Essays on Informational Efficiency in Real Estate Markets

Dissertation zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

vorgelegt von: Johannes Braun

Berichterstatter: Prof. Dr. Wolfgang Schäfers
Prof. Dr. Stephan Bone-Winkel

Tag der Disputation: 11.11.2020

Essays on Informational Efficiency in Real Estate Markets

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Background and Motivation

At the latest since the global financial crisis of 2007-08, the informational efficiency of real estate markets has received a remarkably high level of public attention. The bursting of the US housing bubble as the trigger of the crisis and the temporary downward pressure on all forms of real estate investments in the aftermath, raised various questions as to the degree of susceptibility to agency problems and the extent of sentiment-driven transactions in the markets. This thesis therefore seeks to shed light on two contrasting aspects related to informational efficiency in real estate markets: The first part deals with an example of the role played by noise in the market, the second part with an example of the withholding of information from the market.

With regard to the impact of noise on market efficiency, direct real estate markets represent a particularly interesting and important research field. Many shortcomings that prevail in the comparatively efficient stock market are likely to be aggravated in the context of physical properties. Short-sales required for arbitrage in phases of market exuberance are not feasible in practice, due to the diversity of the traded assets and the lack of exchanges. Transaction- and information costs in distinct, highly segmented property markets are furthermore significantly higher than those in stock markets, thereby further limiting the potential for arbitrage.

Over the course of the last decades, several studies analyzing the effects of noise-traders and their sentiment-driven transaction activity on direct property markets have been published. A crucial element of such studies lies in the definition of market sentiment. While traditional proxies mainly relied on surveys (see e.g. Clayton, Ling and Naranjo, 2009; or Das, Freybote and Marcato, 2015), sentiment analysis has shifted towards the use of text-based indicators in more recent years. The advantages compared to traditional sentiment measures include the high frequency of available data, the low data collection cost and -effort compared to survey-based measures, and the relative directness of the news-based measures, possible through a preselection of suitable news sources. Several more recent studies in the real estate field thus make use of news-based sentiment measures (see e.g. Walker, 2014; or Ruscheinsky, Lang and Schäfers, 2018). These studies have in common the application of predefined sentiment dictionaries, i.e. psychology-based wordlists, indicating sentiment-polarity on a word-to-word basis for the text in question. Despite their advantage of easy applicability, dictionaries inevitably rely on a set of predefined rules and thus lack the ability to identify complex linguistic patterns in texts. Hausler, Ruscheinsky and Lang (2018) take this limitation into account by applying a support-vector-machine (SVM) framework for the classification of news headlines. To train the SVM, Hausler, Ruscheinsky and Lang manually classify a training data set. This practice is common for machine-learning-applications, but is once again time-consuming and subjective with respect to the assignment of news to different sentiment classes.

The central objective of the first two papers of this thesis is therefore to develop a sentiment measure for the real estate market which overcomes the fundamental limitations of existing ones. The approach utilizes an artificial neural network to classify news articles on the US real estate market. The classification process however, refrains from any manual classification, but instead makes use of a large sample of texts with distinct sentiment-polarity, so as to train the network via distant supervision. The constructed sentiment index is employed to make predictions regarding the US direct commercial property market in terms of returns (Paper 1) as well in terms of trading liquidity (Paper 2). The chosen approach contributes to the literature in two ways. First, sentiment research gains from the development of a sentiment measure which fully capitalizes on the benefits of deep-learning, through the complete elimination of human intervention in the classification process. Furthermore, the approach provides new insights into several aspects of the efficiency of property markets by analyzing their susceptibility to noise trading and sentiment.

The second component of the thesis deals with the somewhat converse problem of information which is not available publicly. Possibly inspired by several high-profile accounting scandals at the beginning of the century, Jin and Myers (2006) presented a theory on the mechanisms behind such firm crises and the subsequent crashes in stock price. According to Jin and Myers, firm managers are able to absorb some of the negative information on the firm due to an information asymmetry between management and shareholders. The reasons for such bad news hoarding can be manifold, most obvious are perhaps attempts to secure the own job. However, when the accumulated bad news reaches a certain threshold, Jin and Myers argue that the incentives to continue the hoarding disappear, leading to a sudden release of a bulk of negative information into the market. The effect can be a crash in the firm’s stock price.

Listed real estate firms provide several peculiarities worth subjecting to an empirical test of Jin and Myers' theory. This thesis selects one of them, namely the opportunity for equity REITs to develop their own properties, and analyzes the relationship to stock price crash risk. The particular relevance of this aspect stems from the fact that property developments are one of only a few opportunities for REITs to circumvent the current lows in capitalization rates and thereby to maintain growth. To the best of the author's knowledge, the analysis is the first to examine the REIT business model with regard to its effect on the company's stock price crash risk. The analysis thus contributes to a better understanding of the risk structure of real estate development in the context of equity REITs, and by doing so, to answering the question of whether REIT property-development activity is desirable from an investor perspective.

1.2 Research Questions

The following section provides a short overview of the research questions relevant for each of the papers. While all three papers circle around the central topic of informational efficiency in real estate markets, each addresses several distinct aspects of the general problem field.

Research Questions, Paper 1:

"On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach"

- Are direct property market transaction prices influenced by sentiment-driven ('noise') traders?
- Can deep-learning algorithms contribute to the estimation of real estate market sentiment?
- Can distant supervision replace the manual classification of news-documents necessary for the construction of the deep-learning-based indicator?
- Can the deep-learning-based sentiment indicator explain direct property-market index movements?
- Does news-based sentiment exert a distinct impact on property index returns in up- and down-markets?

Research Questions, Paper 2:
"Artificial Intelligence, News Sentiment, and Property Market Liquidity"

- Do sentiment-driven ('noise') traders exercise a significant influence on direct property market liquidity?
- Can a news-based sentiment indicator predict movements in direct property market liquidity?
- Does a potential relationship hold for different dimensions of market liquidity (i.e. for market depth, -resilience and -breadth)?
- Does the slow nature of property transactions require a lagged model setup to adequately represent the relationship between sentiment and direct property market liquidity?

Research Questions, Paper 3:
"REIT Property Development and Stock Price Crash Risk"

- Are property-developing equity REITs more crash-prone than non-developing equity REITs?
- Do managers of equity REITs exploit the increase in information asymmetry due to property development for self-interested, myopic behavior?
- Can property development of equity REITs serve as a proxy for long-termism of management?
- Should regulators restrict equity REITs in their possibilities to develop own properties?

1.3 Co-Authors, Submissions and Conference Presentations

In the following, information on co-authors, journal submissions, publication status and conference presentations for each of the three papers is provided.

Publication Information, Paper 1:
"On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach"

Authors:

Jochen Hausler, Johannes Braun, Wolfgang Schäfers

Submission Details:

Journal:	Journal of Real Estate Research
Submission date:	08/08/2019
Current status:	Under review, temporarily declined to revise by authors

Conference Presentations:

This paper was presented at the 35th Annual Conference of the American Real Estate Society (ARES) in Paradise Valley, USA (2019), as well as the 24th Asian Real Estate Society (AsRES) International Conference in Shenzhen, China (2019).

Publication Information, Paper 2:
"Artificial Intelligence, News Sentiment, and Property Market Liquidity"

Authors:

Johannes Braun, Jochen Hausler, Wolfgang Schäfers

Submission Details:

Journal:	Journal of Property Investment & Finance
Submission date:	08/01/2019
Current status:	Accepted (11/06/2019)

Publication Information, Paper 3:

"REIT Property Development and Stock Price Crash Risk"

Author:

Johannes Braun

Submission Details:

Journal:	Journal of Real Estate Portfolio Management
Submission date:	07/07/2020
Current status:	Under review

Conference Presentations:

An early version of this paper was presented at the 34th Annual Conference of the American Real Estate Society (ARES) in Bonita Springs, USA (2018) under the preliminary title "The Effect of Liquidity on Stock Price Crash Risk: Evidence from the U.S. REIT Market".

1.4 References

Clayton, J., D. C. Ling and A. Naranjo, Commercial real estate valuation: Fundamentals versus investor sentiment, The Journal of Real Estate Finance and Economics, 2009, 38:1, 5–37.

Das, P. K., J. Freybote and G. Marcato, An investigation into sentiment-induced institutional trading behavior and asset pricing in the REIT market, The Journal of Real Estate Finance and Economics, 2015, 51:2, 160–89.

Hausler, J., J. Ruscheinsky and M. Lang, News-based sentiment analysis in real estate: A machine learning approach, Journal of Property Research, 2018, 35:4, 344–71.

Jin, L. and S. C. Myers, R^2 around the world: New theory and new tests, Journal of Financial Economics, 2006, 79:2, 257–92.

Ruscheinsky, J., M. Lang and W. Schäfers, Real estate media sentiment through textual analysis, Journal of Property Investment & Finance, 2018, 36:5, 410–28.

Walker, C. B., Housing booms and media coverage, Applied Economics, 2014, 46:32, 3954–67.

2 On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach

2.1 Abstract

This paper employs a deep-learning approach to text-based sentiment analysis with regard to the direct commercial real estate market in the United States. By means of an artificial neural network and distant supervision-labeled training data, a market sentiment indicator is derived from news articles and related to market returns, as well as to up- and down-market periods. The created monthly indicator Granger-causes market returns in a vector autoregressive framework during the study period from January 2006 to December 2018. Estimated Markov-switching models reveal a varying influence of the sentiment indicator on market returns in up- and down-market periods. Logit regressions furthermore indicate some forecasting potential in a binary return prediction framework. However, while large market swings are captured well, the indicator struggles with short-term return fluctuations. Through the discussion of the extraction procedure, the potential and also the shortcomings of the sentiment-measuring approach, this paper lays the foundations for further applications of the constructed sentiment indicator.

Keywords: Artificial Neural Network, Deep Learning, Text-based Sentiment, Commercial Real Estate

2.2 Introduction

Compared to other areas of research, artificial intelligence (AI) has not so far gained much attention in the field of real estate. Only a few scholars (e.g. Din, Hoesli and Bender, 2001; and Peterson and Flanagan, 2009) address in their studies the potential of “intelligent agents” such as artificial neural networks (ANNs). Arguably, in particular the sparse data availability compared to other industries, has contributed to the fact that artificial intelligence research for real estate has not yet been able to extend beyond the fledging stage.

However, three rather recent developments have changed the setting and should be able to assist AI in becoming a powerful research instrument: The broad availability of vast amounts of online data through social networks or crowd-sourced information platforms has laid the basis for the data-hungry concepts of machine- and in particular deep-learning. This is aided by a drastic increase in computational power available to researchers through GPU (Graphics Processing Unit) and IaaS (Infrastructure as a Service) computing. Additionally, AI research has overcome several theoretical bottlenecks by developing new and better algorithms.

Due to this evolution, a new field of sentiment analysis, which surpasses the more traditional concepts of survey-based estimates and market proxies such as mortgage fund flows, has become accessible. For the first time, machines can be trained to assess and extract not only the content, but also opinions from textual documents via what is referred to as opinion mining. The research in this context started with sentiment dictionaries and proceeded to sentiment engines, such as Thomson Reuters News Scope (see e.g. Groß-Klußmann and Hautsch, 2011) and more recently, machine-learning approaches. However, to the best of the authors’ knowledge, no research in real estate so far has addressed the most recent subfield of sentiment analysis, namely ANN-based deep-learning. Through better scalability and the possibility of real-time analysis, which consequentially leads to an advantage in ‘big data’ applications, and the ability to identify more complex relationships by analyzing a richer data structure compared to other machine-learning approaches, artificial neural networks may have the potential to surpass other sentiment indicators when a large quantity of good quality training data is available. The bottleneck of traditional deep-learning-based textual sentiment analysis lies in the provision of a sufficient amount of manually sentiment-labeled text documents.[1] This paper is therefore not only the first to test a deep-learning framework for text-based sentiment analysis in real estate, but also seeks to overcome the aforementioned labeled data shortage by utilizing a new source of distant-labeled sentimental text data, namely *Seeking Alpha* long and short idea sections. Because of the slow pace of real estate transactions, the heterogeneity of real assets, as well as non-transparent regional markets, assessing the potential of a scalable sentiment indicator, which is also adaptable to local circumstances through the use of regionally published news articles as training data, seems especially worthwhile.[2]

After looking into the sentiment extraction procedure, the qualities of the resulting sentiment indicator are subject to critical scrutiny in a vector autoregression (VAR), a Markov-switching (MS) and a logit framework. The vector autoregression serves as a starting point, in order to shed light on the question of whether the indicator is able to explain direct real estate market returns. Beyond that, the VAR model can help to clarify the pressing question of causality.[3] Despite the advantages of VAR models, they imply the possibility of ignoring a potential non-linear relationship between the variables in question. In particular for the REIT market, past research has provided resilient evidence that in order to reflect bull and bear markets, the use of Markov-switching models is preferable (see e.g. Lizieri, Satchell, Worzala and Dacco, 1998; Bianchi and Guidolin, 2014). The cyclical nature of direct real estate markets suggests the need to control for the possibility of differing regimes likewise in their specific context. Freybote and Seagraves (2018) suggest a Markov-switching model in their paper on the relationship between sentiment and direct real estate market liquidity, and find strong differences in the relationship for both up- and down-markets. In order to evaluate the possibility of a non-linear relationship between sentiment and returns, this paper applies a Markov-switching model as the second component of its econometric analysis section. In the final econometric section, the paper considers aspects with relevance for the real estate industry. Within a logit framework, the ability of the sentiment measure to forecast up- and down-market periods is investigated. In- and out-of-sample forecasts are performed for this purpose. Besides being required in

[1] To gradually improve a deep-learning algorithm's capabilities, permanent human intervention is required.

[2] A publication assessing a potential link of the constructed sentiment indicator to direct real estate market liquidity is intended by the authors.

[3] Both a case for a causal relationship of sentiment explaining returns, as well as a converse relationship can be made. By the use of Granger-causality tests within a VAR model, this potential issue can be untangled.

terms of econometric diligence, this threefold approach is expected to help identify possible room for improvement in the construction procedure of the sentiment measure, which might allow for the creation for more comprehensive measures in future research.

The paper proceeds as follows: In Section 2.3, research with respect to text-based sentiment in finance and real estate is re-considered as an introduction to the more theory-driven Sections 2.4 and 2.5. These sections depict the structure of the news corpus from the *S&P Global Market Intelligence* database, as well as the training data from *Seeking Alpha,* before showing the sentiment extraction process via ANN and the econometric approaches. Subsequently, Section 2.6 presents the results of the VAR, Markov-switching and logit procedure. Section 2.7 discusses implications and provides concluding remarks.

2.3 Literature Review

2.3.1 Text-Based Sentiment Analysis in Finance

As demonstrated by Loughran and McDonald (2016), textual analysis and parsing in various forms has a history spanning several centuries. Likewise, analyzing the influence of news on stocks or entire markets in the finance literature is by no means a recent development. Starting more than 30 years ago, Roll (1988) made use of news from the *Wall Street Journal* and the *Dow-Jones Newswire* to explain stock price changes in his seminal R^2 paper. Other early studies such as Cutler, Poterba and Summers (1989) and French and Roll (1986) treated news as a mere measure of incoming information, without explicitly analyzing the content of the documents themselves. More recently, with the increase of computational power and driven by the requirements of internet search engines, as well as the rapid growth of social media, natural language processing and especially the subcategories of sentiment analysis and opinion mining have become an increasingly active research area, extending from computer science to the social and management sciences (Liu, 2012). Accordingly, the finance literature has recently been accommodating an ever-growing body of textual sentiment studies.

Kearney and Liu (2014) provide a comprehensive survey on how textual sentiment impacts on firm- and market level performance, sorted by methods and information sources. Most studies in that context focus on the sentiment analysis of news articles and seek to link the constructed sentiment proxies to stock market returns, market prices, trading volumes, volatility, bid-ask spreads as well as firm earnings (see e.g. Groß-Klußmann and Hautsch, 2011; Ozik and Sadka, 2012; Engelberg, Reed and Ringgenberg, 2012; García, 2013; Boudoukh, Feldman, Kogan and Richardson, 2013; Ferguson, Philip, Lam and Guo, 2015; Heston and Sinha, 2016; Sinha, 2016; Sun, Najand and Shen, 2016; Hanna, Turner and Walker, 2017; as well as the seminal articles by Tetlock, 2007; and Tetlock, Saar-Tsechansky and Macskassy, 2008). Another stream of literature addresses the influence of earnings press releases on a broad variety of performance measures (see e.g. Henry, 2008; Davis, Piger and Sedor, 2012; Davis and Tama-Sweet, 2012; Huang, Teoh and Zhang, 2014; Henry and Leone, 2016) and annual reports (see e.g. Kothari, Li and Short, 2009; Li, 2010; Feldman, Govindaraj, Livnat and Segal, 2010; Loughran and McDonald, 2011, 2015; Jegadeesh and Wu, 2013).

The vast majority of those studies either uses a sentiment dictionary such as the *General Inquirer (GI)/Harvard IV-4* for classification purposes or an adapted finance-specific word list. Only a small fraction of papers facilitates text analysis programs (see e.g. Davis, Piger and Sedor, 2012; Davis and Tama-Sweet, 2012; Henry and Leone, 2016; for an application of the program *DICTION*). Basic machine-learning techniques and classification algorithms such as Naïve Bayes and support-vector machines are seldom applied, and more common in literature referring to the inherent sentiment expressed in stock message boards (see e.g. Antweiler and Frank, 2004; and Das and Chen, 2007). However, there are some initial attempts at more advanced deep-learning methods such as artificial neural networks (ANN) in the recent literature. For example Smales (2014), as well as Borovkova and Dijkstra (2018), rely on ANNs as well as news analytics from *Thomson Reuters* and its respective newswire, to investigate the relationship with gold future returns as well as to provide intraday forecasts on the *EUROSTOXX 50*.

2.3.2 Sentiment Analysis in the Realm of Real Estate

Sentiment analysis in real estate research relies predominantly on other, non-text-based, sentiment indicators, although being well established and drawing on an extensive range of resources. Sentiment gauges extend from market-related sentiment proxies such as NAV discounts (see e.g. Barkham and Ward, 1999, for an early study of NAV discounts of property companies in the UK, as well as Lin, Rahman and Yung, 2009, for an analysis of the influence on investor sentiment and REIT returns) to mortgage fund flows, properties sold from the *NCREIF Property Index* (NPI), the ratio of transaction-based (TBI) and constant-liquidity-based versions of the NPI value index, as well as past NPI and TBI total returns (Clayton, Ling and Naranjo, 2009). Freybote and Seagraves (2017) adopt buy-sell imbalances when examining whether multi-asset institutional investors rely on the sentiment of real-estate-specific investors for investment decision making. In addition to such so-called "indirect" measures, surveys – especially the *Real Estate Research Corporation (RERC)* survey – are frequently used as a direct indicator, when linking investor sentiment to commercial real estate valuation (Clayton, Ling and Naranjo, 2009), private market returns (Ling, Naranjo and Scheick, 2014), trading behavior (Das, Freybote and Marcato, 2015) and REIT bond pricing (Freybote, 2016). For residential real estate sentiment, Marcato and Nanda (2016) use the *National Association of Home Builders (NAHB)* and *Wells Fargo* index and evaluate their ability to forecast demand and supply activities.

Furthermore, following a pioneering article by Ginsberg, Mohebbi, Patel, Brammer, Smolinski and Brilliant (2009), several scholars drew on *Google* search query data to analyze various aspects of the real estate market in the United States. Hohenstatt, Käsbauer and Schäfers (2011) provide evidence that *Google Trends*[4] enables inferences on the housing market in the near future, as well as on financing decisions. Similarly, there is evidence that abnormal search activity in US cities can help to predict future abnormal house price changes (Beracha and Wintoki, 2013) and *Google Trends* can serve as an indicator for housing market turning points (Dietzel, 2016). With respect to the commercial real estate market, the results were likewise promising. Dietzel, Braun and Schäfers (2014), Rochdi and Dietzel (2015) as well as Braun (2016) demonstrate the ability of *Google Trends* data to forecast commercial real estate transaction and price indices, REIT market volatility, as well as to serve as a successful application in trading strategies.

Besides such indirect proxies, surveys and search query data, some text-based indicators have found their way into real estate research in more recent years. Facilitating news articles, Soo (2015) uses sentiment expressed in local newspapers to predict house prices in 34 US cities. Walker (2014, 2016) makes use of the

[4] *Google Trends* provides search volume indices of search queries that can be filtered by various different categories, according to the topic of interest.

aforementioned *DICTION* software to investigate the relationship between the UK housing market boom from 1993 to 2008, and media coverage as well as stock returns of firms engaging in the housing market. Analyzing news headlines from *Bloomberg*, *The Financial Times* and *The Wall Street Journal*, Ruscheinsky, Lang and Schäfers (2018) reveal a leading relationship of media-expressed sentiment to the *FTSE/NAREIT All Equity Total Return Index*. With respect to machine-learning and deep-learning, so far, the only available research is apparently provided by Hausler, Ruscheinsky and Lang (2018), in which the authors show that sentiment indicators extracted by means of machine-learning lead the direct as well as the securitized real estate market in the United States. It seems that no research is published exploring the power of deep-learning in general, and artificial neural networks (ANN) in particular in a real estate market context.

