory thresholds will likely be subject to a carbon pricing mechanism. As previously mentioned, only a specific portion of the total energy consumption in the real estate sector can be met through renewable energy sources in the foreseeable future. This scenario underscores the critical need for enhancing energy efficiency and incorporating the gradual elimination of fossil fuels into asset management and development strategies. Looking ahead, the implementation of these measures is likely to be more about averting financial losses than about achieving actual value increases. This proactive approach is essential not only for sustainability but also for safeguarding the long-term financial viability and resilience of assets in an increasingly environmentally-conscious market. The importance for lenders of incorporating these transition risk drivers into their risk considerations is also evident in other sectors of the economy (Chava, 2014; Krueger et al., 2020; Bolton & Kacperczyk, 2021). With regard to the systematic risks, Geltner (1989) and Naranjo and Ling (1997) emphasize the risk premium resulting from the comovement between economic consumption and property performance. If lenders assume green properties to have lower correlation with general economic demand, driven by factors such as regulatory support, higher tenant demand, or cost efficiencies, this might also lead to a green discount within the systematic risk component of spreads.

Growing regulatory requirements and the impact of internal and external stakeholders on climate goals in the banking industry are adding a new dimension to the concept of a green discount. Although EU taxonomy does not require a specific minimum GAR, and several other economies have not yet adopted mandatory reporting of such ratios, the proportion of green real estate is having an increasingly significant impact on transparency and comparability of lenders' portfolios. This may result in increased tensions among management, stakeholders, and shareholders with respect to a sustainability strategy on one hand, and the financial implications of reputational harm and liability concerns stemming from greenwashing or even detrimental environmental effects on the other (Brühl, 2023). In addition to the conventional components of the spread mentioned above, lenders may intentionally price in a green component that gives them room for negotiation in the loan origination process. Permitting a green discount, as evidenced by our findings, could facilitate the origination of green loans and consequently increase the share of green properties in the lenders' loan portfolios. As shown in Figure 6.5, the proportion of capital made available for green loans has been steadily increasing over recent years. While the green discount may factor into internal spread calculations, there is currently no evidence of any commercial lender in the European market actively offering such finance concessions. Regarding the US market this observation is also noted by Mathew et al. (2021).

The distinction between the target spread and the contracted spread in our analysis provides a valuable framework for evaluating the green discount from both the lender's and borrower's perspectives. Under the previously established hypothesis that lenders deliberately price a green component into spreads, the green discount in the target spread can be understood as the internal value that lenders associate with the origination of green loans. Assuming economic rationality, lenders will seek to minimize the green discount. Hence, the difference between the green discount within the target spread and the contracted spread may be interpreted as an indicator of how the supply side of green buildings perceives its bargaining position during loan origination. As shown in Tables 6.2 and 6.3, the comparison of both regressions in Model 3 indicates that the green discount is 36% higher for contracted spreads. Even though the difference of approximately 2 basis points between the two is negligible from an economic standpoint, it signifies that lenders are required to offer a marginally higher green discount in loan origination than their initial calculations indicate. With regard to borrowers holding green properties, our results suggest that they are conscious of their leverage and pursue the greatest possible green discount.

Table 6.4 presents the regression results for the dependent variable LTV. Model 3 again has the best fit and is used as the basis for interpreting the coefficients. However, the model only achieves an adjusted R^2 of 0.379, which is significantly lower than in the (target) spread analysis. We attribute this lower explanatory power to the fact that, unlike spreads, the absolute level of the LTV depends not only on property- and borrower-specific factors, but also on exogenous drivers that are not captured by the model.

Table 6.4: Hedonic Model Results where $y = LTV$ (in %)