Considering the drawbacks of alternative sentiment indicators (i.e. a long reaction time and, in the case of market surveys, a restricted availability), this research gap provides a unique opportunity to explore the potential of a deep-learning approach with respect to text-based sentiment analysis in real estate. Simultaneously, the disadvantages of abstract, theory-loaded proxies are avoided, as deep-learning frameworks do not rely on pre-defined theoretical rules, but independently "master" potential relationships from provided training data. Accordingly, with the help of distant supervision-labeled training documents from *Seeking Alpha,* as well as news articles from the *S&P Market Intelligence Database*, the application of an ANN sentiment classifier for predicting returns and turning points in the *CoStar Commercial Repeat-Sale Index* is assessed. Hence, the present paper is the first to combine text-based sentiment analysis, a deep-learning approach and distant supervision labeling in real estate research.

2.4 Data

The outlined study utilizes four types of data: *Seeking Alpha*[5] (SA) long and short idea sections (as explained below) serve as the training data set for the artificial neural network, and *S&P Global Market Intelligence* (S&P) real estate news articles on the US market constitute the text corpus of the constructed sentiment index. The *CoStar*

[5] *Seeking Alpha* is a crowd-sourced website providing investment content delivered by independent contributing authors. The required long and short ideas are subsections of the SA website, containing opinions on either single financial assets or asset markets in general. In each long idea, an author outlines why he expects the asset or market in question to be a favourable buying opportunity, and conversely for short ideas. Since 2014, long and short idea articles have started with a summary section that delivers the quintessence of the buy or sell recommendation in several short bullet points.

Commercial Repeat-Sale Index (CCRSI) is used as a measure of development of the direct real estate market in the United States. Furthermore, a set of control variables will be added to the regression equations. The time series limiting factor is the S&P news database, which only provides articles back to November 2005. The empirical models thus incorporate data from January 2006 to December 2018.

2.4.1 Seeking Alpha

For the construction of the sentiment index, a two-part process is proposed. As this paper refrains from manually labeling training data for the ANN, a data set of distant supervision-labeled text documents[6] is required. Summary sections of *Seeking Alpha* long and short ideas are collected for this objective. The following example from the data set illustrates the structure of those summary sections for a short idea:

> *"Consumer complaints are everywhere. Particularly concerning are those surrounding false billing and unwillingness to share work invoices. [...]"*

The summary sections of those investment ideas either contain a distilled version of negative sentiment (short ideas) or positive sentiment (long ideas) towards the equity or market in question. It thus can be argued that SA long and short ideas represent an almost ideal data set for training an ANN on the distinction between positive and negative sentiment.

In total 69,773 investment ideas were collected from SA. With only 8,911 of the summaries being short ideas, the ratio is heavily skewed. In order to receive a symmetric training procedure, a random sample of 8,911 long ideas is drawn and joined with the short ideas to constitute the ANN's training data set. The final training data set consists of a balanced sample of 17,822 SA texts provided by 3,107 different authors and containing an average of 381 characters.

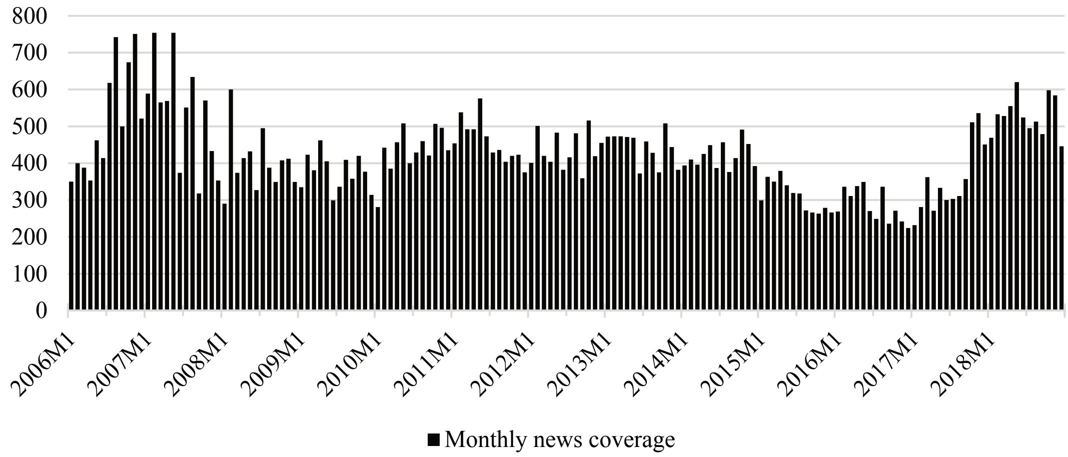
2.4.2 S&P News Database

For the second step in the process of constructing the sentiment index, real estate market news articles are required. Due to their widespread availability among real estate professionals, articles from the *Standard & Poor's Global Market Intelligence* news database with respect to the US real estate market are collected. These articles serve as the basis for estimating of the level of the monthly sentiment index. The total

[6] Distant supervision labeling is defined as the absence of an annotator providing the classification of the training data manually.

number of news articles for the study period between January 2006 and December 2018 is 66,070, with a monthly mean of 424 articles, a minimum number of 224 articles per month and an average of 1,125 characters per article (see Figure 2.1).

Figure 2.1: S&P News Distribution over Study Period



Notes: Figure 2.1 plots the monthly distribution of the 66,070 news articles serving as the basis for constructing of the sentiment index in this study. The articles in the sample were collected from the *S&P Global Market Intelligence* news archive, covering the US real estate market between 2006:M1 and 2018:M12. The monthly mean of news articles per month is 424, and the minimum, 224 articles per month.

2.4.3 Direct Market Return and Macroeconomic Controls

The dependent variable of the regression analysis is the *CoStar Commercial Repeat-Sale Index* (CCRSI) which represents the development of commercial real estate prices in the United States. For this study, monthly percentage changes in the value-weighted US composite price index are used. When running regression analyses for real estate returns, other influencing factors such as the general economy as well as capital markets, have to be taken into account. All control variables are selected in accordance with previous research, mainly Clayton, Ling and Naranjo (2009), Ling, Naranjo and Scheick (2014) and Hausler, Ruscheinsky and Lang (2018). At the capital market level, this study includes credit spread, term structure and general stock market return variables. This allows controlling for the state of debt, as well as equity markets and financing conditions (see e.g. Freybote and Seagraves, 2017). More specifically, future expectations of overall economic development are controlled for by incorporating a term structure variable (TERM, i.e. the spread between 10-year Treasury bonds and 3-month Treasury bill yields). Furthermore, the spread between Moody’s seasoned Baa- and Aaa-rated corporate bond yields is added to the regression equations (SPREAD) in order to control for general economic default risk (see e.g. Clayton, Ling and

Naranjo, 2009). Following Das, Freybote and Marcato (2015), the performance of the general stock market is accounted for by including monthly returns on the *S&P500* composite index (S&P500). To additionally allow for the fact that direct real estate is considered as an inflation hedge (Hoesli, Lizieri and MacGregor, 2008), consumer price index growth is used to control for inflation (INFLATION). Altogether, those variables should also capture the overall demand for real assets. The current state of the supply side however, is reflected by adding percentage changes in seasonally adjusted total construction spending (CONSTRUCTION) on a monthly basis. Summary statistics of the described variables can be obtained from Table 2.1.

Table 2.1: Descriptive Statistics

Statistic	Mean	Median	St. Dev.	Min	Max
CCRSI (%)	0.26	0.46	1.53	-6.82	3.05
TERM (pp)	1.83	1.95	1.05	-0.52	3.69
SPREAD (pp)	1.10	0.94	0.50	0.55	3.38
S&P500 (%)	0.71	1.29	4.10	-16.80	10.93
INFLATION (%)	0.16	0.17	0.39	-1.92	1.01
CONSTRUCTION	86,536	88,709	14,038	62,893	110,362

Notes: This table reports summary statistics of the monthly real estate return data and macroeconomic time series. CCRSI is the total return of the *CoStar Commercial Repeat-Sale Index*. TERM is the difference between the 10-year US Treasury bond and the 3-month Treasury bill yields in percentage points (pp). SPREAD is the difference between Baa- and Aaa-rated corporate bond yields. S&P500 is the total return of the *S&P 500 composite index*. INFLATION is the percentage change of the consumer price index. CONSTRUCTION is the amount of seasonal adjusted construction spending in millions of dollars. The sample period is 2006:M1–2018:M12.

2.5 Methodology

2.5.1 Artificial Neural Network

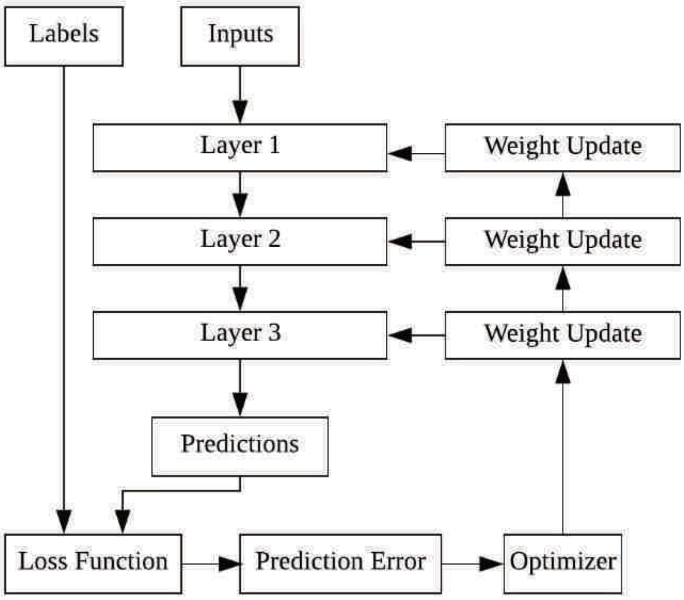
Artificial neural network research, often falsely perceived as a young field, actually emerged as early as the 1950s, with Rosenblatt (1958) often being considered the inventor of the first "real" ANN. Due to the extensive computational requirements and lack of mathematical algorithms to back the concepts, research effort in the field stagnated soon after. With the introduction of the backpropagation algorithm in the context of ANNs, Werbos (1974) drastically increased the possibilities for training complex models efficiently. The newly-wakened research interest was however, again retarded by the breakthroughs in the related machine-learning field of support vector machines (SVMs) in the early 1990s (see Cortes and Vapnik, 1995). As "shallow" learning methods however, SVMs require the application of feature engineering, which regularly renders them inferior to ANNs in solving perceptual problems.

Furthermore, in comparison to ANNs, practical applications of SVM approaches turned out to be less scalable in conjunction with large data sets. The widespread availability of massive amounts of data accompanying the rise of the internet, new algorithms as well as a drastic increase in computational power on hand, have all contributed to a resurgence of ANN research and applications in recent years. Hence, a recent milestone in ANN development is commonly seen in the development of "*AlexNet"* (Krizhevsky, Sutskever and Hinton, 2012), which won the widely recognized *ImageNet* picture classification task in 2012 and heralded a period of dominance of ANN methods in the *ImageNet* and similar competitions since then.

Despite developments in the theoretical foundations of ANN research, the field rests on relatively little mathematical theory. ANN development can thus rather be seen as an engineering than a statistical discipline; models are regularly justified empirically instead of theoretically. The intuitive but simplistic analogy to human brains lending artificial neural networks their name, results from their shape, which combines consecutive layers of interconnected "neurons" (or nodes). Comparable to the human brain, the involved neurons require a certain signal threshold to fire and deliver a transformed signal to the subsequent layer. By directing an input signal through the layers, stepwise transformations of the input signal are performed.[7] The goal of the transformation process executed by the network layers is the minimization of prediction errors, i.e. the "distance" between the network's predictions and the assigned labels defined by the network's loss function. Error reduction is achieved by the gradual alteration of the weight parameters defining the functions of each layer's nodes. Simultaneous optimization of the parameter values is achieved through the application of a backpropagation algorithm. By using backpropagation, the gradient function of the chained derivatives for all network nodes is calculated and thereby also the direction in which the parameter values have to be changed in order to reduce the overall prediction error. The general structure of an ANN is shown in Figure 2.2.

[7] In the context of text sentiment analysis, the input data consists of vectorized text data assigned with sentiment labels.

Figure 2.2: Basic Structure of an Artificial Neural Network



Notes: Figure 2.2 shows the basic circular structure of an artificial neural network (ANN). Training data is channeled through a sequence of transformations. A loss function evaluates the predictions by comparing them to true data labels. Subsequently the predictions are optimized by performing updates of the weight parameters in each layer. Then the process is repeated with the updated weight parameters.

2.5.1.1 Text Pre-Processing

To obtain vectorized, machine-readable text data, several pre-processing steps on the raw *Seeking Alpha* and *Standard & Poor's* text data have to be undertaken. Firstly, Unicode categories P, S, Z and C, as well as separate numbers are removed, and upper case replaced by lower case letters.[8] Intra-word contractions and hyphens are split up into the respective single words, possessive forms of words converted into their regular equivalents (e.g.: "company's" is transformed into "company"). Additionally, the texts are compared to a stopword list to remove words with presumably no or very low sentiment polarity. For this paper, written forms of numbers and any form of calendar terminology are included in the stopword list. These additions to the standard list are performed to remove uninformative patterns related to expressions of time in the SA text data, as these patterns might otherwise be incorporated into the ANN's learning algorithm in the upcoming steps.

Furthermore, an analysis of the structure of both text sources exhibits a considerable number of company names, executive names and similar terms. These terms

[8] Unicode categories P, S, Z and C contain punctuation, symbols, separators and control characters respectively.

presumably do not carry any sentiment polarity themselves. However, due to the structure of SA's long and short ideas, an unintentional influence of such terms on the sentiment prediction of the ANN has to be considered.[9] For this reason, both S&P and SA text data has to be aligned to a dictionary containing a complete set of English vocabulary used in written language. Thus, each text is compared to the broadly used *Hunspell* spell checking dictionary.[10] By doing so, words that are not part of the general English language corpus (i.e. most company names or names) are removed from the text documents. As a final pre-processing step, all words contained in the SA and S&P texts are reduced to their word stem form.

2.5.1.2 ANN Training and Validation

Next, each SA long and short idea is annotated with the distant supervision label (i.e. long ideas are annotated with 1, short ideas with 0). To reduce noise in the ANN's learning process and limit computational requirements, the word universe for all SA texts is restricted to the 1,000 most frequent words.

For the validation of the network after the training process, 20 percent of the SA data is selected at random and excluded from training. The remaining 80 percent of the pre-processed SA data (i.e. 14,258 texts) is supplied to the ANN. This is done with the use of a document feature matrix.[11]

The ANN is set up as a multilayer perceptron with the following structure: 4 fully connected layers with ReLU (Rectified Linear Unit) activation functions and declining node amounts (64, 48, 32, and 16) are used to gradually reduce the feature space. The ReLU layers are defined by the transformation:[12]

$$max(0, dot(Input, W) + b)\,. \tag{2.1}$$

[9] Suppose, for example, a high amount of SA long ideas on *Equinix* REIT. The ANN will inevitably connect the term "*Equinix*" to positive sentiment, if this issue remains unaccounted for.

[10] *Hunspell* word lists are available under http://app.aspell.net/create for downloading. For this paper, a list containing the common spelling of the *Hunspell* default number of words, including American and British English spelling, is used. Variants with and without diacritic marks of respective words are included.

[11] A document feature matrix, also referred to as a sparse matrix, contains a column for each word in the respective data set and a row for each text document in the data set. Each cell of the matrix is filled with 1, if the text document in question contains the respective word, and 0 otherwise. Note that several specifications containing the use of embedding layers, together with an integer matrix, were tested. However, as the classification results did not change drastically, the more intuitive concept of a document feature matrix was given the preference in this paper.

[12] For clarity, the subscripts of the weight parameters W and b are not included in the equations describing the layout of the ANN.

Input constitutes the input matrix resulting from the vectorized text documents for the first ReLU layer and the output of the preceding layer for layers 2 to 4. W and b are the weight parameters.

A final layer of the ANN is constituted by a sigmoid squashing function, so as to obtain a one-dimensional output parameter between 0 and 1:

$$\frac{1}{1+e^{-t}} \quad with\ t = dot(Input, W) + b\ . \tag{2.2}$$

Here, *Input* denotes the output of the last ReLU layer, W and b are again weight parameters. During the training process, the pre-processed SA data is fed into the ANN (starting initially with random weight parameters) in batches of 500 articles with a gradient update following each new batch. In total, 6 epochs, each containing all batches, are performed.[13] The optimization process thus contains a total of 174 gradient updates.[14]

The loss score after each batch is calculated by applying a binary cross-entropy loss function:

$$\frac{1}{n}\sum_{k=1}^{n} -1 * \left(y_k * log(p_k) + (1-y_k) * log(1-p_k)\right). \tag{2.3}$$

y_k is a binary variable taking the value 1 if *Seeking Alpha* text k is labeled as a long idea, and 0 if *Seeking Alpha* text k is labeled as a short idea. p_k is the probability value resulting from the sigmoid function for text k.

The optimization of the ANN is executed by using the *Root Mean Square Propagation* (RMSprop) algorithm (Tieleman and Hinton, 2012).[15] The updates for all parameters W and b are calculated with the following equations:

$$\begin{aligned} v_{dW_t} &= \beta * v_{dW_{t-1}} + (1-\beta) * {dW_t}^2 \\ v_{db_t} &= \beta * v_{db_{t-1}} + (1-\beta) * {db_t}^2 \end{aligned} \tag{2.4}$$

[13] Other specifications were tried, but a lower number of texts per batch did not increase the predictive power. A higher number of epochs lead to a gradual overtraining of the ANN.
[14] Updates per epoch: 29 (≈14,258/500); Updates over all epochs: 174 (=29*6).
[15] RMSprop, first suggested by Geoffrey Hinton during a Coursera online class in 2012, developed into one of the most frequently used ANN optimization algorithms. However, it was never formally published.

$$W_{t+1} = W_t - \frac{\eta}{\sqrt{v_{dW_t} + \varepsilon}} * dW_t$$

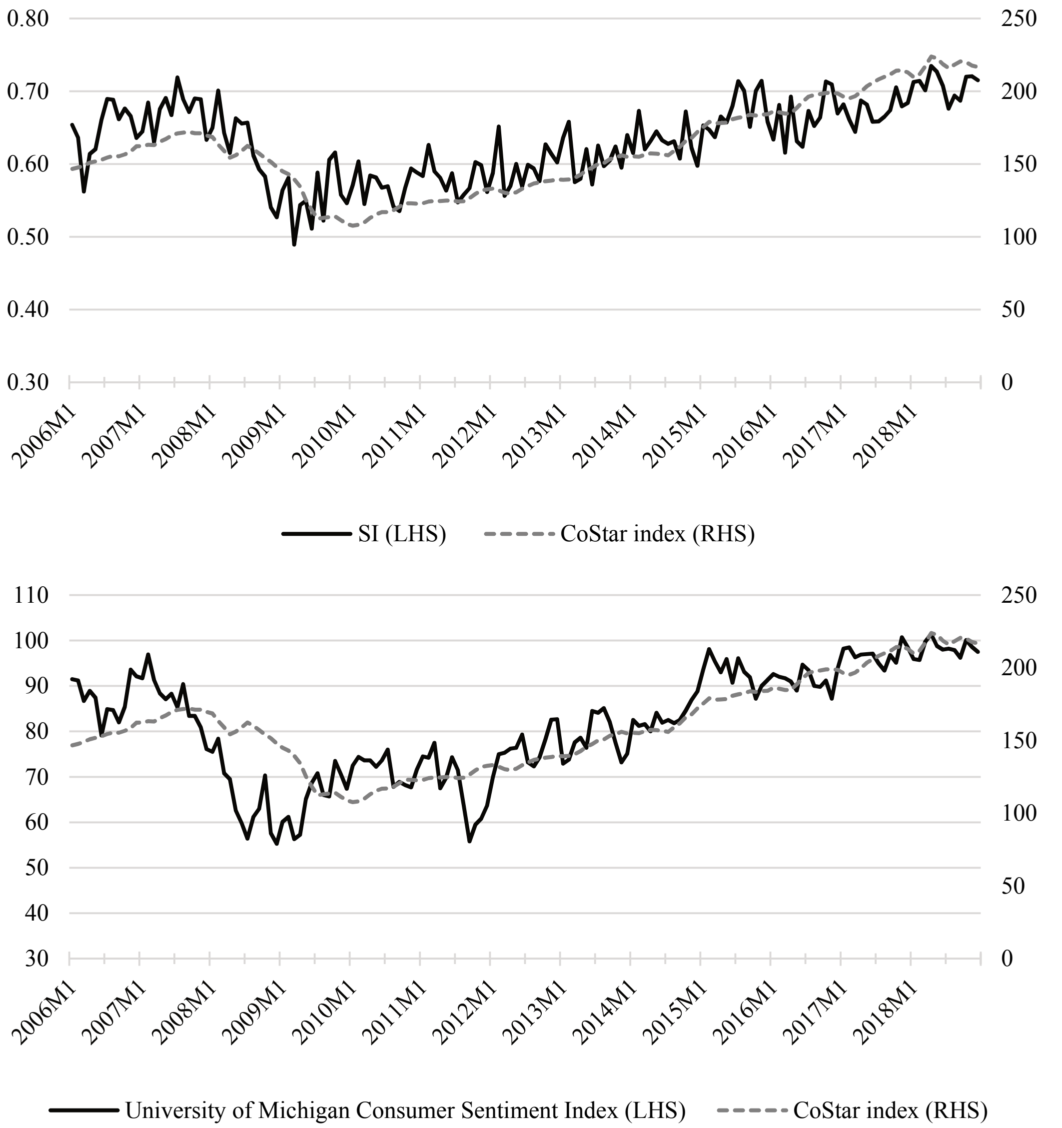
$$b_{t+1} = b_t - \frac{\eta}{\sqrt{v_{db_t} + \varepsilon}} * db_t \ .$$

dW_t and db_t are the gradients of the weight parameters at time *t*, $v_{dW_{t-1}}$ is the moving average of the squared gradient for weight parameter W at time *t-1*, $v_{db_{t-1}}$ the equivalent for weight parameter b at time *t-1*. β is a hyperparameter constituting the computation of the gradients' moving average. For β, Hinton's (for details see Tieleman and Hinton, 2012) initially suggested value of 0.9 is used. η defines the learning rate of the optimizer, for this paper η is set to 0.001. The hyperparameter ε constitutes a fuzz factor to avoid division by zero, in this paper the value of e^{-7} is chosen.

The training process described above is used to train 10 ANN models, in order to increase the robustness of the predictions. The average prediction value for each S&P news article is used to calculate its sentiment score. The monthly sentiment index value is then computed as the average sentiment score of all S&P news articles of the respective month. Due to the application of the sigmoid function in the last ANN layer, the sentiment index (SI) ranges between 0 and 1 in the spectrum and can thus be interpreted as a probability value. In the regression analyses, first differences of the monthly sentiment index score are used.

SI yields a mean value of 0.63 and a standard deviation of 0.05. This matches the average positive market performance of the CCRSI of 0.26% during the sample period. To provide some initial visual results, Figure 2.3 contrasts the SI with the CCRSI, as well as the *University of Michigan Consumer Sentiment Index* (MCSI). To justify the general concept of the sentiment index suggested in this paper, the SI should not differ vastly from existing sentiment measures over the study period. Indeed, MCSI and SI exhibit an index correlation of 73.00%. The index correlations with the direct market are 78.23% and 79.80% for the MCSI and the SI, respectively. Those findings are encouraging with respect to possible results of more in-depth econometric analyses in the future.

Figure 2.3: Temporal Progression of the SI



Notes: The top chart in Figure 2.3 contrasts the temporal progress of the created ANN-based textual sentiment indicator (SI) with the progress of the *CoStar Commercial Repeat Sales* value-weighted index. For a comparison, the bottom picture in Figure 2.3 repeats the same lineup for the *University of Michigan Consumer Sentiment Index* (MCSI). The sample period is always 2006:M1 to 2018:M12.

2.5.2 Econometric Approaches

To examine the full potential of the ANN-based sentiment indicator, three different econometric models are tested. This extensive econometric framework aims to shed light on the indicator's capability to predict both turning points, as well as market returns. With respect to a potential relationship between the proposed sentiment indicator and returns on the direct real estate market in the United States, a vector autoregression as well as a Markov-switching model are implemented. A logit approach further explores the indicator's predictive potential for up- and down-market phases within a binary response model framework. Additionally, in-sample and one-step-ahead out-of-sample forecasts with continuously updated estimations are calculated for the logit model. This combination of econometric models may seem excessive. However, the paper seeks to test the robustness of the influence of the proposed sentiment on the real estate market and find potential improvement opportunities for the chosen sentiment estimation procedure. The comparison of different models thus seems promising for that purpose.

2.5.2.1 Vector Autoregression

To model the relationship between the proposed sentiment indicator SI and CCRSI returns, a VAR framework is deployed in a first step. Because news on real estate markets and therefore arguably also sentiment measures extracted from such news are dynamically and potentially bi-directionally related to market performance, VAR is a reasonable choice, as no a priori causality assumptions are required.

Accordingly, a bivariate framework with two regression equations and two endogenous variables $y_{1,t}$ and $y_{2,t}$ is adopted (i.e. CCRSI returns as well as first differences of the sentiment indicator). Both variables are expressed as linear functions of their own lagged values, the lagged values of additional regression variables, as well as an error term:

$$\begin{aligned} y_{1,t} = \alpha_{1,0} + \alpha_{1,1}\, y_{1,t-1} + \cdots + \alpha_{1,k}\, y_{1,t-k} + \alpha_{1,1}\, y_{2,t-1} + \cdots \\ + \alpha_{1,k}\, y_{2,t-k} + u_{1,t} \\ y_{2,t} = \alpha_{2,0} + \alpha_{2,1}\, y_{2,t-1} + \cdots + \alpha_{2,k}\, y_{2,t-k} + \alpha_{2,1}\, y_{1,t-1} + \cdots \\ + \alpha_{2,k}\, y_{1,t-k} + u_{2,t}\ . \end{aligned} \tag{2.5}$$

$u_{i,t}$ denotes a white noise error term with $E(u_{i,t}) = 0$, $(i = 1,2)$, $E(u_{1,t},\ u_{2,t}) = 0$ and k denotes the number of lags. The model's optimal lag length is determined from a set of information criteria: Akaike (AIC), Schwarz (BIC) as well as Hannan-Quinn (HQ). The model displaying the lowest value for two of the three criteria is selected. Whenever results were ambiguous, as the most rigorous criterion, HQ guided the lag-length selection.

Both equations of (2.5) are eventually adjusted by including a combined set of additional exogenous controls $\boldsymbol{z_t}$ with coefficient matrix $\boldsymbol{B}$.[16] This leads to the widely

[16] Bold characters denote matrices.

used standard-form VAR which can be estimated using ordinary least squares (OLS):

$$\boldsymbol{y_t} = \boldsymbol{A_0} + \boldsymbol{A_1}\boldsymbol{y_{t-1}} + \cdots + \boldsymbol{A_k}\boldsymbol{y_{t-k}} + \boldsymbol{B}\boldsymbol{z_t} + \boldsymbol{u_t}\,. \tag{2.6}$$

Furthermore, a set of diagnostic tests was performed in order to ensure robustness of the results. All explanatory time series are analyzed for the existence of unit roots by means of an Augmented Dickey-Fuller Test (ADF). In all cases, first differences or growth rates are used. A Breusch-Godfrey Lagrange Multiplier further ensures that residuals are not serially correlated. In addition, normality and heteroscedasticity tests were conducted to ascertain statistical appropriateness.

2.5.2.2 Markov-Switching

Switching models are based on the assumption that a variable of interest y_t (i.e. CCRSI returns) follows a process that is dependent on an unobserved state variable s_t. This study assumes two distinct market regimes, corresponding to periods of either positive or negative market returns. The market is assumed to be in state m at period *t* when $s_t = m$ $(m = 1{,}2)$. Given a row vector of regressors $\boldsymbol{x_t}$, the conditional mean of regressand y_t in regime *m* shall be linear, i.e. $\mu_t(m) = \boldsymbol{x_t}\boldsymbol{\beta_m}$ where $\boldsymbol{\beta_m}$ is a column vector of coefficients (indexed by regime). Further assuming that regression errors are normally distributed (ϵ_t is *iid), y_t* is specified by the following model:

$$y_t = \mu_t(m) + \sigma(m)\epsilon_t = \boldsymbol{x_t}\boldsymbol{\beta_m} + \sigma(m)\epsilon_t\,.^{17} \tag{2.7}$$

In the special case of a Markov-switching model with only two regimes, as introduced by Hamilton (1989), s_t follows a first order Markov chain with the following transition matrix, where element *ij* shows the (time-invariant) probability of switching from regime *i* in period *t-1* to regime *j* in period *t*:

$$p = \begin{bmatrix} P(s_t = 1|s_{t-1} = 1) & P(s_t = 2|s_{t-1} = 1) \\ P(s_t = 1|s_{t-1} = 2) & P(s_t = 2|s_{t-1} = 2) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}. \tag{2.8}$$

By using the one-step-ahead probabilities of being in regime *m* as the weights of the density function in each regime, the likelihood contribution of a given observation y_t is received:

[17] Note that the standard deviation may or may not be regime-specific $\sigma(m) = \sigma_m$.

$$L_t(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}) = \sum_{m=1}^{2} \frac{1}{\sigma_m} \phi\left(\frac{y_t - \mu_t(m)}{\sigma_m}\right) * P(s_t = m | \Im_{t-1}, \boldsymbol{\delta}) \,, \qquad (2.9)$$

where $\boldsymbol{\delta}$ are parameters determining the regime probabilities (i.e. determining the elements of the transition matrix), $\boldsymbol{\sigma}$ is the standard deviation of all regimes and $\Im_{t-1}$ the information set available at period *t-1*. Thus, the full log-likelihood for all time periods *T* is given by Equation (2.10):

$$l(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}) = \sum_{t=1}^{T} log \left\{ \sum_{m=1}^{2} \frac{1}{\sigma_m} \phi\left(\frac{y_t - \mu_t(m)}{\sigma_m}\right) * P(s_t = m | \Im_{t-1}, \delta) \right\}. \qquad (2.10)$$

Equation (2.10) can then be maximized with respect to $\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}$. Due to the nature of transition probabilities, Equation (2.10) must be calculated recursively. A demonstration of the detailed procedure is beyond the scope of this paper, but it should be sufficient to state that starting with the initial filtered probability $P(s_{t-1} = m | \Im_{t-1})$ (i.e. filtered means based on available information at time *t*) one-step ahead regime prediction probabilities $P(s_t = m | \Im_{t-1})$ are computed repeatedly by a three-step procedure for all time periods $t = 1, \dots, T$. Afterwards, the results are used to update one-step-ahead filtered probabilities $P(s_t = m | \Im_t)$. Hence, Equation (2.10) can be solved by adopting a numerical-search algorithm, e.g. the Broyden-Fletcher-Goldfarb-Shanno approach (see e.g. Broyden, 1970).

Furthermore, smoothed estimates for regime probabilities, using the full information set in the final period *T*, are provided for all periods *t*, deploying the smoothing algorithm introduced by Kim (1994). Aiming to obtain the most accurate smoothed probabilities in-sample, choosing the optimal lag length of regressors $\boldsymbol{x}$ is once again performed by computing and minimizing the average of the AIC, BIC and HQ information criterion for up to three different lags of the sentiment indicator and up to 15 months in the past.

2.5.2.3 Logit Model

Finally, in order to examine the in- and out-of-sample predictive power with respect to the sign of future returns of the direct real estate market, a logit model is proposed. As stated by Wooldridge (2016), the class of binary response models can be written as:

$$P(y = 1 | \boldsymbol{x}) = P(y = 1 | x_1, x_2 \; \dots \; x_k) \,, \qquad (2.11)$$

where $\boldsymbol{x}$ is a (*1* x *k*) - matrix of explanatory variables and y a binary response variable taking either value one or zero. Assuming that the response probability is linear in a set of parameters β_k, Equation (2.11) can be written as:

$$P(y = 1|\boldsymbol{x}) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) = G(\beta_0 + \boldsymbol{x\beta}) , \qquad (2.12)$$

with *G* being a nonlinear function taking values between one and zero, $0 < G(\,) < 1$, and $\boldsymbol{\beta}$ a (*k* x *1*) - matrix of coefficients. From the set of possible functions *G*, this paper employs the "logit"-link[18] $G(\beta_0 + \boldsymbol{x\beta}) = exp(\beta_0 + \boldsymbol{x\beta})/[exp(\beta_0 + \boldsymbol{x\beta}) + 1]$. Using maximum-likelihood estimation, coefficients can be calculated from the following equation:

$$\text{logit}[P(y = 1|\boldsymbol{x})] = \ln\left(\frac{P(y = 1|\boldsymbol{x})}{1 - P(y = 1|\boldsymbol{x})}\right) = \beta_0 + \boldsymbol{x\beta} + u\,. \qquad (2.13)$$

In order to analyze the relationship between market turns and the ANN-based sentiment indicator, y_t for month *t* is set to one for periods in which the *CCRSI* return is greater than or equal to zero and zero otherwise. The matrix $\boldsymbol{x_t}$ incorporates the aforementioned set of macroeconomic controls at time *t*. With text-based sentiment indicator SI_t separately stated from $\boldsymbol{x}$, Equation (2.13) becomes:

$$\text{logit}[P(y_t = 1|\boldsymbol{x}, SI)] = \beta_0 + \boldsymbol{x_t\beta} + \sum_i \gamma_{t-i}\, SI_{t-i} + u_t\,. \qquad (2.14)$$

The optimal lag length of the sentiment indicator i is chosen analogously to the Markov-switching model. However, five (seven) lags are selected for the in-sample (out-of-sample) forecasting logit model, as the optimization procedure proposes a combination of more recent as well as more distant lags. Whenever necessary, variables are again used in first difference form or as growth rates, in order to ensure stationarity. Detailed specifications of the estimated VAR, MS and logit models can be found in the result section.

[18] Note that logit or log-odds is the natural logarithm of the odds: p/(1-p).

2.5.2.4 Forecast Evaluation

Among a variety of potential forecast accuracy measures, this paper employs two forecast evaluation criteria that are particularly suitable for binary response models, used, for example, by Diebold and Rudebusch (1989) to score leading indicators. These first metric is Brier's (1950) Quadratic Probability Score (QPS):

$$\text{QPS} = \mathrm{T}^{-1} \sum_{t=1}^{T} (\hat{y}_t - y_t)^2 \,, \tag{2.15}$$

where $\hat{y}_t$ is the ex-ante probability of an event and y_t the true binary value in period *t*. T is the total number of observations. Due to the construction of the measure, a QPS score of 0 represents a perfect model, a score of 1 implies the complete absence of predictive power. In contrast, the second metric, namely the Log Probability Score, ranges from 0 to infinity with smaller scores indicating a more accurate forecast:

$$\text{LPS} = -\mathrm{T}^{-1} \sum_{t=1}^{T} [(1 - y_t) ln(1 - \hat{y}_t) + y_t ln(\hat{y}_t)] \,. \tag{2.16}$$

2.6 Results

For the study at hand, a two-step approach was implemented: In a first step, a meaningful procedure for deriving a monthly sentiment indicator from news articles provided by the *S&P Global Market Intelligence Database* via the utilization of artificial neural networks was developed. In a second step, the usefulness of the proposed sentiment measure as an explanatory factor in a direct commercial real estate market setting is outlined. As introduced in the methodology section, three econometric methods are undertaken. Running a VAR highlights the link to direct market returns. Due to the slow nature of real assets, investigation on whether the derived sentiment indicator reacts to past market movements or vice versa is necessary. More formally, Granger-causality between CCRSI returns and changes of the sentiment indicator are examined. Afterwards, a simple MS model provides some first insights into whether the indicator's impact differs during different states of the market cycle, reflecting the boom and bust nature of the direct real estate market. Filtered probabilities are depicted over the full sample period. Following Tsolacos, Brooks and Nneji (2014), the MS approach is eventually complemented by a more elaborate logit approach, given that past research indicates that logit models provide better results in a real estate sentiment context. Moreover, a strict out-of-sample forecast framework allows for an evaluation of a future practical use, both of the suggested and similar sentiment measures. Overall, the described threefold procedure should be suitable for illustrating whether ANN-based textual sentiment indicators can achieve a robust predictive performance and therefore yield a valuable contribution to the sentiment literature in real estate.

2.6.1 Linking Sentiment to Market Returns

In accordance with the assumption of a possible bi-directional relationship between direct market returns and news-based sentiment, Table 2.2 shows the results of estimating the endogenous relationship between the constructed monthly sentiment indicator and CCRSI returns, following Equation (2.6). The presented Models 1, 2 and 3 differ in the use of macroeconomic controls, as well as the way sentiment measures are calculated. While Model 1 refrains from including controls, Models 2 and 3 include the TERM, SPREAD, INFLATION, S&P500 and CONSTRUCTION variables. Model 2 applies the sentiment measure in first differences while Model 3 uses growth rates. This implies that positive and negative indicator changes are treated relative to the prevailing level of market sentiment and thus serves as a robustness check.

For the ease of demonstration, only real estate return equations are reported. However, Granger-causalities for both directions are shown at the end of Table 2.2, as well as the commonly used model assessment criteria. The optimal lag length throughout, for all three models, is 8 months. This is reasonable, considering the sluggish direct market, and seems to be driven mainly by the strong autocorrelation of CCRSI returns. Lagged return values are statistically significant at a 1% level except for the second (and fifth) lag of Model 1, 2 and 3, respectively. Even though the incorporation of more lags in the macroeconomic controls would be preferable, available degrees of freedom limit the number of lags. By the incorporation of additional lags, it seems likely that a robust estimation will be threatened. Therefore, only the first, second and third lag of controls are used for Model 2 and 3. All three specifications are tested for statistical robustness. Although the results are quite similar, it is worth noting that the extended Models 2 and 3 appear more robust than Model 1.[19]

[19] When running a White test, Model 1 shows some evidence of heteroscedasticity. However, further discussion focuses on the results of Model 2.

Table 2.2: VAR Estimation Results

	CoStar Commercial Repeat-Sales Index (CCRSI)					
	Model 1		**Model 2**		**Model 3**	
	Δ(Sentiment) no controls		Δ(Sentiment) incl. controls		g(Sentiment) incl. controls	
CCRSI (-1)	1.177	***	1.099	***	1.096	***
	[14.3492]		[12.0824]		[12.0976]	
CCRSI (-2)	-0.218	*	-0.200	*	-0.194	
	[-1.87171]		[-1.66502]		[-1.61711]	
CCRSI (-3)	-1.008	***	-0.977	***	-0.987	***
	[-9.12445]		[-8.37788]		[-8.51011]	
CCRSI (-4)	1.326	***	1.208	***	1.209	***
	[9.63748]		[8.20966]		[8.22422]	
CCRSI (-5)	-0.412	***	-0.322	**	-0.317	**
	[-2.93737]		[-2.14498]		[-2.10848]	
CCRSI (-6)	-0.433	***	-0.444	***	-0.447	***
	[-3.90108]		[-3.74250]		[-3.76839]	
CCRSI (-7)	0.652	***	0.572	***	0.569	***
	[5.57111]		[4.68253]		[4.65108]	
CCRSI (-8)	-0.308	***	-0.299	***	-0.292	***
	[-3.73911]		[-3.41830]		[-3.34233]	
Sentiment indicator (-1)	0.011		0.000		-0.001	
	[0.46503]		[0.00562]		[-0.07789]	
Sentiment indicator (-2)	0.052	*	0.053		0.034	*
	[1.89684]		[1.65666]		[1.68505]	
Sentiment indicator (-3)	0.008		-0.004		-0.003	
	[0.27250]		[-0.12782]		[-0.11616]	
Sentiment indicator (-4)	0.026		0.000		0.001	
	[0.93510]		[0.00636]		[0.04570]	
Sentiment indicator (-5)	-0.006		-0.020		-0.011	
	[-0.22498]		[-0.66262]		[-0.56351]	
Sentiment indicator (-6)	0.063	**	0.050	*	0.034	*
	[2.28135]		[1.72157]		[1.90561]	
Sentiment indicator (-7)	0.049	*	0.045	*	0.032	**
	[1.89580]		[1.67656]		[1.98395]	
Sentiment indicator (-8)	0.039	*	0.030		0.019	
	[1.73121]		[1.35536]		[1.40080]	
TERM (-1)			-0.667	*	-0.664	*
			[-1.84425]		[-1.85674]	
TERM (-2)			-0.170		-0.178	
			[-0.48043]		[-0.50582]	
TERM (-3)			0.285		0.285	
			[0.83080]		[0.83423]	

(Table continues on next page.)

Table 2.2: VAR Estimation Results (cont.)