Model	1	2	3	4
(Intercept)	67.100*** (2.012)	66.332*** (2.135)	70.669*** (2.346)	71.099*** (2.428)
Green building	-0.016 (1.134)	0.058 (1.149)	1.123 (1.184)	-1.795 (3.335)
LogMargin	4.039* (2.052)	4.683* (2.094)	8.339*** (2.271)	8.525*** (2.278)
Volume (in t. €)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Maturity (in months)	-0.038** (0.014)	-0.033* (0.014)	-0.038** (0.014)	-0.037** (0.014)
Investor profile (ref: developer)				
Fund	-13.461*** (2.855)	-14.171*** (2.928)	-11.930*** (2.940)	-11.779*** (2.943)
Foreign investor	-16.561*** (1.572)	-13.930*** (2.067)	-12.115*** (2.099)	-12.170*** (2.106)
Domestic investor	-8.995*** (1.680)	-9.301*** (1.695)	-8.214*** (1.700)	-8.205*** (1.703)
Investment management	-19.822*** (2.271)	-20.272*** (2.339)	-18.333*** (2.354)	-18.348*** (2.360)
Housing company	-3.207 (2.802)	-3.833 (2.927)	-1.863 (2.961)	-1.645 (2.982)
Customer rating (ref: 1)				
Customer rating 2	3.342* (1.527)	3.423* (1.535)	2.306 (1.540)	2.254 (1.544)
Customer rating 3	8.061*** (1.760)	7.947*** (1.765)	6.292*** (1.790)	6.267*** (1.799)
Customer rating 4	11.396*** (2.158)	11.514*** (2.164)	9.533*** (2.193)	9.478*** (2.200)
Customer rating 5	15.812** (5.081)	14.874** (5.077)	12.090* (5.054)	12.294* (5.064)
Property type (ref: office)				
Retail	2.229 (1.800)	0.788 (1.889)	0.668 (1.869)	0.615 (1.877)
Specialty retail centers	6.609 (4.345)	5.177 (4.429)	7.107 (4.418)	7.590+ (4.437)
Hotel	-6.638 (6.109)	-5.923 (6.095)	-5.716 (6.038)	-5.986 (6.056)
Department store	-2.798 (6.057)	-2.330 (6.155)	-0.711 (6.135)	-0.027 (6.169)
Logistics	-3.071+ (1.692)	-4.843** (1.821)	-4.148* (1.807)	-4.034* (1.821)
Management properties	-1.772 (3.575)	-2.755 (3.570)	-2.888 (3.536)	-3.238 (3.546)
Parking garage	1.988 (6.915)	5.036 (7.037)	5.902 (7.024)	7.291 (7.075)
Mall	-1.126 (7.816)	-2.734 (7.861)	0.095 (7.842)	1.173 (7.906)
Residential investment	2.233 (1.364)	2.204 (1.384)	3.347* (1.400)	3.455* (1.407)
Social property	0.294 (9.659)	-1.439 (9.628)	-1.542 (9.550)	-2.049 (9.565)
Other	11.167* (4.593)	10.337* (4.585)	10.635* (4.563)	10.869* (4.589)
Region (ref: DE: old states A-cities)				
DE: old states B-cities		4.140* (1.836)	3.944* (1.823)	3.853* (1.836)
DE: old states (other)		3.370+ (1.801)	2.254 (1.809)	2.201 (1.823)
DE: new states B-cities		0.278 (2.939)	-0.190 (2.913)	-0.589 (2.923)
DE: new states (other)		2.285 (3.141)	1.958 (3.125)	1.809 (3.139)
DE: Berlin		-1.490 (1.636)	-1.882 (1.622)	-1.982 (1.627)
Belgium		-4.201 (6.825)	-6.266 (6.807)	-5.438 (6.919)
France		-3.549 (2.802)	-4.977+ (2.793)	-5.257+ (2.803)
Great Britain		-11.876 (13.435)	-11.851 (13.308)	-9.877 (13.369)

Continued on next page

Table 6.4: Hedonic Model Results where $y = LTV$ (in %) (Continued)

Luxembourg		20.682 (13.431)	17.093 (13.353)	18.140 (13.460)
Netherlands		-3.439 (2.182)	-4.302* (2.172)	-4.111+ (2.196)
Poland		-2.007 (2.901)	-3.079 (2.887)	-2.844 (2.904)
Other countries		-0.025 (13.415)	0.654 (13.298)	1.266 (13.326)
Year (ref: 2018)				
Year 2019			-4.152* (1.633)	-4.231* (1.844)
Year 2020			-6.817*** (1.722)	-6.882*** (1.937)
Year 2021			-6.464*** (1.820)	-7.756*** (2.147)
Year 2022			-6.921*** (1.870)	-9.338*** (2.293)
Year 2023			-8.508** (2.695)	-8.035* (3.410)
Green building x 2019				1.403 (4.078)
Green building x 2020				1.030 (4.174)
Green building x 2021				4.650 (4.067)
Green building x 2022				6.287 (4.008)
Green building x 2023				0.391 (5.796)
Num.obs.	728	728	728	728
R^2	0.377	0.396	0.414	0.418
R^2 adj.	0.356	0.364	0.379	0.379
AIC	5852.0	5854.1	5841.5	5846.5
BIC	5971.3	6028.5	6038.8	6066.8
RMSE	13.07	12.87	12.67	12.63
Regional effects	N	Y	Y	Y
Time effects	N	N	Y	Y

Note: The standard errors are indicated in parentheses. The levels of statistical significance are represented as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models are estimated using heteroskedasticity-robust covariance matrix estimators. The variable green building is a dummy variable that is one if the respective asset is classified as a green building by the lender, according to the classification described in the text. The categorical variables for investor profile describe the classification of the type of investor made by the lender and is independent of the financed asset. The variable customer rating describes the risk classification of the customer made by the lender, where 1 = least risky and 5 = most risky, and is also independent of the asset. The year refers to the year of financing issuance. The interaction term between year and green building is one if a financed asset was classified as green in the respective years.

With a statistically significant coefficient of 8.339, the spread variable has a positive impact on the LTV. This observation is consistent with the results of the models with *LogSpread* and *LogTargetSpread* as dependent variables, where the LTV coefficient is consistently positive and significant. However, the interaction is marginal and holds relatively little importance in the overall context. Our results align with the findings of Titman et al. (2005) who attribute the weak link between LTV and spreads to the fact that the LTV is determined endogenously. Factors such as negotiations between lenders and borrowers may have an impact on the final LTV choice. The coefficients further indicate that the term to maturity has only limited impact on LTV ratios.

The analysis of the coefficients of the investor profile variables shows that the reference category developer has the highest LTV, while the other variables are associated with negative coefficients. This is in line with our initial hypothesis that developers are constrained by their equity capital and have to rely heavily on debt capital. In contrast, investment managers have the largest negative coefficient at -18.333, which is expected since they typically execute equity investments for their clients with a limited amount of debt. However, the results for customer ratings are not straightforward to interpret. Similar to the spread models, higher risk ratings are associated with higher average LTVs. This could be attributed, on the one hand, to the risk-return profile of the lender and, on the other hand, to the fact that the types of investors involved in highly leveraged projects generally represent a higher risk for banks at the corporate level.