SPREAD (-1)		0.731	0.761
		[1.03711]	[1.08951]
SPREAD (-2)		-0.615	-0.614
		[-0.82511]	[-0.83098]
SPREAD (-3)		0.895	0.949
		[1.33723]	[1.42116]
INFLATION (-1)		-0.114	-0.117
		[-0.49048]	[-0.50671]
INFLATION (-2)		0.351	0.369
		[1.24480]	[1.31615]
INFLATION (-3)		-0.171	-0.185
		[-0.70392]	[-0.76261]
S&P500 (-1)		0.032 *	0.033 *
		[1.78345]	[1.84759]
S&P500 (-2)		0.036 *	0.035 *
		[1.74785]	[1.75105]
S&P500 (-3)		0.003	0.003
		[0.15135]	[0.14987]
CONSTRUCTION (-1)		0.049	0.047
		[0.74773]	[0.72241]
CONSTRUCTION (-2)		0.077	0.077
		[1.19657]	[1.19771]
CONSTRUCTION (-3)		0.072	0.076
		[1.12230]	[1.18581]
Constant	0.000	0.000	0.000
	[0.65223]	[0.08183]	[-0.17754]
Adj. R-squared	0.77	0.78	0.78
F-statistic	31.58	17.65	17.98
Log likelihood	519.63	531.69	532.83
Akaike AIC	-6.84	-6.80	-6.81
Schwarz SC	-6.49	-6.15	-6.16
Granger Causality			
Sentiment indicator	**0.09**	**0.07**	**0.03**
CCRSI	**0.05**	0.12	0.13

Significance levels: **p<0.1; **p<0.05; ***p<0.01*

Notes: This table reports results for the estimated VAR models with monthly CCRSI returns and news-based sentiment as endogenous variables. The set of macroeconomic control variables includes the difference between the 10-year US Treasury bond and 3-month Treasury bill yields (TERM), the difference between Baa- and Aaa-rated corporate bond yields (SPREAD), the inflation rate (INFLATION), S&P 500 returns (S&P500) as well as the amount of monthly seasonal adjusted construction spending (CONSTRUCTION). The table only shows the results of the real estate return equations. T-statistics are reported in square brackets underneath the coefficient estimates. In terms of Granger causality, values are reported for both directions. P-values in bold indicate a level of significance up to 10%. The sample period is 2006:M10 to 2018:M12.

All models show an adjusted R^2 of about 78% with slightly better results when macroeconomic controls are included. Due to the construction of SI as a probability score of positive market attitude, a positive coefficient sign is expected. The results indeed reveal a positive value, except for the third and the fifth lag. However, these lags are statistically insignificant. In Model 1, lags 2, 6, 7 and 8 are significant at a 5% and 10% level. When including macroeconomic controls (Model 2), the second lag of the sentiment indicator now slightly misses the 10% level of significance, while lags 6 and 7 remain significant with somewhat lower coefficients. With added controls, lag 8 is no longer significant.

Although single lags do not show high levels of significance, the text-based indicator overall does Granger-cause market returns at a 10% level of significance in both models. While a reverse relationship also holds true for Model 1, a more pronounced causality from indicator to market returns is proposed in Model 2. Considering the relatively high level of monthly fluctuation (see Figure 2.3), this had to be anticipated. While values in individual months might be noisy, the overall change in market attitude over the last couple of months can be considered a more accurate indicator of future market returns. Cholesky variance decomposition over 36 months indeed shows a contribution of the sentiment indicator in Model 1 and 2 of 7.78% and 5.21%, respectively. As a further robustness check, Model 3 employs growth rates instead of first differences. Thus, sentiment changes at high sentiment levels have a diminished impact, which reduces the overall amplitude of the sentiment indicator. The standard deviation of Δ(Sentiment) is 0.0358, while the standard deviation of g(Sentiment) is 0.0153. This is also in line with the idea that market participants react more strongly to newly arriving sentiment in contrast, for example to positive news in addition to an overall positive market attitude. Consequently, the second lag of Model 3 becomes significant again and t-statistics for the 6^{th} and 7^{th} lag increase. Furthermore, the sentiment indicator now Granger-causes CCRSI returns at a 5%, instead of a 10% level and the contribution in the variance decomposition increases slightly to 5.67%.

Overall, these findings indicate that the cumulative ANN-based sentiment measure has some return-signaling effect with respect to the direct real estate market in the United States, although the impact of individual lags is less distinct. Especially the more pronounced link from the sentiment indicator to market returns shown by all three models seems promising with respect to further evaluation.

2.6.2 Accounting for Market Regimes

In the second approach, the SI is employed in a simple Markov-switching model to explore the behavior of the SI in different market regimes and account for a potential non-linear relationship at the same time. Table 2.3 shows the estimation results of Equation (2.7). Minimizing the average of AIC, HQ and BIC suggests a need to include the 7^{th}, 8^{th} and 9^{th} lag of the SI. As can be seen, the numerical-search algorithm clearly states two distinct regimes. Average returns are positive and significant in regime 1 (up-market), while the opposite is true for regime 2 (down-market). This is

Table 2.3: Markov-Switching Model Estimations

Regime 1		**Regime 2**	
C	0.007	*C*	-0.021
	[< 1E-4]		[< 1E-4]
Sentiment indicator (-7)	0.007	*Sentiment indicator (-7)*	0.323
	[0.829]		[0.0001]
Sentiment indicator (-8)	0.039	*Sentiment indicator (-8)*	0.311
	[0.295]		[< 1E-4]
Sentiment indicator (-9)	0.001	*Sentiment indicator (-9)*	0.315
	[0.981]		[< 1E-4]

Constant transition probabilities:

	Regime 1	*Regime 2*
Regime 1	0.962	0.038
Regime 2	0.204	0.796

Constant expected durations:

	Regime 1	*Regime 2*
(months)	25.99	4.90

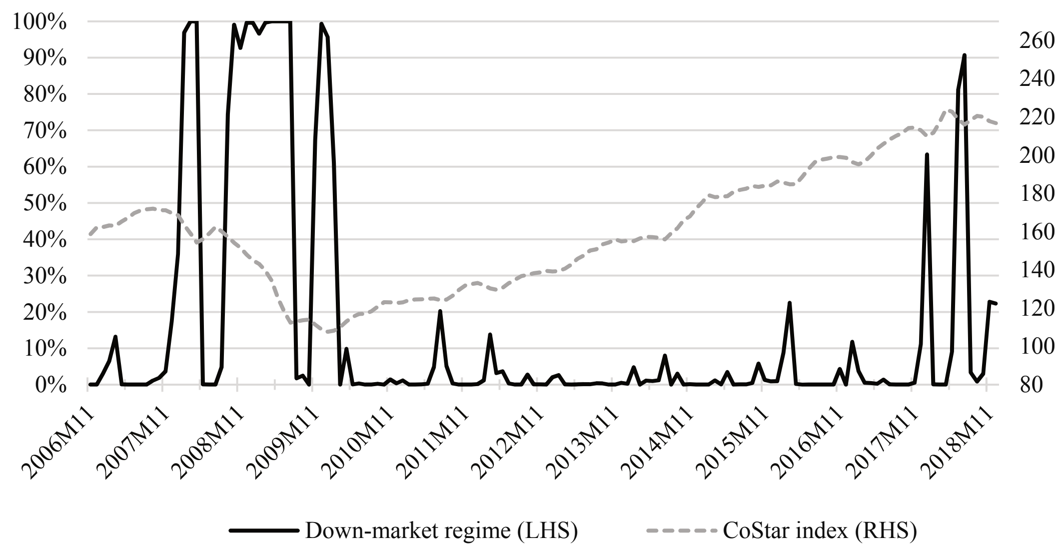
Akaike (AIC)	-5.957
Hannan-Quinn (HQ)	-5.866
Log likelihood	445.877
Schwarz (BIC)	-5.732

Notes: This table reports results for the estimated Markov-switching model with monthly CCRSI returns as the exogenous variable, and news-based sentiment as the endogenous variable. Errors are not regime-specific. No macroeconomic controls are included. T-statistics are reported in square brackets underneath the coefficient estimates. The sample period is 2006:M11 to 2018:M12.

indicated by the significantly positive (negative) values of C in regime 1 (2). However, only regime 2 shows a statistically significant relationship with lagged SI values. Estimated coefficients are highly significant and large in magnitude for all three lags. Looking at the constant transition probability matrix, both regimes – the up-market regime 1 as well as the down-market regime 2 – are very stable with switching probabilities out of the up-market (regime 1) of 3.8% and out of the down-market (regime 2) of 20.4%. In accordance with the development of the CCRSI over the study period, the expected duration is almost 26 months for regime 1 and only 5 months for regime 2. Because the MS model is presented mainly as a supplement to the following logit model, no controls are included in the model shown in Table 2.3. However, the results do not change substantially when similar controls with identical lags as in the VAR are included.

Figure 2.4 provides an initial visual indication of the predictive potential of the SI, depicting the estimated filtered probabilities of being in the down-market regime using all information available up to 2018:M12. Probability scores are stated on the left and CCRSI values on the right. The model seems to achieve acceptable in-sample performance. In 2007:M10, the *CoStar* index began to fall and the filtered probabilities of being in the down-market regime started to rise one month earlier. Interestingly, the market rebounds in March 2008 and September 2009 are captured in the model as well. Afterwards, no prediction values above 0.5 are reported until January 2018, which indeed corresponds to a 1.51% index decrease. It is worth noting that this month was the biggest dip since January 2010. Furthermore, the negative growth period from May 2018 to July 2018 is identified by the model. While the model apparently depicts larger swings quite accurately, smaller index decreases are identified in the form of short-term probability rises only, without reaching the required 50 percent threshold.

Figure 2.4: Markov-Switching - Filtered Probabilities



Notes: This figure depicts filtered probabilities computed by the Markov-switching model estimated in Table 2.3. The *CoStar Commercial Repeat-Sales Index* is plotted on the RHS. The up-market regime (1 - filtered probability of the down-market regime) is not shown for ease of demonstration. The sample period is 2006:M11 to 2018:M12.

In order to also control for this regime-varying nature of the SI in the VAR model, Equation (2.6) is re-estimated for subsamples of positive and negative market returns, only. In accordance to the findings of the MS model, a weaker influence of the SI is expected during up-market periods and a more pronounced one during down-market phases. Therefore, Table 2.4 facilitates Model 2 of Table 2.2 and recalculates the results in the form of Model 4 and Model 5 for up- and down months, respectively. Once again, robustness checks were conducted for the two additional models.

In Model 2, the 6^{th} and 7^{th} lag of the sentiment indicator are significant at a 10% level. However, not a single lag remains its level of significance when only accounting for months with positive market returns in Model 4. This is also reflected in the massively decreasing adjusted R^2 of 47% compared to the former value of 78%. Neither the sentiment indicator nor CCRSI returns Granger-cause each other. In contrast, the adjusted R^2 rises to almost 90% in Model 5 although all three models include the same controls and show a similar autoregressive behavior of the CCRSI. The sentiment indicator becomes highly significant at a 1% and 5% level for lags 6 and 8 and the 7^{th} lag slightly misses the 10% level of significance. Accordingly, the sentiment indicator Granger-causes market returns at a 1% level in Model 5.

It is also worth noting that all sentiment coefficients of Model 4 show a negative sign while this was only occasionally true for the other VAR models. Although not being significant, this could further imply that positive sentiment changes do not only have no impact on returns during boom periods but may even dampen returns. While a positive relationship of market sentiment and market returns is more obvious, the reverse relationship could be the result of skepticism during longer boom periods such as the market run-up after the financial crisis.

Table 2.4: VAR Estimation Results in Up- and Down-Market Periods

	CoStar Commercial Repeat-Sales Index (CCRSI)					
	Model 2		**Model 4**		**Model 5**	
	Δ(Sentiment)		Δ(Sentiment) up-market		Δ(Sentiment) down-market	
CCRSI (-1)	1.099	***	0.565	***	0.859	***
	[12.0824]		[4.71986]		[4.67879]	
CCRSI (-2)	-0.200	*	-0.022		-0.155	
	[-1.66502]		[-0.17402]		[-0.56743]	
CCRSI (-3)	-0.977	***	-0.696	***	-0.567	**
	[-8.37788]		[-5.99956]		[-2.10357]	
CCRSI (-4)	1.208	***	0.595	***	0.962	***
	[8.20966]		[3.77978]		[3.70730]	
CCRSI (-5)	-0.322	**	-0.067		-0.316	
	[-2.14498]		[-0.44378]		[-1.14289]	
CCRSI (-6)	-0.444	***	-0.302	**	-0.505	**
	[-3.74250]		[-2.51994]		[-2.24446]	
CCRSI (-7)	0.572	***	0.309	**	0.584	**
	[4.68253]		[2.62415]		[2.55502]	
CCRSI (-8)	-0.299	***	-0.100		-0.196	
	[-3.41830]		[-1.10348]		[-1.14012]	
Sentiment indicator (-1)	0.000		-0.037		-0.071	
	[0.00562]		[-1.43699]		[-1.54383]	
Sentiment indicator (-2)	0.053		-0.032		0.013	
	[1.65666]		[-0.96833]		[0.26336]	
Sentiment indicator (-3)	-0.004		-0.057		-0.084	
	[-0.12782]		[-1.52082]		[-1.34026]	
Sentiment indicator (-4)	0.000		-0.048		-0.049	
	[0.00636]		[-1.22552]		[-0.71172]	
Sentiment indicator (-5)	-0.020		-0.051		-0.031	
	[-0.66262]		[-1.50435]		[-0.42428]	
Sentiment indicator (-6)	0.050	*	-0.025		0.130	**
	[1.72157]		[-0.79392]		[2.09827]	
Sentiment indicator (-7)	0.045	*	-0.011		0.107	
	[1.67656]		[-0.43780]		[1.59995]	
Sentiment indicator (-8)	0.030		-0.001		0.126	***
	[1.35536]		[-0.05361]		[2.99813]	
TERM (-1)	-0.667	*	-0.223		-0.254	
	[-1.84425]		[-0.51861]		[-0.41091]	
TERM (-2)	-0.170		0.056		0.276	
	[-0.48043]		[0.14011]		[0.47093]	
TERM (-3)	0.285		0.030		0.690	
	[0.83080]		[0.08376]		[0.82694]	

(Table continues on next page.)

Table 2.4: VAR Estimation Results in Up- and Down-Market Periods (cont.)

SPREAD (-1)	0.731		0.813		0.025	
	[1.03711]		[0.74043]		[0.02761]	
SPREAD (-2)	-0.615		0.331		-1.366	
	[-0.82511]		[0.30833]		[-1.57395]	
SPREAD (-3)	0.895		0.683		0.447	
	[1.33723]		[0.81408]		[0.45329]	
INFLATION (-1)	-0.114		-0.163		-0.640	*
	[-0.49048]		[-0.65668]		[-1.72362]	
INFLATION (-2)	0.351		0.073		0.759	
	[1.24480]		[0.24085]		[1.52136]	
INFLATION (-3)	-0.171		-0.139		-0.885	**
	[-0.70392]		[-0.53010]		[-2.08401]	
S&P500 (-1)	0.032	*	0.003		-0.015	
	[1.78345]		[0.12483]		[-0.40924]	
S&P500 (-2)	0.036	*	0.015		0.031	
	[1.74785]		[0.52224]		[0.81317]	
S&P500 (-3)	0.003		-0.022		-0.010	
	[0.15135]		[-0.86599]		[-0.31594]	
CONSTRUCTION (-1)	0.049		0.025		0.075	
	[0.74773]		[0.37886]		[0.66415]	
CONSTRUCTION (-2)	0.077		0.028		0.135	
	[1.19657]		[0.43427]		[0.99440]	
CONSTRUCTION (-3)	0.072		-0.017		0.144	
	[1.12230]		[-0.27362]		[1.04782]	
Constant	0.000		0.008	***	-0.006	***
	[0.08183]		[5.73315]		[-3.83994]	
Adj. R-squared	0.78		0.47		0.87	
F-statistic	17.65		3.77		11.59	
Log likelihood	531.69		386.94		217.34	
Akaike AIC	-6.80		-7.32		-7.41	
Schwarz SC	-6.15		-6.47		-6.19	
Granger Causality						
Sentiment indicator	**0.071**		0.755		**0.004**	
CCRSI	**0.117**		0.366		0.572	

Significance levels: **p<0.1; **p<0.05; ***p<0.01*

Notes: This table reports results for the estimated VAR models with monthly CCRSI returns and news-based sentiment as endogenous variables for the whole sample period as well as for months with positive returns (up-market) and negative returns (down-market), only. The set of macroeconomic control variables includes the difference between the 10-year US Treasury bond and 3-month Treasury bill yields (TERM), the difference between Baa- and Aaa-rated corporate bond yields (SPREAD), the inflation rate (INFLATION), S&P 500 returns (S&P500) as well as the amount of monthly seasonal adjusted construction spending (CONSTRUCTION). The table only shows the results of the real estate return equations. T-statistics are reported in square brackets underneath the coefficient estimates. In terms of Granger causality, values are reported for both directions. P-values in bold indicate a level of significance up to 10%. The sample period is 2006:M10 to 2018:M12.

2.6.3 Binary Return Forecasts

Finally, following the reasoning of Tsolacos, Brooks and Nneji (2014), the market return models are complimented by a logit approach. By doing so, one can study the influences of the constructed sentiment indicator in a binary return prediction framework, which is presumably of greater practical use for market participants than the derivation of point return forecasts. The SI, as well as macroeconomic controls are used as the predictor series in Model 7, according to Equation (2.14). Model 6 is a reduced version with sentiment indicators and a constant only. Lags were selected for both models, based on the lowest average of HQ, BIC and AIC, thus facilitating information for the full observation period 2007:M03 to 2018:M12. Hence, the 1st, 2nd, 11th, 12th and 13th lags are chosen, with that including sentiment information for more than one year in the past. The information criterion results evidently imply the importance of some seasonal information, as the model captures the effect of the 1st and 13th lags (i.e. the same month) in the preceding year. With regressand values of 1 for direct market returns equal to or greater than zero, a positive sign of SI is expected and confirmed in Table 2.5. Furthermore, both times, SI lags are significant at a 5% or 1% level for 3 (4) of 5 lags. The likelihood ratio test for joint significance is passed by both models, and the full model reaches a McFadden's R^2 of 27.1%. The hypothesis of good-fit in the conducted Hosmer-Lemeshow tests with 10 quantiles cannot be rejected. The percentage gain in comparison to a constant probability model is 10% and 32% for Model 6 and 7, respectively.