Looking at the different property types in our analysis, there is no significant effect on the LTV, except for the coefficients of logistics and residential properties. In contrast to the office reference category, the residential category exerts a positive influence on the LTV. Consequently, residential properties within our sample tend to secure higher LTVs, assuming all other factors being equal. This implies that lenders perceive residential properties as less risky when compared to office properties. Additionally, consistent with the results from spread models, promotional loans for residential properties may enable higher LTV ratios at lower spreads. The lender's approach appears to be different for properties in the Netherlands. Compared to the reference category Germany: old states A-cities, the Netherlands dummy results in a spread premium and a negative LTV adjustment of -4.302, indicating a more conservative lending approach.

The time effects introduced in Model 3, using 2018 as the reference year, consistently show significant negative values and lead to an improvement in the adjusted R^2 compared to Model 2. While loans had higher LTV ratios on average in 2018, there is a modest downward trend up to and including 2023. This observation is in line with a European-wide trend in recent years, according to which the LTV ratios for commercial real estate have declined on average (European Central Bank, 2023).

Unlike the green discount observed in credit spreads, none of our models with LTV as the dependent variable indicates a significant influence from the green building coefficient. As with the other models in our analysis, the interaction terms in Model 4 fail to enhance the goodness of fit and lack statistical significance. A further descriptive analysis of our dataset with a focus on developers does not reveal any difference in LTVs between green and non-green projects. Given the market barriers and market failures highlighted by Chegut et al. (2019), our findings suggest that green properties and especially green development projects are not incentivized with additional loan proceeds. As a result, developers are unable to secure extra senior debt to cover shortages in capital during the initial stages of green construction projects, which involve higher upfront costs and extended construction timelines. Instead, they must seek alternative financing, such as additional equity or mezzanine capital, thus leaving market barriers for developers of green buildings in place. However, findings by Chegut et al. (2019) also reveal that the marginal costs associated with green construction projects are offset by the green value premium they achieve after completion. The lack of statistical significance of the green variable in our sample may, therefore, be explained by the fact that investing in properties or financing green projects after their completion does not specifically require more debt capital compared to their non-green peers due to the value premium they achieve.

6.6 Conclusions

Despite the robustness of our statistical models and the high quality of the data utilized, there are some limitations in interpreting the results of any hedonic model. A primary limitation is the challenge of completely controlling for endogeneity issues between green and non-green properties, even with the application of comprehensive control variables. Incorporating cash flow-related metrics like the DSCR into the existing dataset could significantly mitigate potential biases in loan risk assessment, thereby enhancing the overall robustness of the results. Furthermore, it is important to note that the dataset is representative of a specific sample, confined to a particular period and region. Consequently, the insights derived from this dataset may not be universally extendable to all commercial real estate lenders.

Our study emphasizes that the assessment of a property's environmental performance has become an integral part of loan origination for lenders in the commercial real estate sector. Due to the rise in regulations, such as EU taxonomy, lenders are expanding their assessment criteria beyond commercial certification schemes and are placing greater emphasis on the actual energy efficiency of buildings. At the same time, this implies that borrowers have to meet more stringent property standards in order to qualify for a green loan.

Although confirming previous research that finance concessions associated with green properties are not publicly offered by commercial lenders, our findings still provide evidence for a green discount in loan spreads. Applying a hedonic regression analysis while accounting for both observed and unobserved heterogeneity in loan data, we identified a reduction of 3.92% in lender's target spreads for green properties and 5.35% in contract spreads, respectively. The size of the green discount is consistent with previous findings for the U.S. market. This reduction seems to remain constant throughout the observed timeframe, without any statistically notable changes in the green discount across the years. We attribute the green discount to two potential drivers. Firstly, lenders perceive lower risk associated with green properties, which has a positive impact on the risk components in spreads. Secondly, lenders provide some leeway for negotiation within the commission spreads during the loan origination process, allowing them both to incentivize green loans and thus to increase their own green asset ratio. The spread of the green discount between the calculated target spread and the contractually agreed spread can be interpreted as the deviation between the internal value that lenders place on green loans and the bargaining power of borrowers holding green properties. In light of the economically marginal spread of approximately 2 basis points, borrowers align with lenders' perceptions of fair market conditions. However, the supply side of green buildings is well aware of its bargaining power, pushing lenders to grant a slightly higher green discount than their internal calculations indicate.

While incentives for borrowers are already reflected in lower spreads, our examination does not indicate higher LTV ratios for green properties, leading to additional loan proceeds for borrowers. As a result, lenders are not fully addressing the potential shortage of capital on the supply side, stemming from the higher up-front costs and and extended construction timelines associated with green buildings, by providing additional debt capital. Furthermore, banks are forced to further improve their own green asset ratio in the face of new regulations and net zero targets. Our sample highlights a gradual greening of lenders' portfolios in recent years, yet opportunities for improvement still remain. Introducing stronger incentives for the supply side of green buildings is one way to increase this ratio in the future, simultaneously stimulating more green construction and retrofits. Future research may show whether lenders globally are adopting a comprehensive green discount, thereby contributing to the elimination of market inefficiencies and promoting the adoption of energy efficient buildings, as suggested by the green premium observed on the equity side of real estate.