To provide insights into the forecast performance of the SI in a binary return setting, forecasting accuracy has to be evaluated. Thus, in- and out-of-sample forecasts are provided for the logit framework. Figure 2.5 depicts periods of non-negative market growth, as well as one-month-ahead forecasts. Note that for this in-sample performance test, Model 7, optimized with information criteria calculated for the whole sample, can be applied. For evaluation of the out-of-sample performance described later on, the model is optimized based on information until the end of 2015 only. During the shaded periods, probabilities above 50% are expected from the logit model. Similar to the MS model, the large swings from 2007:M4 until 2009:M07 are well captured. There are some incorrectly forecasted returns – notably September 2007 and June 2009 – but usually, periods of negative market growth are associated with

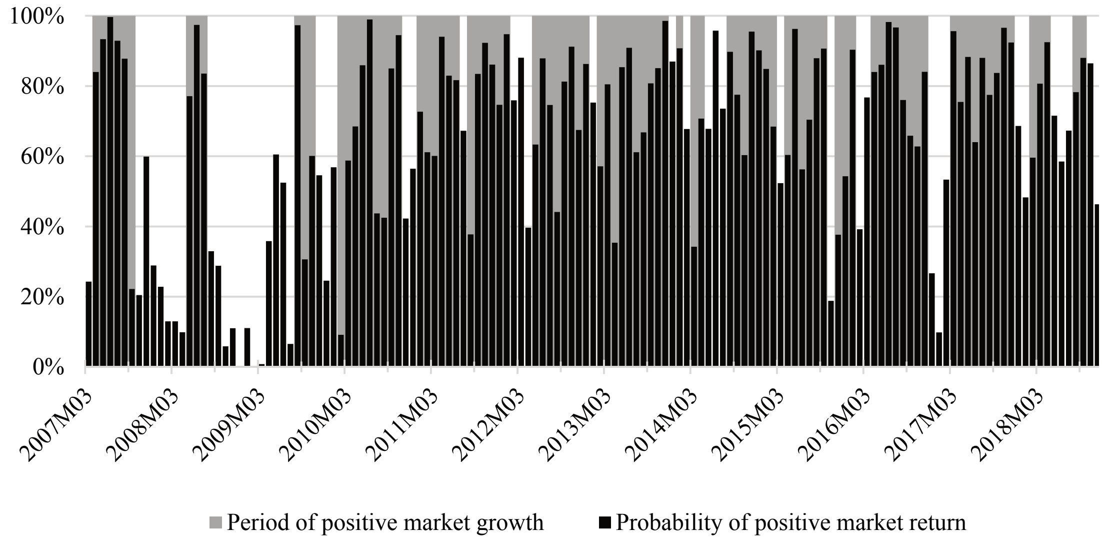
Table 2.5: Logit Estimation Results

	Pr[CCRSI return = 1]			
	Model 6		Model 7	
	no macroeconomic controls		with macroeconomic controls	
Sentiment indicator (-1)	12.011	*	10.687	
Sentiment indicator (-2)	9.551		20.399	**
Sentiment indicator (-11)	19.419	***	27.333	***
Sentiment indicator (-12)	25.133	***	39.435	***
Sentiment indicator (-13)	16.508	**	34.882	***
TERM (-1)			-383.207	***
TERM (-2)			16.972	
TERM (-3)			86.966	
SPREAD (-1)			632.361	**
SPREAD (-2)			-643.139	*
SPREAD (-3)			22.216	
INFLATION (-1)			-28.192	
INFLATION (-2)			73.209	
INFLATION (-3)			39.480	
S&P500 (-1)			12.461	*
S&P500 (-2)			16.022	**
S&P500 (-3)			5.100	
CONSTRUCTION (-1)			24.055	
CONSTRUCTION (-2)			16.385	
CONSTRUCTION (-3)			-0.995	
Constant	0.643	***	0.294	
McFadden R-squared	0.086		0.271	
Akaike info criterion (AIC)	1.271		1.241	
Schwarz criterion (BIC)	1.396		1.679	
Hannan-Quinn criterion (HQ)	1.322		1.419	
LR statistic	15.788		49.966	
Prob (LR statistic)	0.0075		0.0002	
Significance levels:			**p<0.1; **p<0.05; ***p<0.01*	

Notes: This table reports results for the estimated logit models with monthly Pr[CCRSI returns = 1] as the endogenous variable. The constructed sentiment indicator, as well as a set of macroeconomic controls, are included in the extended model, while the reduced model includes a constant and the sentiment measures only. Utilized macroeconomic control variables are the difference between the 10-year US Treasury bond and 3-month Treasury bill yields (TERM), the difference between Baa- and Aaa-rated corporate bond yields (SPREAD), the inflation rate (INFLATION), S&P 500 returns (S&P500), as well as the amount of monthly seasonal-adjusted construction spending (CONSTRUCTION). The sample period is 2007:M03 to 2018:M12.

probabilities below 50% and vice versa. Looking at the following years, the model once again struggles with shorter swings. Nevertheless, as depicted in the top panel of Table 2.6, the hit rate/correct sign prediction is 76.06% from March 2007 until the end of 2018. A naïve model facilitating the average return over the 13-year sample period yields a hit rate of 64.79% only. Additionally, the QPS and LPS are 31.94% and 27.21% lower, respectively.[20]

Figure 2.5: In-Sample Probability Forecast for Market Return Directions



Notes: This figure depicts one-step-ahead in-sample forecasts computed by means of the logit model of Table 2.5. CCRSI returns are included as a second series to indicate periods of positive market growth. The sample period is 2007:M03 to 2018:M12.

From a market participant standpoint, only out-of-sample performance provides real insight into SI's predictive potential. As the last four years of the study period provide an especially challenging environment with four distinct periods of positive returns, as well as five periods of negative returns (compare with Figure 2.6), factual out of sample forecasting performance from 2016:M01 to 2018:M12 is worth investigating. Thus, based on the information available up to end of 2015, a logit model is optimized and estimated. In contrast to Model 7, the 1st, 2nd, 3rd, 4th, 11th, 12th and 13th lag are suggested by the AIC, BIC and HQ. A one-month-ahead forecast for January 2016 is provided with controls included in the equation. Afterwards, the information period is extended by one month, the model is re-estimated and the next forecasting value is derived. Overall, 36 forecasts are made for 36 months, based on an individually estimated model each time. The results are contrasted to a naïve model using the average direct market return derived from preceding months in the study period, when prediction and forecasting accuracy measures are calculated. With respect to correct predictions, the logit model yields 66.67% accuracy in contrast to 63.89% for the naïve model. However, note that the naïve model benefits from a surplus of positive return

[20] When excluding controls, the model still yields better results than the naïve model, but outperforms by a smaller margin.

periods in the past, as well as during the forecasting period. Figure 2.6 helps to explain the mediocre out-of-sample results. Although the model reacts to periods of negative market returns by reducing the forecasting values accordingly, the adjustments are once again not strong enough. As down-market phases in the period facilitated for the forecast last no longer than 3 months, the logit model does not adapt appropriately, leading to a high error rate during those market periods.

Table 2.6: Forecast Performance

In-sample forecast performance

	Logit model	**Naïve model**
Hit rate / correct-sign prediction	76.06 %	64.79 %
Brier's Quadratic Probability Score (QPS)	0.156	0.229
Log Probability Score (LPS)	0.473	0.650

Period: 2007:M03 - 2018:M12
Lagged terms: -1, -2, -11, -12, -13

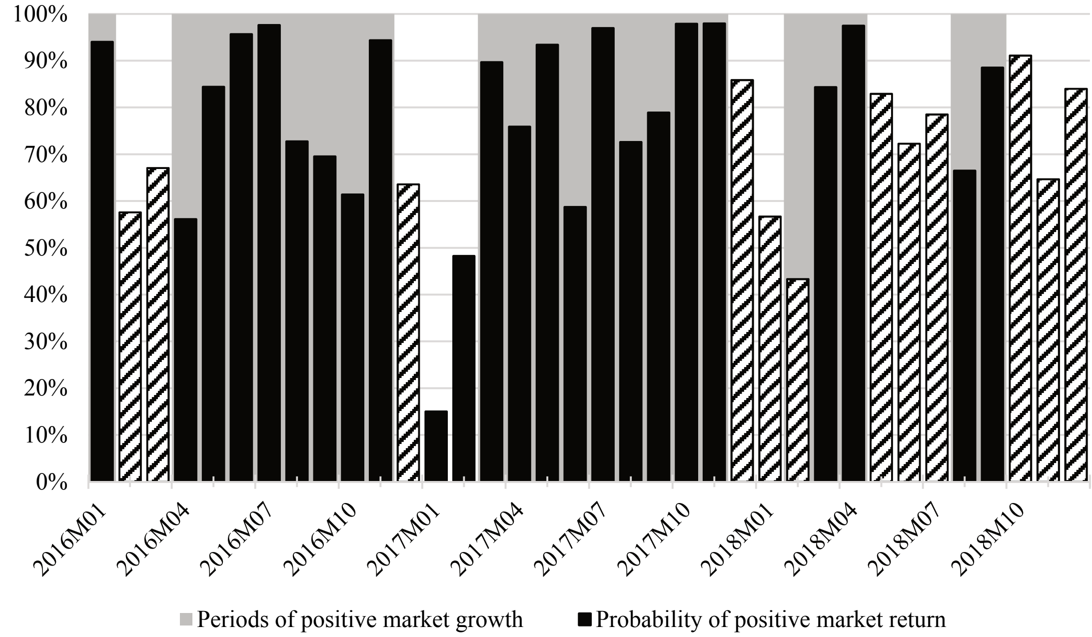
Out-of-sample forecast performance

	Logit model	**Naïve model**
Hit rate / correct-sign prediction	66.67 %	63.89 %
Brier's Quadratic Probability Score (QPS)	0.213	0.233
Log Probability Score (LPS)	0.604	0.660

Period: 2016:M01 - 2018:M12
Lagged terms: -1, -2, -3, -4, -11, -12, -13

Notes: This table reports in- and out-of-sample forecast performance for estimated logit models with monthly Pr[CCRSI returns = 1] as endogenous variable. The constructed sentiment indicator, as well as the same set of macroeconomic controls as in Table 2.5, are included as exogenous variables. Chosen lags for in-sample and out-of-sample models are based on minimizing the AIC, HQ and BC for the full sample period 2006:M01-2018:M12 and 2006:M01-2015:M12, respectively. For in-sample performance, the optimal model is estimated, including all information available up to 2018:M12. The resulting model is used to make all one-month-ahead predictions without continuously updating the model coefficients. The first out-of-sample model facilitates information until 2015:M12 only. One-step-ahead forecasts are conducted by estimating the model with given information from the past and extending the estimation window gradually by one month afterwards (i.e. coefficient estimates and forecasts are updated every month). As return directions are forecasted only, the hit rate and correct sign prediction measure yield the same result. QPS ranges from 0 to 1 with a better model exhibiting a lower QPS value. LPS ranges from 0 to infinity, with lower scores indicating a more accurate forecasting model. In cases of in-sample performance, the naïve model facilitates the share of positive return periods from 2006:M01 to 2018:M12 for the forecast. For out-of-sample performance, the average percentage of past positive returns is used for the forecast and this value is updated every month in accordance with the logit model.

Figure 2.6: Out-of-Sample Forecasting Performance



Notes: The figure depicts one-step-ahead out-of-sample forecasts computed by means of a logit model. CCRSI returns are included as a second series to indicate periods of positive market growth. Shaded periods indicate wrong predictions. The sample period is 2016:M01 to 2018:M12.

2.6.4 Synopsis

Taking into account all presented results, the ANN-based textual sentiment indicator shows explanatory and predictive potential, but also exhibits some shortcomings. Some return-signaling effect with respect to the direct real estate market was demonstrated, as indicated by Granger-causality and significant returns in the VAR model. The MS framework showed that the sentiment indicator's impact differs during up- and down-market phases and may even have reversed impact during boom periods. In-sample calculations within the logit framework further highlighted forecasting potential in terms of indication of binary market development, with a hit rate of 76.06%. However, the findings also revealed that the SI has problems capturing sudden swings in the market. This first became evident with the depiction of filtered probabilities in Figure 2.4 and was later confirmed within the logit frameworks. While SI did recognize the changes, it did not adopt fast enough. This could be due to several (potentially contrarian) reasons. Either information available within one period is not sufficient, and consequently, more textual documents have to be aggregated to obtain a more pronounced signal, or there is a high level of ambiguous information. This would mean that the measure is too noisy to allow more timely reactions. Thus, training of the classifier could be improved or the measure could to be passed through a subsequent filtering process to extract and distil more accurate information. The more pronounced results of VAR Model 3 (using relative changes of the sentiment indicator) compared to Model 2 (facilitating absolute changes) suggest this conjecture. Hence, this study showed that the ANN-based sentiment extraction procedure can be considered a promising alternative in the realm of real estate, which still provides a vast range of optimization opportunities for future research.

2.7 Conclusion

By analyzing and extracting market sentiment from 66,070 news articles on the real estate market in the United States, this paper is centered on exploring the explanatory and predictive potential of text-based sentiment indicators by means of deep-learning. In a novel approach, a densely-connected ANN is trained via distant supervision-labeled data comprising long and short ideas provided by *Seeking Alpha.* The gained knowledge is applied to *S&P Global Market Intelligence* news articles, which are classified accordingly and aggregated in a monthly sentiment index. A threefold econometric approach assesses the link to direct market returns and forecast potential with respect to return estimates and periods of positive/negative market growth. In doing so, the SI reveals potential, but also some shortcomings. Especially the weak capabilities of fully capturing faster swings are noteworthy.

In a global environment, multi-asset-class portfolio investors require early signals when assessing risks and comparing asset classes for future investment decisions. As direct real estate is slow by nature and less transparent due to heterogeneous assets, sentiment indicators evidently do provide useful information. The VAR and Markov-switching models showed that the sentiment indicator has some return signaling potential but its influence may differ during boom and bust periods of the market. With respect to the more practically applicable forecast of up- and down-market periods, the results are mixed. While in-sample forecasts provide satisfactory results, out-of-sample forecast precision suffers in a high volatility forecasting period. A more pronounced adjustment of the indicator would be required for more accurate results.

However, the relationship between the ANN-based indicator and market returns is not negligible. The indicator did Granger-cause direct market returns during the study period both with and without accounting for its regime-specific behavior. Hence, future research should try to overcome the remaining deficiencies of the sentiment indicator.

Bearing in mind the shortcomings of alternatives, any improvement of the proposed methodology seems worthwhile. Surveys are not provided at high frequency and are both time consuming and expensive by nature. Other market proxies such as closed-end fund discounts or mortgage fund flows are heavily theory-driven, possibly leading to decreased operationality. Neither such direct nor indirect indicators provide the flexibility of text-based sentiment measures with respect to temporal aggregation periods and transferability to other key figures of the real estate industry. Forecasting potential with respect to rents, cap rates and market volatility has yet to be assessed.

It should also be stressed that the use of text-based deep-learning sentiment indicators is not limited to commercial real estate. Especially the application of text mining in a housing context seems promising. Due to distant supervision-labeled data that, for example, local broker recommendations can provide, as well as the capability of a deep-learning framework to independently create classification rules, an adaption to regional or sector-specific markets is certainly possible. This is a clear advantage of the ANN-based textual sentiment gauge, in contrast to other and more widespread dictionary-based measures.

Altogether, those findings highlight the importance of news-analytics for direct real estate markets in general, as well as the potential of deep-learning text-based sentiment indicators in particular. With respect to the securitized real estate market, the indicator’s reaction time presumably has to be shortened significantly. However, as shown by related research in finance, the use of filtering techniques, as well as an extended text corpus, might allow a high-frequency application of the sentiment indicator in the realm of listed real estate as well. This seems worth investigating in future research.

2.8 References

Antweiler, W. and M. Z. Frank, Is all that talk just noise? The information content of internet stock message boards, The Journal of Finance, 2004, 59:3, 1259–94.

Barkham, R. J. and C. W. R. Ward, Investor sentiment and noise traders: Discount to net asset value in listed property companies in the U.K, Journal of Real Estate Research, 1999, 18:2, 291–312.

Beracha, E. and M. B. Wintoki, Forecasting residential real estate price changes from online search activity, Journal of Real Estate Research, 2013, 35:3, 283–312.

Bianchi, D. and M. Guidolin, Can linear predictability models time bull and bear real estate markets? Out-of-sample evidence from REIT portfolios, The Journal of Real Estate Finance and Economics, 2014, 49:1, 116–64.

Borovkova, S. and M. Dijkstra, Deep learning prediction of the EUROSTOXX 50 with news sentiment, SSRN Electronic Journal, 2018.

Boudoukh, J., R. Feldman, S. Kogan and M. Richardson, Which news moves stock prices? A textual analysis, NBER Working Paper No. 18725, National Bureau of Economic Research, Cambridge, 2013.

Braun, N., Google search volume sentiment and its impact on REIT market movements, Journal of Property Investment & Finance, 2016, 34:3, 249–62.

Brier, G. W., Verification of forecasts expressed in terms of probability, Monthly Weather Review, 1950, 78:1, 1–3.

Broyden, C. G., The convergence of a class of double-rank minimization algorithms 1. General considerations, IMA Journal of Applied Mathematics, 1970, 6:1, 76–90.

Clayton, J., D. C. Ling and A. Naranjo, Commercial real estate valuation: Fundamentals versus investor sentiment, The Journal of Real Estate Finance and Economics, 2009, 38:1, 5–37.

Cortes, C. and V. Vapnik, Support-vector networks, Machine Learning, 1995, 20:3, 273–97.

Cutler, D. M., J. M. Poterba and L. H. Summers, What moves stock prices?, The Journal of Portfolio Management, 1989, 15:3, 4–12.

Das, P. K., J. Freybote and G. Marcato, An investigation into sentiment-induced institutional trading behavior and asset pricing in the REIT market, The Journal of Real Estate Finance and Economics, 2015, 51:2, 160–89.

Das, S. R. and M. Y. Chen, Yahoo! for Amazon: Sentiment extraction from small talk on the web, Management Science, 2007, 53:9, 1375–88.

Davis, A. K., J. M. Piger and L. M. Sedor, Beyond the numbers: Measuring the information content of earnings press release language, Contemporary Accounting Research, 2012, 29:3, 845–68.

Davis, A. K. and I. Tama-Sweet, Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A, Contemporary Accounting Research, 2012, 29:3, 804–37.

Diebold, F. X. and G. D. Rudebusch, Scoring the leading indicators, The Journal of Business, 1989, 62:3, 369–91.

Dietzel, M. A., Sentiment-based predictions of housing market turning points with Google trends, International Journal of Housing Markets and Analysis, 2016, 9:1, 108–36.

Dietzel, M. A., N. Braun and W. Schäfers, Sentiment-based commercial real estate forecasting with Google search volume data, Journal of Property Investment & Finance, 2014, 32:6, 540–69.

Din, A., M. Hoesli and A. Bender, Environmental variables and real estate prices, Urban Studies, 2001, 38:11, 1989–2000.

Engelberg, J. E., A. V. Reed and M. C. Ringgenberg, How are shorts informed?, Journal of Financial Economics, 2012, 105:2, 260–78.

Feldman, R., S. Govindaraj, J. Livnat and B. Segal, Management's tone change, post earnings announcement drift and accruals, Review of Accounting Studies, 2010, 15:4, 915–53.

Ferguson, N., D. Philip, H. Y. T. Lam and J. M. Guo, Media content and stock returns: The predictive power of press, Multinational Finance Journal, 2015, 19:1/1, 1–31.

French, K. R. and R. Roll, Stock return variances, Journal of Financial Economics, 1986, 17:1, 5–26.

Freybote, J., Real estate sentiment as information for REIT bond pricing, Journal of Property Research, 2016, 33:1, 18–36.

Freybote, J. and P. A. Seagraves, Heterogeneous investor sentiment and institutional real estate investments, Real Estate Economics, 2017, 45:1, 154–76.

Freybote, J. and P. A. Seagraves, The impact of investor sentiment on commercial real estate market liquidity, Journal of Real Estate Research, 2018, 40:4, 597–627.

García, D., Sentiment during recessions, The Journal of Finance, 2013, 68:3, 1267–300.

Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski and L. Brilliant, Detecting influenza epidemics using search engine query data, Nature, 2009, 457:7232, 1012–4.

Groß-Klußmann, A. and N. Hautsch, When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions, Journal of Empirical Finance, 2011, 18:2, 321–40.

Hamilton, J. D., A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica, 1989, 57:2, 357–84.

Hanna, A. J., J. D. Turner and C. B. Walker, News media and investor sentiment over the long run, QUCEH Working Paper Series 2017-06, Queen's University Centre for Economic History, Queen's University, Belfast, 2017.

Hausler, J., J. Ruscheinsky and M. Lang, News-based sentiment analysis in real estate: A machine learning approach, Journal of Property Research, 2018, 35:4, 344–71.

Henry, E., Are investors influenced by how earnings press releases are written?, Journal of Business Communication, 2008, 45:4, 363–407.

Henry, E. and A. J. Leone, Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone, The Accounting Review, 2016, 91:1, 153–78.

Heston, S. L. and N. R. Sinha, News versus sentiment: Predicting stock returns from news stories, FEDS Working Paper No. 2016-048, Board of Governors of the Federal Reserve System, Washington, 2016.

Hoesli, M., C. Lizieri and B. MacGregor, The inflation hedging characteristics of US and UK investments: A multi-factor error correction approach, The Journal of Real Estate Finance and Economics, 2008, 36:2, 183–206.

Hohenstatt, R., M. Käsbauer and W. Schäfers, “Geco” and its potential for real estate research: Evidence from the U.S. housing market, Journal of Real Estate Research, 2011, 33:4, 471–506.

Huang, X., S. H. Teoh and Y. Zhang, Tone management, The Accounting Review, 2014, 89:3, 1083–113.

Jegadeesh, N. and D. Wu, Word power: A new approach for content analysis, Journal of Financial Economics, 2013, 110:3, 712–29.

Kearney, C. and S. Liu, Textual sentiment in finance: A survey of methods and models, International Review of Financial Analysis, 2014, 33, 171–85.

Kim, C.-J., Dynamic linear models with Markov-switching, Journal of Econometrics, 1994, 60:1-2, 1–22.

Kothari, S. P., X. Li and J. E. Short, The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis, The Accounting Review, 2009, 84:5, 1639–70.

Krizhevsky, A., I. Sutskever and G. E. Hinton, ImageNet classification with deep convolutional neural networks, In F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger, editors, Advances in neural information processing systems 25, Curran Associates, Inc, 2012, 1097–105.

Li, F., The information content of forward-looking statements in corporate filings - A naïve bayesian machine learning approach, Journal of Accounting Research, 2010, 48:5, 1049–102.

Lin, C. Y., H. Rahman and K. Yung, Investor sentiment and REIT returns, The Journal of Real Estate Finance and Economics, 2009, 39:4, 450–71.

Ling, D. C., A. Naranjo and B. Scheick, Investor sentiment, limits to arbitrage and private market returns, Real Estate Economics, 2014, 42:3, 531–77.

Liu, B., Sentiment analysis and opinion mining, Synthesis Lectures on Human Language Technologies, 2012, 5:1, 1–167.

Lizieri, C., S. Satchell, E. Worzala and R. Dacco, Real interest regimes and real estate performance: A comparison of UK and US markets, Journal of Real Estate Research, 1998, 16:3, 339–56.

Loughran, T. and B. McDonald, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, The Journal of Finance, 2011, 66:1, 35–65.

Loughran, T. and B. McDonald, The use of word lists in textual analysis, Journal of Behavioral Finance, 2015, 16:1, 1–11.

Loughran, T. and B. McDonald, Textual analysis in accounting and finance: A survey, Journal of Accounting Research, 2016, 54:4, 1187–230.

Marcato, G. and A. Nanda, Information content and forecasting ability of sentiment indicators: Case of real estate market, Journal of Real Estate Research, 2016, 38:2, 165–203.

Ozik, G. and R. Sadka, Media and investment management, SSRN Electronic Journal, 2012.

Peterson, S. and A. Flanagan, Neural network hedonic pricing models in mass real estate appraisal, Journal of Real Estate Research, 2009, 31:2, 147–64.

Rochdi, K. and M. Dietzel, Outperforming the benchmark: Online information demand and REIT market performance, Journal of Property Investment & Finance, 2015, 33:2, 169–95.

Roll, R., R^2, The Journal of Finance, 1988, 43:3, 541–66.

Rosenblatt, F., The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review, 1958, 65:6, 386–408.

Ruscheinsky, J., M. Lang and W. Schäfers, Real estate media sentiment through textual analysis, Journal of Property Investment & Finance, 2018, 36:5, 410–28.

Sinha, N. R., Underreaction to news in the US stock market, Quarterly Journal of Finance, 2016, 6:2, 1–46.

Smales, L. A., News sentiment in the gold futures market, Journal of Banking & Finance, 2014, 49, 275–86.

Soo, C. K., Quantifying animal spirits: News media and sentiment in the housing market, Ross School of Business Working Paper No. 1200, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, 2015.

Sun, L., M. Najand and J. Shen, Stock return predictability and investor sentiment: A high-frequency perspective, Journal of Banking & Finance, 2016, 73, 147–64.

Tetlock, P. C., Giving content to investor sentiment: The role of media in the stock market, The Journal of Finance, 2007, 62:3, 1139–68.

Tetlock, P. C., M. Saar-Tsechansky and S. Macskassy, More than words: Quantifying language to measure firms' fundamentals, The Journal of Finance, 2008, 63:3, 1437–67.

Tieleman, T. and G. E. Hinton, Lecture 6.5-Rmsprop, Coursera: Neural networks for machine learning, Technical Report, Coursera, Inc., 2012.

Tsolacos, S., C. Brooks and O. Nneji, On the predictive content of leading indicators: The case of U.S. real estate markets, Journal of Real Estate Research, 2014, 36:4, 541–73.

Walker, C. B., Housing booms and media coverage, Applied Economics, 2014, 46:32, 3954–67.

Walker, C. B., The direction of media influence: Real-estate news and the stock market, Journal of Behavioral and Experimental Finance, 2016, 10, 20–31.

Werbos, P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences, Thesis (Ph. D.), Cambridge, 1974.

Wooldridge, J. M., Introductory econometrics: A modern approach, Sixth Edition, Cengage Learning: Australia, 2016.

3 Artificial Intelligence, News Sentiment, and Property Market Liquidity

3.1 Abstract

The purpose of this paper is to use a text-based sentiment indicator to explain variations in direct property market liquidity in the USA. By means of an artificial neural network, market sentiment is extracted from 66,070 US real estate market news articles from the *S&P Global Market Intelligence* database. For training of the network, a distant supervision approach utilizing 17,822 labeled investment ideas from the crowd-sourced investment advisory platform *Seeking Alpha* is applied. According to the results of autoregressive distributed lag models including contemporary and lagged sentiment as independent variables, the derived textual sentiment indicator is not only significantly linked to the depth and resilience dimensions of market liquidity (proxied by Amihud's 2002 price impact measure), but also to the breadth dimension (proxied by transaction volume). These results suggest an intertemporal effect of sentiment on liquidity for the direct property market. Market participants should account for this effect in terms of their investment decisions, and also when assessing and pricing liquidity risk. This paper not only extends the literature on text-based sentiment indicators in real estate, but is also the first to apply artificial intelligence for sentiment extraction from news articles in a market liquidity setting.

Keywords: Artificial Intelligence, Sentiment, Deep Learning, Commercial Real Estate, Market Liquidity, News Analytics

3.2 Introduction

Recent increases in media attention and public enthusiasm about the field of artificial intelligence might lead one to draw the incorrect conclusion that artificial neural networks (ANNs) are a new field of research. In fact, with Rosenblatt (1958) often being considered the inventor of the first "real" ANN, the theoretical foundations of deep learning methods are more than half a century old. Due to the vast computational requirements and lack of mathematical algorithms to support the concept, research efforts dried up soon after the initial suggestion of ANN approaches. Werbos' (1974) introduction of the seminal backpropagation algorithm certainly pushed the borders of efficiently training complex models. But only the possibility to accumulate massive amounts of exploitable data through the internet, and an exponential increase in available computational power during the last few decades, facilitated the recent renewal of interest in ANN research and applications.

A moderate number of studies employing deep learning in a real estate context have been published, although the majority of contributions addresses ANN-based property valuation (see e.g. Kathmann, 1993; Worzala, Lenk and Silva, 1995; Nguyen and Cripps, 2001; Din, Hoesli and Bender, 2001; Lam, Yu and Lam, 2008; Peterson and Flanagan, 2009; Poursaeed, Matera and Belongie, 2018). Apart from valuation research, ANN studies in the field of real estate are sparse. Ellis and Wilson (2005) suggest a portfolio-selection approach for the Australian property stock market, applying ANNs, Zhang, Gao, Seiler and Zhang (2015) use ANNs to identify real estate market cycles in China, and Chen, Chang, Ho and Diaz (2014) develop an ANN-approach for REIT return forecasts.

By introducing a deep learning-based approach to extract market sentiment from news articles, this study not only extends research on sentiment in real estate markets, but also the limited literature on investor sentiment as a factor explaining the variation in direct real estate market liquidity. Scholars such as Fisher, Gatzlaff, Geltner and Haurin (2003) and Clayton, Ling and Naranjo (2009) have pointed out the time-varying nature of direct real estate market liquidity compared to other asset classes. Empirically demonstrated during the last market cycle, the "ease" of trading increases during up-market periods, and decreases accordingly in down-markets. This feature of the property market is caused partially by the characteristics of real estate assets which are usually large-volume, heterogeneous and traded infrequently in segmented, local markets. However, in accordance with Liu (2015), who shows a relationship between sentiment and liquidity for the stock market, Freybote and Seagraves (2018) recently demonstrated the influence of market participants' sentiment as an additional driver of liquidity in the US office property market. The promising results of Freybote and Seagraves (2018) are a good entry point for additional research. In particular, testing a model incorporating shorter data-aggregation periods, the possibility of a lagged relationship and the use of refined sentiment measures seems worthwhile.

Hence, in this study, an ANN is trained on a data set collected from the investment advisory platform *Seeking Alpha* (SA). In a second step, the trained network is used to evaluate news articles from the *S&P Global Market Intelligence* database regarding their inherent sentiment. By averaging the sentiment scores of the news articles within each month of the study period, an aggregate index is calculated. The resulting monthly market sentiment indicator can then be analyzed for its influence on direct real estate market liquidity.

The chosen approach enables the extraction of a rich information structure from news articles, as ANNs do not rely on a predefined set of rules to indicate the sentiment polarity expressed by the respective article's author. Unlike conventional deep learning sentiment analysis, which requires the time-consuming and subjective practice of manually classifying a sufficiently large training data set, this paper furthermore applies distant supervision labeling. The available categorization of articles on the SA website into long- and short investment ideas is utilized as a natural indicator for the sentiment polarity prevailing in the respective text. By automatizing the sentiment-extraction procedure and making use of a vast online source for text data with a distinct sentiment polarity, the paper overcomes one of the limitations of existing ANN sentiment analysis. In addition, it might be a starting point for other studies exploiting alternative distant supervision-labeled data sources.

For the observation period from January 2006 to December 2018, the findings provide strong evidence of a dynamic link between sentiment and different dimensions of market liquidity. While there is a significant contemporary link for both of the two tested liquidity proxies, in the case of the market-depth proxy, sentiment leads market liquidity by up to more than two quarters. Market participants in the direct commercial real estate market seem to exhibit sentiment-induced behavior as a trigger of transaction decisions and by doing so, stimulate future market liquidity.

The remainder of this paper is structured as follows. The third section provides an overview of related literature, identifies existing research gaps and in this context outlines the motivation for this study. The fourth and fifth sections describe the data sets, the sentiment-extraction procedure, and the econometric approach used to estimate the results, which follow in the sixth section. The seventh section concludes.

3.3 Literature and Motivation

The properties of market **liquidity for the general stock market** have been the subject of extensive empirical research during the last few decades. Chordia, Roll and Subrahmanyam (2000) find a market-wide co-movement, Amihud (2002) shows an effect of market liquidity on returns, and Pastor and Stambaugh (2003) as well as Acharya and Pedersen (2005) provide empirical evidence supporting the existence of a systematic liquidity-risk factor. Compared to the effects of market liquidity on returns and asset prices, literature on the effects causing the marked-wide variation in liquidity is limited. Investor sentiment, as one relevant explanatory factor for market liquidity in the general stock market, was empirically analyzed by Liu (2015). However, the first theoretical foundations for the relationship were established by the seminal papers of Kyle (1985) and DeLong, Shleifer, Summers and Waldmann (1990), showing a connection between sentiment (i.e. bullishness or bearishness of investors), the resulting proportion of noise trading in the market and market liquidity, through the degree of market makers' price adjustment to order flow. Nevertheless, applying the framework of Kyle (1985) and DeLong, Shleifer, Summers and Waldmann (1990) to direct property markets poses some difficulties. No short-sale constraints exist in the models, so that noise traders increase trading both when sentiment is high and low. Additionally, the framework relies on the existence of perfect competition between market making agents, who unconditionally absorb the entire order flow. Both assumptions seem unrealistic in a direct property market setting. Baker and Stein (2004) suggest a model providing a better match for the peculiarities of the direct property market.[21] In their model, sentiment-driven investors underreact to information contained in the order flow. A higher share of such investors consequently results in a reduced price impact of trading. As a consequence of the lower price impact of trades in sentiment-driven market phases, insiders furthermore increase their trading activity and by doing so, boost trading volume in the market. In contrast to Kyle (1985)

[21] Baker and Stein (2004) explicitly suggest empirical tests of their model in real asset markets.

and DeLong, Shleifer, Summers and Waldmann (1990), Baker and Stein's (2004) model predicts higher liquidity only in phases of high sentiment. This one-directional behavior results from the introduction of short-sale constraints and provides a more realistic model setup particularly for a direct property market application. In an extension of their model, the authors additionally incorporate a higher propensity of sentiment-driven investors to churn their positions after receiving private signals, thus further stimulating trading volume in the market. This extension allows for an interesting empirical test of the direct property market. On the one hand, market imperfections are particularly strong in property markets compared to the highly efficient stock market, thus leaving additional space for contrary private signals. On the other hand, the high transaction fees in the property market might stifle this behavior. The answer to the question as to which effect prevails is insofar an empirical one.

The links between **sentiment and various dimensions of real estate markets** have been researched for more than two decades. In an early study, Barkham and Ward (1999) test the hypothesis of DeLong, Shleifer, Summers and Waldmann (1990) and find NAV discounts of listed property companies in the UK to be caused partially by price misperceptions of noise traders. Correspondingly, Gallimore and Gray (2002) verify the influence of sentiment in property decision making with a survey of UK property investment professionals. By using a survey-based sentiment measure from the *Real Estate Research Corporation (RERC)*, Das, Freybote and Marcato (2015) find a relationship between institutional investor sentiment and trading behavior in the US REIT market. The authors attribute their findings to a sentiment spillover effect between private and public real estate markets. In more recent years, *Google Trends* data was suggested as explaining REIT market movements and volatility (Rochdi and Dietzel, 2015; Braun, 2016) as well as forecasting the direct commercial real estate market in the USA (Dietzel, Braun and Schäfers, 2014). As an alternative sentiment proxy, text-based sentiment indicators also found their way into the real estate literature. Walker (2014) and Soo (2015) study the influence of media on housing markets in the UK and the USA. Ruscheinsky, Lang and Schäfers (2018) furthermore argue that sentiment extracted from newspaper headlines can be used as a leading indicator for the US REIT market.

The connection between sentiment and real estate market liquidity has, however, attracted little attention from scholars. The first paper to analyze the potential

relationship between sentiment and liquidity for the commercial real estate market is provided by Clayton, MacKinnon and Peng (2008). The authors examine several possible explanations of time variation in commercial real estate market liquidity. In a subsequent empirical analysis utilizing quarterly NCREIF data and a vector autoregression approach, they do not, however, find evidence of an influence of over-optimistic (noise) traders on market liquidity. In a related study, Freybote and Seagraves (2018) conduct a detailed analysis of the sentiment-liquidity relationship for the US office market, using Markov-switching models. Freybote and Seagraves (2018) use quarterly data for their analyses, employ activity (turnover) and market depth (Amihud) liquidity measures, and the *Real Estate Research Corporation (RERC)/Situs* survey as well as *Real Capital Analytics (RCA) buy-sell index (BSI)* data for their sentiment measures. They find that the relationship between sentiment and liquidity might be non-linear, with a larger impact of sentiment on turnover measures in phases of high liquidity, and a larger impact on the market depth dimension of liquidity in phases of low liquidity. The study furthermore shows that the effect of sentiment on liquidity varies for different investor types.

Despite the preceding investigation of Freybote and Seagraves (2018), this paper posits that additional insights can be gained from an analysis which refines several dimensions of previous work on the topic. First, despite the high quality of NCREIF data, quarterly analysis prevents a fine-grained examination of a potential mix of contemporary and lagged effects of sentiment on liquidity, due to its high degree of aggregation. It might be useful to break the effect down into its time-dependent components by adding a distributed lag structure into quantitative analyses. The rationale behind this approach lies in the specifics of the direct property market; Ametefe, Devaney and Marcato (2016) analyze the inefficiencies in direct property markets and among other factors, emphasize the decentralized structure of the market and the resulting, often time-consuming need to find a transaction counterparty. Together with long time frames to complete transactions (see Investment Property Forum, 2004; Scofield, 2013; Devaney and Scofield, 2015), sentiment-driven buy or sell decisions may merely influence market periods in the future. More specifically, Devaney and Scofield (2015) find, for a sample of UK property transactions from 2004 to 2013, that the mean time for a purchase (introduction to completion) is 144 days, and the mean time for a sale (marketing to completion) is 165 days.[22] With many of the transactions in Devaney and Scofield's sample terminating substantially faster or slower, a sufficiently long time period for the market-wide sentiment-liquidity relationship has to be considered. Accordingly, this paper analyzes the existence of an intertemporal relationship between real estate investor sentiment and direct market liquidity. The authors hope that by revising investor understanding of the timely structure of the relationship, decision-making processes in real estate transactions might be improved.

Second, the use of an alternative measure of real estate investor sentiment might strengthen the empirical power of the analyses. The use of text documents for sentiment determination allows for a more unmediated analysis, compared for example, to buy-sell indices, which constitute the aggregated results of potentially month-long transaction processes, possibly triggered initially by sentiment. Hence, this paper applies a novel news-based approach, and suggests a sentiment measure developed by means of a deep learning framework. More precisely, a multilayer perceptron is trained to distinguish between the degree of positive and negative sentiment in real estate news articles. Based on information extracted from training data, the application of AI reveals a rich information structure from news articles which might not only result in a superior sentiment indicator, but can also be applied to short aggregation periods. The obtained sentiment scores are, thus, used to create an index proxying overall investor sentiment in the US property market on a monthly basis. With the described approach, this paper extends the so far only AI-based sentiment-extraction attempt in real estate research of Hausler, Ruscheinsky and Lang (2018) by making use of ANNs and furthermore objectivizing the sentiment extraction procedure. In addition, the authors believe that the suggested use of distant supervision training data for sentiment analysis can advance sentiment research regardless of the field of application.

3.4 Data and Methodology

The paper makes use of several data sources. For the ANN training procedure, text data are gathered from the crowd-sourced financial content platform SA. The sentiment measure itself is based on the vast *S&P Global Market Intelligence* (S&P)

[22] Although Devaney and Scofield (2015) analyze the UK real estate transaction market, their conclusions should also apply to the US market, as both markets are highly developed.

news database. In order to construct the liquidity measures required for the regression analyses, both *CoStar* and RCA data are utilized. Finally, data required for several control variables are obtained from the webpage of the *Federal Reserve Bank of St. Louis* (FRED).

3.4.1 Sentiment Index

The chosen distant supervision approach for training the ANN requires a sufficient amount of financial text data with distinct, unambiguous sentiment polarity. SA, as a crowd-sourced platform providing investment information in its large long idea/short idea sections is well suited for this approach and has already found its way into academic research through an application as a news provision database for Chen, De, Hu and Hwang (2014). Each idea text contains the personal opinion of a freelance author on an equity or market, with long ideas suggesting a positive development of the equity or market in question and short ideas suggesting a negative development. Since 2014, SA's long and short ideas contain a summary section which outlines the essential reasoning of the text in a catchy and mostly undifferentiated fashion.[23] As those summary sections concisely cover the authors' positive or negative opinion on the equity or market in question, they serve as a reliable data source for isolating textual sentiment in a financial context. For the ANN's training process, a balanced sample of long and short summary sections containing 17,822 SA texts is collected.[24]

The text corpus for the sentiment index is obtained from the *S&P Global Market Intelligence* news database. S&P's news are widely used among real estate professionals and available in large quantities. Accordingly, it can be argued that the news articles' mean monthly polarity represents a reasonably accurate measure of the sentiment prevailing in the market for that month. In total, 66,070 US real estate market news articles for the study period between January 2006 and December 2018 serve as the study's textual sentiment sample. The monthly mean number of articles over the study period is 424, and the minimum amount is 224 articles per month.

To provide the ANN with comparable data for the training and prediction tasks, pre-processing steps have to be carried out both on the S&P and the SA text data sets (see

[23] An example from *Seeking Alpha's* long idea sample of this study is: "Newmont Mining's bottom line is improving rapidly, and a strong asset profile should improve its performance in the future." A representative short idea excerpt is: "MCD is at a critical juncture. All signs are pointing to a likely break lower."

[24] The sample consists of texts from 3,107 different freelance authors, the average length of each text is 381 characters.

e.g. Uysal and Gunal, 2014). Unicode categories of punctuation (P), symbols (S), separators (Z) and numbers (N), as well as intra-word contractions, are removed. Words are converted to lower case, tokenized and stemmed, using Porter's (1980) algorithm for suffix stripping. With respect to stop-word removal, this study starts with a common list of English stopwords and extends that list with written numbers and calendar terminology. This method avoids any unintended association of sentiment with certain date or time expressions. As a further extension, the training and estimation data sets are compared to a full list of written English vocabulary. By excluding non-standard words (e.g.: company and executive names), a false association of those words with positive (negative) sentiment resulting from their incidence in SA's long (short) ideas can be avoided. For this task, the widely used *Hunspell* spell-checking dictionary is employed.[25]

Next, SA investment ideas are annotated with a distant supervision label of 0 if they are from the short idea category, and 1 if they are from the long idea category. A sparse matrix based on the 1,000 most frequent words of the SA training data is computed, in order to one-hot-encode the S&P and SA data sets. In this way, textual documents are expressed as binary vectors and thus interpretable by the neural network. Note that the use of embedding layers and a larger word corpus were tested, but did not increase performance.

This study uses a random sample of 80 percent of the 17,822 one-hot encoded SA texts for the training of the sentiment estimation ANN. The remaining 20 percent are set aside for out-of-sample validation and comparison of alternative network setups.

The final ANN contains four fully connected layers with a declining node amount of 64, 48, 32 and 16 nodes per layer. The four layers make use of ReLU (Rectified Linear Unit) activation functions. The reduction of nodes per layer is applied in order to gradually reduce the complexity of the feature space. In formal terms, each of the ReLU layers processes data according to the equation:

$$max(0, dot(Input, W) + b)\,, \tag{3.1}$$

[25] This paper uses the default *Hunspell* list with common word spelling. The list including British as well as American spelling, and also, diacritic and non-diacritic marks, was derived from http://app.aspell.net/create.