6.7 Appendix

Table 6.5: Results of Variance Inflation Factor Models where $y = LogSpread$

Model	Variable	$GVIF$	Df	$GVIF^{(1/(2*Df))}$
Model 1	Investor profile	4.322232	6	1.129733
	Customer rating	2.199901	8	1.05051
	Property type	2.334517	11	1.039289
	Green building	1.166907	1	1.080235
	LTV (in %)	1.636945	1	1.279432
	Volume (in t. €)	1.091463	1	1.044731
	Maturity (in months)	1.576548	1	1.255606
Model 2	Investor profile	10.138731	6	1.212919
	Customer rating	2.641558	8	1.062591
	Property type	3.825711	11	1.062887
	Green building	1.2155	1	1.102497
	LTV (in %)	1.693176	1	1.301221
	Volume (in t. €)	1.133989	1	1.064889
	Maturity (in months)	1.654996	1	1.286466
	Region	6.306539	12	1.079754
Model 3	Investor profile	10.848957	6	1.219782
	Customer rating	2.819923	8	1.06694
	Property type	4.422519	11	1.069913
	Green building	1.3162	1	1.147257
	LTV (in %)	1.727224	1	1.314239
	Volume (in t. €)	1.138377	1	1.066947
	Maturity (in months)	1.715995	1	1.30996
	Region	7.200627	12	1.085735
	Year	1.708522	5	1.055023
Model 4	Investor profile	11.322948	6	1.224137
	Customer rating	2.983629	8	1.070709
	Property type	4.840777	11	1.074317
	Green building	10.521224	1	3.243644
	LTV (in %)	1.735328	1	1.317318
	Volume (in t. €)	1.147941	1	1.07142
	Maturity (in months)	1.729108	1	1.314956
	Region	8.316685	12	1.092273
	Year	15.46262	5	1.315008
	Green building x year	100.797984	5	1.586153

6.8 References

An, X., & Pivo, G. (2020). Green buildings in commercial mortgage-backed securities: The effects of LEED and energy star certification on default risk and loan terms. *Real Estate Economics*, *48*(1), 7-42.

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of financial economics*, *142*(2), 517–549.

Bond, S. A., & Devine, A. (2016). Certification matters: Is green talk cheap talk? *The Journal of Real Estate Finance and Economics*, *52*(2), 117–140.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, 1287–1294.

Brühl, V. (2023). The green asset ratio (GAR)–a new KPI for credit institutions. *Center for Financial Studies Working Paper*, *683*.

Cadena, A., & Thomson, T. A. (2021). The value of "green" in resale residential real estate: Premium by neighborhood value quintile, homestead status and year. *Journal of Sustainable Real Estate*, *13*(1), 12–29.

Cajias, M., Fuerst, F., & Bienert, S. (2019). Tearing down the information barrier: the price impacts of energy efficiency ratings for buildings in the German rental market. *Energy Research & Social Science*, *47*, 177–191.

Chava, S. (2014). Environmental externalities and cost of capital. *Management science*, *60*(9), 2223–2247.

Chegut, A., Eichholtz, P., & Kok, N. (2014). Supply, demand and the value of green buildings. *Urban studies*, *51*(1), 22–43.

Chegut, A., Eichholtz, P., & Kok, N. (2019). The price of innovation: An analysis of the marginal cost of green buildings. *Journal of Environmental Economics and Management*, *98*, Article 102248.

Dell'Anna, F., & Bottero, M. (2021). Green premium in buildings: Evidence from the real estate market of Singapore. *Journal of Cleaner Production*, *286*, 125327.

Deng, Y., Li, Z., & Quigley, J. M. (2012). Economic returns to energy-efficient investments in the housing market: evidence from Singapore. *Regional Science and Urban Economics*, *42*(3), 506–515.

Devine, A., & McCollum, M. (2021). Advancing energy efficiency through green bond policies. *Available at SSRN 3856312*.

Devine, A., & Yönder, E. (2021). Impact of environmental investments on corporate financial performance: Decomposing valuation and cash flow effects. *The Journal of Real Estate Finance and Economics*, 1–28.

Eichholtz, P., Holtermans, R., Kok, N., & Yönder, E. (2019). Environmental performance and the cost of debt: Evidence from commercial mortgages and REIT bonds. *Journal of Banking & Finance*, *102*, 19–32.

Eichholtz, P., Kok, N., & Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *American Economic Review*, *100*(5), 2492–2509.

Eichholtz, P., Kok, N., & Quigley, J. M. (2013). The economics of green building. *Review of Economics and Statistics*, *95*(1), 50–63.

Eichholtz, P., Kok, N., & Yonder, E. (2012). Portfolio greenness and the financial performance of REITs. *Journal of International Money and Finance*, *31*(7), 1911–1929.

European Central Bank. (2023). Financial stability review, May 2023. *Published Online*. Retrieved from https://www.ecb.europa.eu/pub/financial-stability/fsr/special/html/ecb.fsrart202005_01~762d09d7a2.en.html

European Parliament. (2023). Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending regulation (EU) 2023/955 (recast). *Official Journal of the European Union*.

Fisher, I. (1930). The theory of interest. *New York*, *43*, 1–19.

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, *87*(417), 178–183.

Fuerst, F. (2015). The financial rewards of sustainability: A global performance study of real estate investment trusts. *Available at SSRN 2619434*.

Fuerst, F., & McAllister, P. (2011a). Green noise or green value? Measuring the effects of environmental certification on office values. *Real Estate Economics*, *39*(1), 45–69.

Fuerst, F., & McAllister, P. (2011b). The impact of energy performance certificates on the rental and capital values of commercial property assets. *Energy Policy*, *39*(10), 6608–6614.