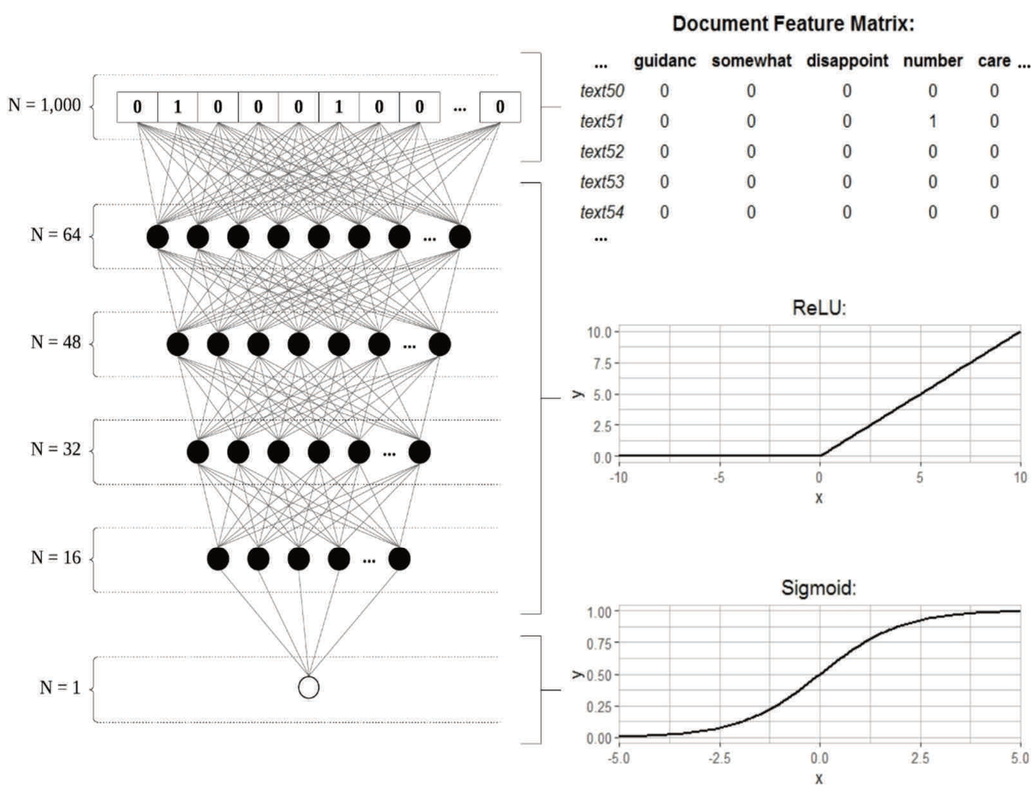
where *Input* denotes one-hot encoded textual data in the form of a tensor of rank 2. W and b are the trainable weight tensors of the respective layer.[26]

While initially-set ANN weights are random, the training process carries out a step-wise weight adjustment based on a feedback signal. This is provided by the combination of a sigmoid layer and a loss function. The sigmoid squashing function, as the last layer of the ANN, compresses output values into the spectrum between 0 and 1 and thus provides a label prediction $\hat{y}_k$ for each textual document:

$$\hat{y}_k = \frac{1}{1+e^{-t}} \; with \; t = dot(Input, W) + b \,. \qquad (3.2)$$

Figure 3.1 provides an overview of the conceptual layout of the multilayer perceptron developed for this paper.

Figure 3.1: ANN Layout



Notes: Figure 3.1 shows the conceptual layout of the multilayer perceptron. Based on the 1,000 most frequent words in the *Seeking Alpha* training sample, articles from the *S&P Global Intelligence* database are expressed in the form of a document feature matrix. This matrix is processed by four fully connected ReLU layers with a decreasing number of nodes. The final node provides a sentiment score for each news article, ranging from 0 (negative) to 1 (positive), by using a sigmoid activation function.

[26] All equations describing the ANN setup skip subscripts for ease of demonstration.

The network's overall classification error (or prediction loss) L is calculated via binary cross-entropy, i.e. by comparing $\hat{y}_k$ to the true binary distant supervision label value y_k for each textual document k:

$$L = \frac{1}{n}\sum_{k=1}^{n} -1 * \left(y_k * log(\hat{y}_k) + (1 - y_k) * log(1 - \hat{y}_k)\right). \quad (3.3)$$

SA texts are fed into the ANN in batches of 500, and after each batch, the prediction loss L is calculated and backpropagated through the network, using Root Mean Square Propagation (RMSprop) as the optimizer algorithm (Tieleman and Hinton, 2012). 6 epochs, each containing all batches, are performed. Hence, weights W and b undergo a total number of 174 updates specified by the equations:

$$v_{dW}(t) = \beta * v_{dW}(t-1) + (1-\beta) * \left(\frac{\partial L}{\partial W}(t)\right)^2$$

$$v_{db}(t) = \beta * v_{db}(t-1) + (1-\beta) * \left(\frac{\partial L}{\partial b}(t)\right)^2$$

$$\Delta W(t) = -\frac{\eta}{\sqrt{v_{dW}(t) + \varepsilon}} * \frac{\partial L}{\partial W}(t) \quad (3.4)$$

$$\Delta b(t) = -\frac{\eta}{\sqrt{v_{db}(t) + \varepsilon}} * \frac{\partial L}{\partial b}(t),$$

where $v_{dW}(t)$ is the moving average of the squared gradient of W at time t, and $v_{db}(t)$ the squared gradient of b at time t. η defines the optimizer's learning rate (set to 0.001 for this paper) and β is a hyperparameter defining the influence of past gradient updates (here, the value of β is set to 0.9, as suggested by Tieleman and Hinton, 2012). ε is a fuzz factor to avoid division by zero; in this paper the value is set to e^{-7}.

The described ANN model is trained independently ten times, and for each trained model, a sentiment score for each document in the S&P data set is estimated. After calculating the mean sentiment score for each individual document, the mean scores of all documents published in a respective month are calculated. For the study period between January 2006 and December 2018, the resulting average monthly sentiment score (*SM*) is 0.63, and the standard deviation 0.05.

3.4.2 Liquidity Proxies

In their analysis of the literature on liquidity in financial markets, Ametefe, Devaney and Marcato (2016) identify the five liquidity dimensions of *tightness*, *depth*, *resilience*, *breadth* and *immediacy*. The authors describe *tightness* as the "the cost of trading even in small amounts", *depth* as the "capacity to sell/buy without causing price movements", *resilience* as "the speed at which the marginal price impact increases as trading quantities increase", *breadth* as "the overall volume traded", and *immediacy* as "the cost (discount/premium) to be applied when selling/buying quickly". Although several proxies for each dimension exist for indirect financial markets, measurement for direct property markets is cumbered by limited data availability and conceptual differences between both markets. For the *tightness* dimension of liquidity, Ametefe, Devaney and Marcato suggest several bid-ask spread proxies, although for the direct property market, these proxies are unavailable.[27] For the fifth dimension, namely *immediacy*, Ametefe, Devaney and Marcato merely suggest real estate time on market as a proxy. To depict this dimension, a representative data set of time-on-market information would be required. Due to the unavailability of such data, this study focuses on a representation of the remaining dimensions *depth*, *resilience* and *breadth* of the US direct property market. Therefore, Amihud's (2002) widely used liquidity proxy (see e.g. Brounen, Eichholtz and Ling, 2009; Glascock and Lu-Andrews, 2014; Freybote and Seagraves, 2018) is used to cover the dimensions *depth* and *resilience*. The measure is calculated as:[28]

$$AMI_t = log\left(\frac{|R_t|}{Vol_t}\right). \tag{3.5}$$

AMI_t captures the absolute value of the price impact (R) of one billion USD transaction volume (VOL) for month t. For the denominator Vol_t, RCA's monthly data on US commercial direct real estate transaction volume is obtained.[29] The numerator is the absolute value of the return on the *CoStar Commercial Repeat-Sale Index* for month

[27] The conversion of Ametefe, Devaney and Marcato's (2016) tightness proxy *relative quoted spread* to a direct real estate market use case is theoretically possible, but requires the use of a private data set containing the bid and ask prices of property transactions.

[28] This paper follows the methodology of Amihud's (2002) paper, and takes the natural logarithm of the proxy. The denominator of the proxy is furthermore adjusted for inflation of the transaction volume over time, by scaling it with the consumer price index for the USA.

[29] RCA collects data on transactions of volume USD 2.5 million or greater.

t.[30] Besides its ease of calculation, the application of the Amihud measure furthermore allows for a test of Baker and Stein's (2004) hypothesis of a negative relationship between sentiment and price impact.

The second liquidity measure in this study is suggested by Ametefe, Devaney and Marcato (2016) for their fourth liquidity dimension, *breadth*. The measure VOL_t is the transaction volume of the direct US property market for month *t* in billion USD.[31] By incorporating trading volume into the analysis, Baker and Stein's (2004) supposed positive relationship to sentiment can be examined. A case for volume-based measures of liquidity can be made through their links to easier market-access and lower transaction costs (see e.g. Demsetz, 1968; Glosten and Milgrom, 1985). Monthly transaction volume data for this study are again obtained from RCA.

3.4.3 Control Variables

In order to control for the effect of other factors potentially explaining variation in direct property market liquidity, a set of control variables is added into the regression analyses. Liu (2015) considers the possibility that sentiment merely captures macroeconomic conditions. For this reason, this paper controls for general economic conditions as a factor explaining liquidity. *UNRATE* and *CPI* are the seasonally adjusted civilian unemployment rate and the consumer price index for all urban consumers, respectively. *BAA10YM*, which is the spread between the yield on Moody's seasoned Baa corporate bonds and 10-year treasury constant maturity bonds, represents general economic default risk. Together with *UNRATE* and *CPI*, *BAA10YM* is intended to proxy for the state of the economy. Liu (2015) furthermore adds into his regressions several variables reflecting the general stock market. This paper accordingly controls for the state of the direct property market. The supply side of the direct property market is allowed for by adding seasonally adjusted total construction spending in the USA (*CONST*) in billion USD. In addition, the development of the US direct property market is included in the regressions by adding returns of the *CoStar*

[30] RCA also provides a transaction-based monthly direct real estate index for the US market; however, the construction methodology of the index leads to an untenable level of autocorrelation which inevitably causes problems in any subsequent quantitative analysis.

[31] Turnover, as a generally preferable proxy for market breadth, compared to transaction volume, can only be calculated if the asset universe is defined (e.g. for the NCREIF Index, turnover data are available). This study analyzes monthly time series and uses RCA data, for which no turnover measure is available.

Commercial Repeat-Sale Index (*CCRSI*).[32] Descriptive statistics for the liquidity, sentiment and control variables for the study period between January 2006 and December 2018 can be found in Table 3.1.

Table 3.1: Descriptive Statistics

Statistic	Mean	Median	St. Dev.	Min	Max
SM	0.63	0.63	0.05	0.49	0.73
*AMI (*1000)*	0.83	0.28	1.86	0.01	15.30
VOL (bn USD)	31.94	33.75	16.63	3.64	79.29
PROPS	1,876.28	2,061.00	794.04	391.00	3,651.00
UNITS (bn)	0.13	0.13	0.07	0.02	0.42
CCRSI (%)	0.03	0.03	0.01	0.02	0.06
BAA10YM (pp) (%)	2.69	2.66	0.84	1.56	6.01
CONST (bn USD)	1,038.44	1,064.51	168.45	754.71	1,324.35
CPI (%)	0.16	0.17	0.39	-1.92	1.01
UNRATE (%)	6.37	5.65	1.99	3.70	10.00

Notes: Table 3.1 reports summary statistics of the constructed sentiment measure *SM* as well as four different proxies of direct real estate market liquidity. *AMI* covers the liquidity dimensions *depth* and *resilience*. For better interpretability, *AMI* is displayed without the CPI-adjustment of the denominator or the log transformation and is furthermore multiplied by 1,000. *VOL* represents the market *breadth* dimension and is depicted in bn USD. As alternatives, *PROPS* reflects the number of properties and *UNITS* the number of units traded in a respective month (see the results chapter for details on *PROPS* and *UNITS*). *CCRSI* are monthly returns of the *CoStar Commercial Repeat-Sale Index* and *BAA10YM* is the spread between *Moody's* seasoned Baa corporate bond yields and the yield of 10-year constant maturity treasury bonds in percentage points (pp). *CONST* (in bn USD) and *CPI* are seasonal-adjusted total construction spending and the consumer price index for all urban customers, respectively. *UNRATE* measures seasonal-adjusted unemployment rate. The sample period is 2006:M01-2018:M12.

3.5 Regression Analysis

Given that this paper seeks to respect the slow nature of processes in the direct property market and, thus, to split up the effect of sentiment on liquidity into its contemporary and lag components, the empirical models require the addition of distributed lag terms. An analysis of the liquidity measures utilized in this paper furthermore reveals a strong negative auto-correlation.[33] For this reason, regression analysis requires the addition of autoregressive terms (i.e. lagged liquidity variables). The use of both distributed lags as well as autoregressive components provides the application of autoregressive distributed lag (ARDL) models. By including the dependent variable besides other

[32] Variables proxying the US general stock market or the REIT market (i.e. the S&P 500 and the NAREIT index) were tested as additional control variables. However, the chosen lag selection methodology described in the "Regression Analysis" section rejected their inclusion for the main model containing Amihud's (2002) measure for liquidity as the dependent variable. The same applies to the federal funds rate and a disposable income control variable, which have also been tested.

[33] The empirical explanation of the negative serial correlation lies in the existence of several months in the study period which exhibit an extraordinarily high transaction volume, followed by periods with very low volumes. This pattern most probably exists due to a market dry up effect after periods of particularly strong transaction activity.

explanatory variables as regressors, ARDL models allow a simultaneous examination of a potential long- and short-run relationship between market liquidity, sentiment and macroeconomic controls. ARDL models have gained attention particularly through the work of Pesaran and Shin (1998) and Pesaran, Shin and Smith (2001) on cointegrating relationships. In formal terms, the following equation represents the applied model:

$$LIQ_t = \alpha_0 + \sum_{i=1}^{I} \alpha_i LIQ_{t-i} + \sum_{j=0}^{J} \beta_j SM_{t-j} + \sum_{k=1}^{K} \sum_{l_k=0}^{L} \gamma_{k,l_k} x_{k,t-l_k} + \sum_{m=2}^{12} \delta_m Month_m + \varepsilon_t \,, \tag{3.6}$$

where LIQ_t is a measure of market liquidity in period *t* (i.e. AMI or VOL), SM_{t-j} the ANN-based sentiment indicator, $x_{k,t-l_k}$ the macroeconomic controls, $Month_m$ a set of monthly dummy variables and ε_t a random disturbance term. Running augmented Dickey-Fuller tests shows that some variables are stationary in levels (i.e. I(0)), while others are integrated of order 1. Thus, to secure unbiased and consistent estimates, the research framework has to account for a potential cointegrating relationship. By estimating Equation (3.6) in first differences and including the first lag of all regressors in levels, an unconstrained error correction model is derived. Subsequently, the bound-testing procedure of Pesaran, Shin and Smith (2001) is executed. In the event of a long-run relationship, the OLS residual series of the long-run cointegrating regression $y_t = \alpha_0 + \delta Sentiment_{t-1} + \sum_{k=1}^{K} \theta_k x_{k,t-1} + u_t$ must be added to the model to allow unbiased and consistent estimation. Bound-testing, however, reveals no evidence of a long-run relationship, so that each series of Equation (3.6) is differenced once, and coefficients are derived using standard OLS.

Considering Devaney and Scofield's (2015) results for direct real estate transaction periods, liquidity measures and the sentiment indicator are included in a fixed lag of up to nine months in the OLS models, so as to provide a complete picture of the relationship up to three quarters in the past ($I = J = 1, \dots, 9$). The appropriate lag structure for each macroeconomic control variable is derived analytically, by running all possible continuous lag combinations and choosing the optimal structure based on the minimal Akaike Information Criterion (AIC).[34]

[34] For this purpose, the maximum lag for the control variables was set to 6 and in total, 32,768 models were tested and ranked by AIC.

3.6 Results

Ordinary least squares estimation of Equation (3.6) in first differences leads to the results shown in columns (1) and (2) of Table 3.2. The results in column (1) exhibit an OLS regression using *AMI* as the dependent variable, column (2) the results for *VOL*.

As expected, the autoregressive lag terms in both the *AMI* and *VOL* regressions display a strong negative serial correlation, with coefficients significant at the 1 percent level for the first three lags in the *AMI* regression and two lags in the *VOL* regression. This finding is most probably caused by the drying up of the direct real estate market after periods of very high increases in transaction volumes, which effects *VOL* directly and *AMI* indirectly through the lower denominator value during periods after such "high volume" periods.

Table 3.2: Liquidity - Sentiment: Autoregressive Distributed Lag Models

	Dependent variable:			
	(1) AMI		(2) VOL	
C	0.663		-5.580	
	(0.406)		(3.685)	
SM	-12.645	***	62.042	*
	(4.317)		(35.239)	
SM (-1)	-12.178	**	31.848	
	(5.518)		(43.690)	
SM (-2)	-13.665	**	87.700	*
	(6.063)		(46.613)	
SM (-3)	-5.932		9.342	
	(5.611)		(43.719)	
SM (-4)	-13.729	***	23.936	
	(5.182)		(40.286)	
SM (-5)	-7.946		32.924	
	(4.973)		(39.023)	
SM (-6)	-10.629	**	25.857	
	(4.858)		(38.797)	
SM (-7)	-12.904	***	53.007	
	(4.879)		(39.984)	
SM (-8)	-8.046	*	11.822	
	(4.657)		(37.822)	
SM (-9)	-6.095		28.955	
	(4.057)		(33.929)	
AMI (-1)	-0.489	***		
	(0.103)			
AMI (-2)	-0.387	***		
	(0.113)			
AMI (-3)	-0.565	***		
	(0.111)			

(Table continues on next page.)

Table 3.2: Liquidity - Sentiment: Autoregressive Distributed Lag Models (cont.)

	(1)		(2)	
AMI (-4)	-0.206	*		
	(0.116)			
AMI (-5)	-0.157			
	(0.113)			
AMI (-6)	-0.223	*		
	(0.113)			
AMI (-7)	-0.163			
	(0.108)			
AMI (-8)	-0.165			
	(0.106)			
AMI (-9)	0.019			
	(0.093)			
VOL (-1)			-0.837	***
			(0.095)	
VOL (-2)			-0.555	***
			(0.117)	
VOL (-3)			-0.298	**
			(0.118)	
VOL (-4)			-0.250	**
			(0.118)	
VOL (-5)			-0.202	
			(0.124)	
VOL (-6)			-0.115	
			(0.126)	
VOL (-7)			-0.162	
			(0.133)	
VOL (-8)			0.056	
			(0.124)	
VOL (-9)			0.035	
			(0.103)	
Macroecon. controls	YES		YES	
Month dummies	YES		YES	
Observations	146		146	
R^2	0.613		0.759	
Adjusted R^2	0.376		0.620	
Residual std. error	0.933		8.115	
F-statistics	2.588	***	5.473	***
Significance levels:			**p<0.1; **p<0.5; ***p<0.01*	

Notes: Table 3.2 reports results of the first-difference autoregressive distributed lag (ARDL) models analyzing the relationship between the constructed sentiment index (*SM*) and two different liquidity proxies. Column (1) shows the coefficients of the regression using Amihud's (2002) measure for illiquidity (*AMI*), representing the price impact of transaction volume. Column (2) shows the coefficients for transaction volume (*VOL*). The contemporary value and 9 lags of *SM* were used together with 9 autoregressive terms of either *AMI* or *VOL* in both regressions. The *AMI* (*VOL*) regression furthermore contains an intercept, month dummies, as well as 5 (2) lags for the spread between the yield on Moody's seasoned Baa corporate bonds and 10-year treasury constant maturity bonds (*BAA10YM*), 5 (5) lags of seasonally adjusted construction spending (*CONST*), 5 (2) lags of consumer price index for all urban consumers (*CPI*), 4 (4) lags for the *CoStar Commercial Repeat-Sale Index* (*CCRSI*) and 1 (6) lag(s) for the seasonally adjusted civilian unemployment rate (*UNRATE*). Macroeconomic controls and month dummies are not displayed. Standard errors are included in brackets underneath their respective coefficient. The sample period is 2006:M11-2018:M12.

For the regression containing first differences in Amihud (*AMI*) as a proxy for the depth and resilience dimensions of liquidity as the dependent variable, the contemporary value as well as several lags of the sentiment measure (*SM*) are highly significant in explaining market liquidity changes. Specifically, parameters of the contemporary sentiment value and lags 1, 2, 4, 6 and 7 are significant at least at the 5 percent level, and lag 8 is furthermore still significant at the 10 percent level. All sentiment coefficient values have the expected negative sign, indicating a negative contemporary and lagged relationship between increases in sentiment and increases in *AMI*. This observation supports the hypothesis of an intertemporal relationship between the two variables, resulting from long transaction periods and the generally slow pace of direct property markets. Recalling Devaney and Scofield's (2015) results, the significance pattern of *SM* apparently tracks the pattern of times to completion, with around 87 percent of the property purchases and 86 percent of the property sales transactions requiring a time period of no more than 239 days. The effect of sentiment on liquidity, thus, seems to trickle into the market over an extended period, possibly caused initially by sentiment-induced behavior of market participants. The empirical results furthermore support the theoretically derived relationship between sentiment and price impact suggested by Baker and Stein (2004).