Fuerst, F., McAllister, P., Nanda, A., & Wyatt, P. (2015). Does energy efficiency matter to home-buyers? an investigation of EPC ratings and transaction prices in England. *Energy Economics*, *48*, 145–156.

Geltner, D. (1989). Estimating real estate's systematic risk from aggregate level appraisal-based returns. *Real Estate Economics*, *17*(4), 463–481.

Holtermans, R., & Kok, N. (2019). On the value of environmental certification in the commercial real estate market. *Real Estate Economics*, *47*(3), 685–722.

International Energy Agency. (2021). Net zero by 2050. A roadmap for the global energy sector. *Part of Net Zero Emissions Online Publications*.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259.

Kahn, M. E., & Kok, N. (2014). The capitalization of green labels in the california housing market. *Regional Science and Urban Economics*, *47*, 25-34.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, *33*(3), 1067–1111.

Mathew, P., Issler, P., & Wallace, N. (2021). Should commercial mortgage lenders care about energy efficiency? Lessons from a pilot study. *Energy Policy*, *150*, 112137.

Naranjo, A., & Ling, D. C. (1997). Economic risk factors and commercial real estate returns. *The Journal of Real Estate Finance and Economics*, *14*, 283–307.

Ooi, J. T., & Dung, D. D. (2019). Finding superior returns in green portfolios: Evidence from Singapore REITs. *Journal of Sustainable Real Estate*, *11*(1), 191–215.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673–690.

Reichardt, A., Fuerst, F., Rottke, N., & Zietz, J. (2012). Sustainable building certification and the rent premium: a panel data approach. *Journal of Real Estate Research*, *34*(1), 99–126.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Sah, V., Miller, N., & Ghosh, B. (2013). Are green REITs valued more? *Journal of Real Estate Portfolio Management*, *19*(2), 169–177.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, *13*(1), 1–44.

Titman, S., Tompaidis, S., & Tsyplakov, S. (2005). Determinants of credit spreads in commercial mortgages. *Real Estate Economics*, *33*(4), 711–738.

UNEP FI. (2019). Principles for Responsible Banking. *Published Online*. Retrieved from https://www.unepfi.org/banking/bankingprinciples/

UNEP FI. (2021a). Guidelines for Climate Target Setting for Banks. *Published Online*. Retrieved from https://www.unepfi.org/industries/banking/guidelines-for-climate-target-setting-for-banks/

UNEP FI. (2021b). Net-Zero Banking Alliance. *Published Online*. Retrieved from https://www.unepfi.org/net-zero-banking/

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817–838.

Wiley, J. A., Benefield, J. D., & Johnson, K. H. (2010). Green design and the market for commercial office space. *The Journal of Real Estate Finance and Economics*, *41*, 228–243.

7 Conclusion

7.1 Executive Summary

This section provides a brief overview and summary of each of the five scientific papers. First, the context and objectives are addressed, followed by an explanation of the utilized dataset and methodology, and concluding with a summary of the results.

7.1.1 Rooftop Solar Potential and Financial Performance of REITs

Background and Objectives
Transition risks play a critical role in REIT markets and should therefore be considered when managing and investing in REITs (Salisu et al., 2024). Scientific research quantifies the effects of transition risks by evaluating properties labeled as sustainable (e.g., via certificates) or benchmarking (e.g., via GRESB). While both of these measures may be appropriate for determining the status quo of a building with respect to sustainability performance, they do not address necessary future optimization to minimize carbon emissions. Akhtyrska and Fuerst (2024) have shown that the focus on actual carbon emissions and energy intensity is important to managing transition risk in the portfolio. In this context, electrifying the property with renewable energy is a widely used option to reduce greenhouse gas emissions. Since a property's ability to mitigate greenhouse gas emissions using on-site generated green electricity is highly dependent on the roof area available for solar utilization, it can serve as a proxy for the potential to minimize transitory risks. In the first study, we therefore examine the extent to which solar potential or displaced carbon is reflected in the financial performance of REITs.

Data and Methodology
We use a unique, self-compiled dataset that includes U.S. REIT firm characteristics, daily stock market data, and annualized alphas and betas calculated using the 4-factor model based on Fama and French (1993) and Carhart (1997). Additional variables for each asset owned by the REIT sample are compiled using the Google Solar API, which employs machine learning to estimate the available solar rooftop area (in square feet) and the corresponding carbon emissions (in kilograms of CO_2 per megawatt-hour) that a solar-power-system would displace. These values, referred to as the solar potential

and carbon offset factor, are then weighted according to the available floor area of each property. To ensure meaningful results, particularly with respect to firm effects, we use multiple models and variable combinations to measure the impact of solar potential and carbon offset factors on REIT financials. First, we apply the 4-factor model using alphas (abnormal returns) and betas (systematic risk) as dependent variables. We then examine the impact of our key independent variables on ROA, ROE, M/B, and P/NAV. Using an OLS model, we further regress the same variables with industry fixed effects and heteroskedasticity-robust standard errors adjusted for within-cluster dependencies. Robustness checks include subsample analysis, coefficient calculation using time-clustered standard errors, and analysis for unobserved confounders following Cinelli and Hazlett (2020).

Results and Remarks
We analyzed the extent to which REITs' rooftop solar potential and corresponding offset factors can be considered as drivers of stock performance, operating performance, and firm value. Our results indicate that while the market does not currently fully reflect REITs' solar adoption potential in terms of alphas and betas, there is strong evidence of a significant impact on firm value. We conclude that this impact extends beyond omitted financial characteristics, with the market placing a higher value on REITs' potential future earnings due to the financial benefits of solar adoption. The significant effect of the carbon offset factor suggests that REIT financials are affected by transition risk even beyond firms' direct control over emissions. Policymakers and management can capitalize on these findings by promoting the installation of solar systems, particularly in regions with high emission factors. The solar potential of REITs is significant and offers a way to decarbonize properties and mitigate transition risk in a portfolio context.

7.1.2 Picture This: A Deep Learning Model for Operational Real Estate Emissions

Background and Objectives
Reliable data are needed to make strategic decisions about where to invest in buildings with the lowest abatement costs. This is important both politically and economically, especially with regard to targeted retrofit rates (Bienert & Groh, 2022). However, data on real estate emissions has been highly opaque and largely not centrally collected on a global scale. The European Union has recently attempted to harmonize EPCs through a reform of the EPBD, as EPCs in their current form vary widely across member states and emissions data is not always reported (Semple & Jenkins, 2020). However, as there has been little progress in the collection and assessment of operational real estate emissions

worldwide, and as decarbonization efforts need to be further intensified, there is a gap between the data available and the data needed by policymakers and investors to make informed decisions. The aim of this paper is to investigate the extent to which currently available data sources can be leveraged to make automated out-of-sample predictions of operational emissions from residential buildings using SOTA deep learning methods.

Data and Methodology
We utilized the UK EPC database as our primary dataset. After thoroughly cleaning the dataset to ensure it was up to date, we randomly selected 21,000 EPCs from different regions across England. Each EPC, with ratings ranging from 0 to 140 CO_2e per m^2 and evenly distributed across seven classes, was then matched with and supplemented by the corresponding exterior image of the property obtained from GSV. The resulting variable quality of the images required a manual review, in which approximately 14.66% of the images were found to be of insufficient quality and were therefore removed from the dataset. The remaining images, with their corresponding CO_2e per m^2 values, were used as a training base for five ResNet models, which differed in depth and architecture (He et al., 2016). Each model was trained using 80% training data, 20% validation data and 50 epochs. The results are evaluated using common KPIs such as accuracy and F1-Score. The validity of the results was ensured using the visual Grad-CAM approach according to Selvaraju et al. (2017).

Results and Remarks
The ResNet34 model, which we trained using the building images and emissions data, achieved an accuracy of 43.38%. Using image augmentation and the addition of primary energy sources, the accuracy increased to 47.59%. Particularly high F1-Scores were found for emission-efficient buildings. When the immediate neighboring classes are included, the accuracy reaches 75.34%. The results, or rather the application of the model and the model architecture, can be utilized in practice to increase the transparency of data on operational emissions and to improve the basis for decision-making. From a research point of view, the results provide a first basis on which future research can be extended, for example by using additional data sources or in combination with existing models from other areas.

7.1.3 Real-GPT: Efficiently Tailoring LLMs for Informed Decision-Making in the Real Estate Industry

Background and Objectives
Since the introduction of ChatGPT, LLMs have received considerable attention, leading to the creation of numerous other models with variations in architecture or capabilities.

However, because these models are trained on general datasets, mostly collected from the internet, specific knowledge about less discussed areas is underrepresented in the training data. This leads to reduced capabilities in these areas for established LLMs (Wu et al., 2023; Taylor et al., 2022), including the real estate sector. One method to mitigate these effects is to fine-tune LLMs on custom datasets specifically tailored to the intended use case. Fine-tuning can not only improve knowledge about niche areas, but also lead to better results in other NLP tasks and allow the use of smaller, more resource-efficient models (Wei et al., 2022). In this paper, we present a method for fine-tuning open-source models on niche data—in our case real estate data—and compare the results with models that are not fine-tuned to real estate data. The resulting model, Real-GPT, is then benchmarked against typical real estate tasks such as investment scenarios and knowledge capabilities.

Data and Methodology

To our knowledge, no existing real estate-specific dataset contains structured text suitable for fine-tuning a LLM. Therefore, we first had to create an appropriate dataset, which we accomplished by collecting 670 files from various university databases, including both practical and academic literature. We manually curated each paragraph, resulting in over 500,000 lines of text. Using the GPT-3.5 Turbo API, we generated ten questions and corresponding answers for each section, resulting in a structured dataset of 17,769 real estate-specific entries. Subsequently, we further filtered these question-answer pairs for quality using the GPT-4 API. A similar process was used to create a test dataset. However, in this case, we generated three fictitious answers in addition to one true answer, requiring the model to identify the correct answer during testing. After fine-tuning the model using the QLoRA approach (Dettmers et al., 2023), we evaluated its capabilities on the generated multiple-choice dataset and on a repeated sales dataset where the model's objective was to maximize profits and provide reasons for its decisions.

Results and Remarks

Existing technologies related to NLP and AI are already having a significant impact on process efficiency, including in real estate. As technology continues to advance, it is expected that specific real estate applications will be better captured by custom models. Fine-tuning pre-trained models can be an effective way of adapting these models to specific needs. Real-GPT achieves comparable results to the closed-source SOTA model, despite having a relatively small number of parameters and being deployable on data-secure consumer hardware. This makes it an efficient alternative for integration into the daily tasks of real estate companies. Furthermore, our results suggest that fine-tuned models are able to make better investment decisions. Beyond real estate, the fine-tuning approach we propose can be applied to other niche domains. Future research in this area could explore how other methods, such as RAG or multimodal models, could further enhance the capabilities of real estate models.

7.1.4 Unveiling the Impact of SFDR on Unlisted Real Estate Funds: A J-Curve and Panel Regression Analysis

Background and Objectives
With the Sustainable Finance Disclosure Regulation becoming mandatory in all EU member states in March 2021, market participants must now consider sustainability characteristics when promoting real estate funds. Article 6, 8 or 9 of the legislation defines different criteria, each with increasing requirements for the sustainability characteristics reported. Amidst ongoing debates about the effectiveness of the current SFDR framework, market participants and researchers still have limited information about the financial impact or interactions of SFDR on fund products. Therefore, we examine the financial impact of the SFDR on returns of unlisted real estate funds using a unique dataset from a large European service capital investment company. In particular, our study focuses on how the typical phase-specific J-Curve of these funds is affected by the SFDR classifications.

Data and Methodology
The unbalanced dataset initially consists of 50 funds starting between March 2021 and March 2023. After cleaning, the time series of the remaining 40 funds include various characteristics such as risk classification and open/closed fund status, as well as the return measures IRR and BVI, which serve as dependent variables in the subsequent models. Methodologically, the paper is structured in two steps: First, the J-Curve of the fund classifications is determined and analyzed using cubic spline interpolation. This is followed by an evaluation using random effects models and econometric tests. The models differ in terms of the dependent variable and the inclusion of external variables outside the dataset, such as term- and loan spread.

Results and Remarks
The results suggest that Article 8 funds do not significantly underperform in comparison to Article 6 funds in the initial phase and even generate positive returns earlier. This suggests that initial fund costs are not affected by the SFDR category or that higher costs (e.g., those arising from enhanced due diligence) are recovered more quickly. Moreover, we observe a clear phase-typical progression according to the J-Curve of the funds. The panel regression results confirm that SFDR category has a significant impact on both IRR and BVI. However, this impact is slightly negative for Article 8

funds, contrary to the results of the J-Curve analysis. We attribute these effects in part to a more conservative risk-return profile and the likelihood that Article 8 funds hold newer buildings with lower (cash flow) return requirements. In addition, the relatively short time series is primarily driven by recent acquisitions, with the effects of valuation changes expected to materialize later. It remains to be seen how market participants and regulators will approach the objectives and implementation of the SFDR in the future. Current industry perspectives are critical of the present regulatory framework.

7.1.5 Is there a Green Discount in Commercial Real Estate Lending?

Background and Objectives

Sustainable finance is being prioritized by both regulators and the private sector because of its important role in the transition to a decarbonized economy. The capital-intensive real estate sector is particularly affected by these efforts. Literature on the financial benefits of sustainable buildings suggests that sustainable real estate investments are not only driven by philanthropic motives, but also promise measurable financial benefits on the equity side (Eichholtz et al., 2010; Bond & Devine, 2016). However, there is limited literature on the debt side of sustainable real estate investment. In the EU, lenders are incentivized to increase the share of sustainable real estate due to the mandatory reporting of the GAR in order to remain competitive. To investigate whether sustainable real estate (in our case, "green buildings" that comply with the EU taxonomy) can be financed at more favorable conditions, this paper analyzes a large dataset of commercial real estate financing.

Data and Methodology

The dataset covers over €30 billion worth of financed commercial real estate across Europe, provided by one of the largest financiers in the sector. This volume is spread across 728 loans over the period from January 2018 to June 2023. The dependent variables are the planned and contracted spread on the bank's cost of funds and the LTV ratio. The control variables include the green building indicator, loan volume, maturity, investor profile, client rating, property type, region, and year. Our empirical framework consists of a hedonic regression model where the first two models (target and contracted spread) are specified as semi-log models and the third model (y = LTV) is in linear form. In addition, interaction terms are included to capture the time effect related to the green building variable. Model robustness is ensured by varying the variables and analyzing relevant metrics, including adjusted R^2, AIC, BIC, and $RMSE$.

Results and Remarks

The results suggest that green buildings can be financed at more favorable terms compared to their non-green counterparts. An isolated effect of 3.92% on the target spread and 5.35% on the contracted spread was identified. Translated to the average asset size and spread, this results in an interest saving of 5 bps and €20,776 in the first year for the target spread and 7 bps and €28,356 for the contracted spread. The difference between the target and contracted spreads indicates that the actual interest savings are close to, but slightly higher than, the savings calculated by the lender. This may be due to the borrower's perceived bargaining power. No significant relationship was found between LTV and the green building variable. Future research should examine whether incentives to finance green buildings increase with rising regulatory requirements, especially as discussions move from mere reporting requirements under the GAR to the establishment of minimum standards.

7.2 Final Remarks

Environmental sustainability, *innovation*, and *regulation* stand in stark contrast to traditional real estate practices, requiring both academia and industry professionals to continually and intensely examine the potential impact on the real estate sector. This includes adapting to emerging developments and addressing financial implications and risks.

In terms of *environmental sustainability*, it is expected that even greater efforts will be required in the future. Forward-looking analyses indicate that global commitments to reduce carbon are far from sufficient to limit global warming to the target of 1.5 degrees Celsius. Current legislative efforts to limit climate change are even farther from achieving this goal (UNEP, 2023). In particular, each year that greenhouse gas emissions exceed planned limits further constrains the budget for future years by a corresponding amount, requiring more drastic interventions in the global economy. At the same time, failure to meet climate targets increases the realized climate risks for the real estate industry. There is ample evidence that real estate, as a significant component of the built environment, is affected by these climate risks throughout much of its value chain. As temperatures continue to rise, these risks are expected to intensify, ultimately leading to greater price declines. The dissertation findings indicate that real estate market participants are already considering the future potential for mitigating transitional risks through on-site electrification of properties, and such considerations are being factored in at the firm level.

Additionally, it has been demonstrated that new *innovations*, such as machine learning algorithms, are capable of significantly scaling up the comprehensive mining of data necessary for strategic sustainability decisions. Especially at the macro level using deep learning methods, estimating CO_2e emissions can be accomplished with comparatively high accuracy using widely available images and quality-assured energy certificates. Without sufficient data at the asset, portfolio, or firm level, decision making in real estate processes becomes challenging. Similar difficulties arise when using established transformer-based language models in a real estate context: the training data for these models is likely to contain limited real estate-specific content. In addition, there are no structured text datasets that could provide a remedy. To fully realize the significant economic potential—McKinsey estimates that the use of generative AI in real estate could add up to $180 billion in value—new approaches are needed (Chui et al., 2023). The results of the third paper suggest that adapting language models through techniques such as fine-tuning can improve their knowledge of real estate. This not only supports resource-efficient and privacy-compliant use of the models, but also improves their ability to make investment decisions.

In terms of *regulation*, it is widely expected that legislative measures will become stricter, thus increasing the requirements for both existing and new real estate. By assessing the impact of the SFDR and the EU Taxonomy Regulation on real estate, this dissertation primarily analyzes the impact of the Green Deal components. It was found that from an empirical perspective and considering the views of investors, there were negative financial impacts that were incurred, but also opportunities that presented themselves. Despite the increased effort required for SFDR reporting, the initial costs of the funds studied did not increase when sustainability characteristics had to be reported. In addition, buildings identified as sustainable under the EU taxonomy can be financed at a lower cost. This serves as an indicator of the effectiveness of the measures implemented by the Green Deal, as the private sector is financially incentivized to invest in real estate that is classified as sustainable according to the EU taxonomy. In addition to the SFDR and the EU Taxonomy Regulation, European market participants will face numerous other sustainability requirements in the future. For example, starting with the 2024 reporting year, the Corporate Sustainability Reporting Directive (CSRD), which requires sustainability reporting in management reports, will be gradually enforced and further specified by the European Sustainability Reporting Standards (ESRS). The findings of the dissertation are limited to specific time periods and deadlines, and therefore provide limited insight into future developments beyond the SFDR and the EU Taxonomy Regulation.

In general, the econometric models used are derived from historical data and primarily reflect the conditions prevailing at the time of analysis. It remains to be seen to what extent environmental sustainability will continue to be incorporated into investment decisions, although there are indications that the impact of sustainability on real estate is increasing. This is also true, albeit in a different way, for machine learning models as technological obsolescence progresses rapidly. Thus, for researchers and market participants in this field, there might be a fundamental need for methodological competence rather than the continued use of the specific methods. The need for further research in real estate economics, coupled with the three topic areas, will continue into the future. Both environmental sustainability and innovation, along with regulation, remain dynamic elements on the global agenda.

7.3 References

Akhtyrska, Y., & Fuerst, F. (2024). The effectiveness of climate change regulations in the commercial real estate market. *Energy Policy*, *185*, 113916.

Bienert, S., & Groh, A. M. (2022). Klimaneutralität vermieteter Mehrfamilienhäuser? Aber wie? [Climate neutrality for rented apartment buildings? But how?]. *Beiträge zur Immobilienwirtschaft*, *28*.

Bond, S. A., & Devine, A. (2016). Certification matters: Is green talk cheap talk? *The Journal of Real Estate Finance and Economics*, *52*(2), 117–140.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Chui, M., Roberts, R., Yee, L., Hazan, E., Singla, A., Smaje, K., ... Zemmel, R. (2023). The economic potential of generative AI: The next productivity frontier. *McKinsey Global Institute*. Retrieved from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(1), 39–67.

Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient finetuning of quantized LLMs. *arXiv Pre-Print*.

Eichholtz, P., Kok, N., & Quigley, J. M. (2010). Doing well by doing good? Green office buildings. *American Economic Review*, *100*(5), 2492–2509.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

Salisu, A. A., Ogbonna, A. E., & Vo, X. V. (2024). Climate risks and the REITs market. *International Journal of Finance & Economics, Advance online publication*.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.

Semple, S., & Jenkins, D. (2020). Variation of energy performance certificate assessments in the European Union. *Energy Policy*, *137*, Article 111127.

Taylor, R., Kardas, M., Cucurull, G., Scialom, T., Hartshorn, A., Saravia, E., … Stojnic, R. (2022). Galactica: A large language model for science. *arXiv Pre-Print*.

UNEP. (2023). Emissions gap report 2023: Broken record – temperatures hit new highs, yet world fails to cut emissions (again). *United Nations Environment Programme*. Retrieved from https://doi.org/10.59117/20.500.11822/43922

Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., … Le, Q. V. (2022). Finetuned language models are zero-shot learners. *arXiv Pre-Print*.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., … Mann, G. (2023). Bloomberggpt: A large language model for finance. *arXiv Pre-Print*.