OLS estimation, using differences in transaction volume (*VOL*) as a proxy for the breadth dimension of liquidity, displays similar, but weaker results. All sentiment parameters show a positive relationship with differences in volume, although only the contemporary volume parameter and the second lag are significant at the 10 percent level. These results suggest that positive (negative) sentiment stimulates (stifles) the overall amount of trading, but that the effect of sentiment on price impact (i.e. *AMI*) exceeds the effect of trading volume (i.e. *VOL*).[35] However, the significance of the second lag of *VOL* suggests an intertemporal relationship between sentiment and the breadth dimension of liquidity as well. The increase in market breadth appears to manifest itself partially in future periods, arguably due to the slow transaction process in direct property markets. A possible reason for the weak effect of sentiment on trading volume could lie in the high transaction costs in direct property markets, which

[35] Note that the *VOL* and *AMI* regressions are not perfectly comparable, due to the different lag lengths of the control variables.

moderate the effect of sentiment that the extended model of Baker and Stein (2004) posits on trading volume.

To ensure the robustness of the regression results, several diagnostics tests have been performed. To identify potential auto-correlation problems in the regression residuals, a Breusch-Godfrey test was conducted. While there is no evidence of first order autocorrelation, with the inclusion of residuals up to 9 lags, there is some evidence of serial correlation at the 10 percent level. For this reason, the regressions are re-estimated, using Newey-West standard errors (Newey and West, 1987). The results remain basically unchanged. Furthermore, a Breusch-Pagan test is performed, indicating no heteroscedasticity problems in the regression residuals for both regressions. CUSUM and CUSUM square analyses confirm the stability of the estimated models.

A possible explanation of the strong relationship between sentiment (*SM*) and Amihud (*AMI*) could result from the denominator of the measure. Liu (2015) notes that the effect through the division by trading volume might drive a strong relationship between sentiment and Amihud. To test this possibility, a model including contemporary *VOL* as well as 9 lags is estimated. The untabulated results show a strengthened effect of *SM* on *AMI*.

As an additional robustness check, alternative liquidity measures are tested. Instead of differences in transaction volume, differences in the absolute number of traded properties (*PROPS*) and the number of traded units (*UNITS*) are used in the regressions. The lag structure for the control variables is again determined by AIC. The results can be obtained from columns (1) and (2) of Table 3.3 in the appendix. Contemporary sentiment in the *PROPS* regression shows a positive parameter value which is furthermore significant at the 1 percent level. The second lag of *PROPS* is also positive and significant at the 10 percent level. *UNITS* is significant and positive at the 5 percent level for the contemporary variable. The structure of significant lags of *PROPS* is, thus, similar to the structure of *VOL*, which is not surprising considering the similarity in the construction of the measures. The *UNITS* regression does not exhibit an intertemporal effect of sentiment on liquidity.

Overall, the results provide strong evidence of an intertemporal relationship between sentiment and liquidity. The effect seems to be persistent in particular for the depth and resilience dimensions of liquidity, as proxied by *AMI*. Market participants in the direct commercial real estate market seem to display sentiment-induced behavior as a trigger for a transaction. Due to the long transaction periods, the effect of sentiment on liquidity, however, only manifests itself gradually over the following months.

3.7 Conclusion

This paper introduces a novel approach to the construction of a sentiment index for the US real estate market. The approach is text-based and relies on the application of an ANN for sentiment estimation. Highly sentiment-loaded text documents from the crowd-sourced investment content provider SA serve as a distant supervision-labeled data set and are used to train an ANN to discriminate between positive and negative sentiment. The trained network is then employed to predict the polarity of real estate news articles from the broadly used *S&P Global Market Intelligence* news database for the time period between January 2006 and December 2018. By doing so, and through aggregating monthly polarity scores of the single articles, a monthly real estate sentiment index is designed. In a next step, the ability of the sentiment index to explain changes in liquidity in the direct commercial US real estate market is examined. The slow pace of direct real estate markets, implying long search periods for both sellers and buyers, and complex transaction processes (see e.g. Investment Property Forum, 2004) suggest that an effect of sentiment on liquidity might manifest itself in a lagged fashion. This potential lagged effect is accounted for in the regression models of this paper. Furthermore, an increase in the time series frequency compared to existent studies from quarterly to monthly data enables a more fine-grained analysis of the sentiment-liquidity relationship. The liquidity proxies in this study have been selected in order to capture several dimensions of market liquidity, namely, the depth, resilience and breadth of the market. In this respect, Amihud's (2002) price-impact measure is used as the first proxy to represent market depth and resilience. Transaction volume, as the second proxy, is chosen to represent market breadth. In order to examine the hypothesis of a lagged relationship between sentiment and liquidity empirically, autoregressive distributed lag models are constructed. OLS estimation yields strong evidence supporting an intertemporal relationship between the applied measure of sentiment and the depth and resilience dimensions of liquidity. Regressions show several significant lags for Amihud, which range up to order 8. The intertemporal relationship between sentiment and the breadth dimension of liquidity is somewhat weaker, but with a significant second lag still present.

Despite the high level of aggregation of this current study, useful insights are gained. As pointed out by Freybote and Seagraves (2018), through understanding the dynamics of market sentiment and liquidity, investors may assess and price liquidity risk more appropriately and accurately. The present paper adds a new dimension to this understanding. In the direct property market, contemporary liquidity measures are merely the realization of transactions initiated several months in the past. Investors might thus want to attach secondary importance to those "stale" measures. They might instead want to keep in view the factors determining liquidity measures in the future. This paper analyzes one of those potential factors, namely, real estate market sentiment, and finds a significant influence. Considering the temporal link between investor sentiment and market liquidity can thus help investors to plan transaction decisions ahead, invest or divest preemptively and in this way gain a competitive edge compared to those relying on a direct observation of liquidity. In particular, institutional investors with high transaction rates and, therefore, depending on high market liquidity (e.g. pooled funds, property companies, see e.g. Investment Property Forum, 2004) can benefit from focusing on sentiment indicators instead of delayed liquidity figures.

However, the study is limited by a couple of data constraints. Shortcomings result from the unavailability of real-estate-specific training data and the lack of liquidity proxies for two of the dimensions, as outlined by Ametefe, Devaney and Marcato (2016). The used training texts focus on stocks in general and not on real estate in particular. While the results are promising, especially for the depth and resilience dimensions, real-estate-specific training data could provide further insights and possibly a more precise estimation. Future research might also benefit from the use of data sets which allow for the construction of alternative liquidity proxies representing the dimensions of tightness and immediacy, in order to gain deeper knowledge of the sentiment-liquidity relationship.

Another constraint of the study is that investing in the US commercial real estate market as a whole is only possible in theory. At first glance, this limitation restricts the practical usefulness of the approach for real estate investors. However, the proposed training and sentiment-extraction procedure can be applied to specific property types and regional markets. After collecting additional data from regionally and thematically specific news sources, it should be particularly interesting to analyze the interactions between the geographical and thematic layers of sentiment. The question of whether regional or global, or rather universal or topic-specific news dominate the sentiment of market participants might be worth analyzing empirically.

In this regard, this study facilitates more granular sentiment research in the future. In the context of AI-based sentiment analysis, the authors believe that there is vast potential for future research and practical applications. With the collection of a broader range of distant supervision-labeled data and an increased amount of news constituting the sentiment index, a more complete picture of real estate market sentiment could be obtained. Furthermore, this picture could ultimately be completed by an analysis of the influence of text-based sentiment on rents, cap rates and volatility.

3.8 Appendix

Table 3.3: Liquidity - Sentiment: ARDL Models with Alternative Liquidity Measures

	Dependent variable:			
	(1) Props		(2) Units (bn)	
C	-273.408		-0.007	
	(166.305)		(0.020)	
SM	5,668.750	***	0.441	**
	(1,416.210)		(0.172)	
SM (-1)	2,280.774		0.086	
	(1,737.298)		(0.209)	
SM (-2)	3,753.825	*	0.046	
	(1,909.511)		(0.223)	
SM (-3)	-2,398.345		-0.130	
	(1,871.445)		(0.222)	
SM (-4)	1,388.745		0.272	
	(1,822.705)		(0.209)	
SM (-5)	-1,320.822		-0.030	
	(1,702.090)		(0.195)	
SM (-6)	320.037		0.040	
	(1,671.634)		(0.191)	
SM (-7)	24.047		0.077	
	(1,617.013)		(0.190)	
SM (-8)	-1,400.794		-0.089	
	(1,451.007)		(0.182)	
SM (-9)	-382.467		0.028	
	(1,301.999)		(0.166)	
PROPS (-1)	-0.768	***		
	(0.094)			
PROPS (-2)	-0.546	***		
	(0.115)			
PROPS (-3)	-0.279	**		
	(0.119)			

(Table continues on next page.)

Table 3.3: Liquidity - Sentiment: ARDL Models with Alternative Liquidity Measures (cont.)

	(1)		(2)	
PROPS (-4)	-0.441	***		
	(0.120)			
PROPS (-5)	-0.334	***		
	(0.123)			
PROPS (-6)	-0.171			
	(0.119)			
PROPS (-7)	0.025			
	(0.121)			
PROPS (-8)	0.039			
	(0.117)			
PROPS (-9)	-0.068			
	(0.090)			
UNITS (-1)			-0.849	***
			(0.087)	
UNITS (-2)			-0.650	***
			(0.108)	
UNITS (-3)			-0.468	***
			(0.118)	
UNITS (-4)			-0.353	***
			(0.123)	
UNITS (-5)			-0.216	*
			(0.127)	
UNITS (-6)			-0.210	*
			(0.123)	
UNITS (-7)			-0.089	
			(0.125)	
UNITS (-8)			-0.061	
			(0.117)	
UNITS (-9)			0.056	
			(0.091)	
Macroecon. controls	YES		YES	
Month dummies	YES		YES	
Observations	146		146	
R^2	0.835		0.781	
Adjusted R^2	0.728		0.665	
Residual std. error	316.822		0.041	
F-statistics	7.812	***	6.769	***
Significance levels:			**p<0.1; **p<0.5; ***p<0.01*	

Notes: Table 3.3 reports results of the first-difference autoregressive distributed lag (ARDL) models analyzing the relationship between the constructed sentiment index (*SM*) and two alternative liquidity proxies. Column (1) shows the coefficients of the regression using the total number of traded properties (*PROPS*) as the dependent variable. Column (2) shows the coefficients for the total number of traded units in bn (*UNITS*) used as the dependent variable. The contemporary value and 9 lags of *SM* were included together with 9 autoregressive terms of either *PROPS* or *UNITS* in both regressions. The *PROPS (UNITS)* regression furthermore contains an intercept, month dummies, as well as 6 (2) lags for the spread between the yield on Moody's seasoned Baa corporate bonds and 10-year treasury constant maturity bonds (*BAA10YM*), 5 (5) lags of seasonally adjusted construction spending (*CONST*), 5 (2) lags of consumer price index for all urban consumers (*CPI*), a contemporary value of the *CoStar Commercial Repeat-Sale Index* (*CCRSI*) and 6 (6) lags for the seasonally adjusted civilian unemployment rate (*UNRATE*). Macroeconomic controls and month dummies are not displayed. Standard errors are included in brackets underneath their respective coefficient. The sample period is 2006:M11-2018:M12.

3.9 References

Acharya, V. V. and L. H. Pedersen, Asset pricing with liquidity risk, Journal of Financial Economics, 2005, 77:2, 375–410.

Ametefe, F., S. Devaney and G. Marcato, Liquidity: A review of dimensions, causes, measures, and empirical applications in real estate markets, Journal of Real Estate Literature, 2016, 24:1, 1–29.

Amihud, Y., Illiquidity and stock returns: Cross-section and time-series effects, Journal of Financial Markets, 2002, 5:1, 31–56.

Baker, M. and J. C. Stein, Market liquidity as a sentiment indicator, Journal of Financial Markets, 2004, 7:3, 271–99.

Barkham, R. J. and C. W. R. Ward, Investor sentiment and noise traders: Discount to net asset value in listed property companies in the U.K., Journal of Real Estate Research, 1999, 18:2, 291–312.

Braun, N., Google search volume sentiment and its impact on REIT market movements, Journal of Property Investment & Finance, 2016, 34:3, 249–62.

Brounen, D., P. Eichholtz and D. Ling, The liquidity of property shares: An international comparison, Real Estate Economics, 2009, 37:3, 413–45.

Chen, H., P. De, Y. Hu and B.-H. Hwang, Wisdom of crowds: The value of stock opinions transmitted through social media, Review of Financial Studies, 2014, 27:5, 1367–403.

Chen, J.-H., T.-T. Chang, C.-R. Ho and J. F. Diaz, Grey relational analysis and neural network forecasting of REIT returns, Quantitative Finance, 2014, 14:11, 2033–44.

Chordia, T., R. Roll and A. Subrahmanyam, Commonality in liquidity, Journal of Financial Economics, 2000, 56:1, 3–28.

Clayton, J., D. C. Ling and A. Naranjo, Commercial real estate valuation: Fundamentals versus investor sentiment, The Journal of Real Estate Finance and Economics, 2009, 38:1, 5–37.

Clayton, J., G. MacKinnon and L. Peng, Time variation of liquidity in the private real estate market: An empirical investigation, Journal of Real Estate Research, 2008, 30:2, 125–60.

Das, P. K., J. Freybote and G. Marcato, An investigation into sentiment-induced institutional trading behavior and asset pricing in the REIT market, The Journal of Real Estate Finance and Economics, 2015, 51:2, 160–89.

DeLong, J. B., A. Shleifer, L. H. Summers and R. J. Waldmann, Noise trader risk in financial markets, Journal of Political Economy, 1990, 98:4, 703–38.

Demsetz, H., The cost of transacting, The Quarterly Journal of Economics, 1968, 82:1, 33–53.

Devaney, S. and D. Scofield, Liquidity and the drivers of search, due diligence and transaction times for UK commercial real estate investments, Journal of Property Research, 2015, 32:4, 362–83.

Dietzel, M. A., N. Braun and W. Schäfers, Sentiment-based commercial real estate forecasting with Google search volume data, Journal of Property Investment & Finance, 2014, 32:6, 540–69.

Din, A., M. Hoesli and A. Bender, Environmental variables and real estate prices, Urban Studies, 2001, 38:11, 1989–2000.

Ellis, C. and P. Wilson, Can a neural network property portfolio selection process outperform the property market?, Journal of Real Estate Portfolio Management, 2005, 11:2, 105–21.

Fisher, J., D. Gatzlaff, D. Geltner and D. Haurin, Controlling for the impact of variable liquidity in commercial real estate price indices, Real Estate Economics, 2003, 31:2, 269–303.

Freybote, J. and P. A. Seagraves, The impact of investor sentiment on commercial real estate market liquidity, Journal of Real Estate Research, 2018, 40:4, 597–627.

Gallimore, P. and A. Gray, The role of investor sentiment in property investment decisions, Journal of Property Research, 2002, 19:2, 111–20.

Glascock, J. and R. Lu-Andrews, An examination of macroeconomic effects on the liquidity of REITs, The Journal of Real Estate Finance and Economics, 2014, 49:1, 23–46.

Glosten, L. R. and P. R. Milgrom, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics, 1985, 14:1, 71–100.

Hausler, J., J. Ruscheinsky and M. Lang, News-based sentiment analysis in real estate: A machine learning approach, Journal of Property Research, 2018, 35:4, 344–71.

Investment Property Forum, Liquidity in commercial property markets, Working Paper Five: Liquidity – Findings and Recommendations, Investment Property Forum, London, 2004.

Kathmann, R. M., Neural networks for the mass appraisal of real estate, Computers, Environment and Urban Systems, 1993, 17:4, 373–84.

Kyle, A. S., Continuous auctions and insider trading, Econometrica, 1985, 53:6, 1315–35.

Lam, K. C., C. Y. Yu and K. Y. Lam, An artificial neural network and entropy model for residential property price forecasting in Hong Kong, Journal of Property Research, 2008, 25:4, 321–42.

Liu, S., Investor sentiment and stock market liquidity, Journal of Behavioral Finance, 2015, 16:1, 51–67.

Newey, W. K. and K. D. West, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica, 1987, 55:3, 703–8.

Nguyen, N. and A. Cripps, Predicting housing value: A comparison of multiple regression analysis and artificial neural networks, Journal of Real Estate Research, 2001, 22:3, 313–36.

Pastor, L. and R. F. Stambaugh, Liquidity risk and expected stock returns, Journal of Political Economy, 2003, 111:3, 642–85.

Pesaran, M. H. and Y. Shin, An autoregressive distributed-lag modelling approach to cointegration analysis, In S. Strøm, editor, Econometrics and economic theory in the 20th century: The Ragnar Frisch Centennial Symposium, Econometric Society Monographs, Cambridge University Press, Cambridge, 1998, 371–413.

Pesaran, M. H., Y. Shin and R. J. Smith, Bounds testing approaches to the analysis of level relationships, Journal of Applied Econometrics, 2001, 16:3, 289–326.

Peterson, S. and A. Flanagan, Neural network hedonic pricing models in mass real estate appraisal, Journal of Real Estate Research, 2009, 31:2, 147–64.

Porter, M. F., An algorithm for suffix stripping, Program, 1980, 14:3, 130–7.

Poursaeed, O., T. Matera and S. Belongie, Vision-based real estate price estimation, Machine Vision and Applications, 2018, 29:4, 667–76.

Rochdi, K. and M. Dietzel, Outperforming the benchmark: Online information demand and REIT market performance, Journal of Property Investment & Finance, 2015, 33:2, 169–95.

Rosenblatt, F., The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review, 1958, 65:6, 386–408.

Ruscheinsky, J., M. Lang and W. Schäfers, Real estate media sentiment through textual analysis, Journal of Property Investment & Finance, 2018, 36:5, 410–28.

Scofield, D., Time to completion liquidity in UK commercial real estate investment: 2000-2008, Journal of European Real Estate Research, 2013, 6:1, 34–47.

Soo, C. K., Quantifying animal spirits: News media and sentiment in the housing market, Ross School of Business Working Paper No. 1200, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, 2015.

Tieleman, T. and G. E. Hinton, Lecture 6.5-Rmsprop, Coursera: Neural networks for machine learning, Technical Report, Coursera, Inc., 2012.

Uysal, A. K. and S. Gunal, The impact of preprocessing on text classification, Information Processing & Management, 2014, 50:1, 104–12.

Walker, C. B., Housing booms and media coverage, Applied Economics, 2014, 46:32, 3954–67.

Werbos, P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences, Thesis (Ph. D.), Cambridge, 1974.

Worzala, E., M. Lenk and A. Silva, An exploration of neural networks and its application to real estate valuation, Journal of Real Estate Research, 1995, 10:2, 185–201.

Zhang, H., S. Gao, M. J. Seiler and Y. Zhang, Identification of real estate cycles in China based on artificial neural networks, Journal of Real Estate Literature, 2015, 23:1, 65–84.

4 REIT Property Development and Stock Price Crash Risk

4.1 Abstract

This paper analyses the stock price crash risk of US equity REITs, dependent on their property development activity. Following Jin and Myers' (2006) theory on the formation of stock price crashes, arguments can be made both for a positive as well as a negative relationship. The paper reveals that development activity leads to reduced future crash risk for equity REITs. Developments can be interpreted as a manifestation of management farsightedness, as short-term interests are sacrificed for long-term profits. Property-developing REITs might thus show a reduced tendency for bad news hoarding and other myopic behaviors, which are the precondition for crashes.

Keywords: Real Estate Development, REITs, Stock Price Crash Risk, Managerial Short-termism

4.2 Introduction

In an era of all-time low yields and a decreasing amount of moderately priced property investment opportunities, the in-house execution of development projects becomes increasingly attractive for equity REITs. There is some debate in both practice and academia on the question of whether REITs should deviate from their core competency of owning and managing income-producing properties, and actively seek opportunities for excess profit-generation through development projects (for a discussion of the pros and cons see Brounen and Eichholtz, 2004). Studies on the performance of equity REITs engaging in real estate development, however, are scarce and focus mainly on the analysis of differences in return and volatility between developers and non-developers (Gerbich, Levis and Venmore-Rowland, 1999; Brounen, Kanters and Eichholtz, 2000; Brounen and Eichholtz, 2004).

While attracting ever-increasing research interest in the field of the general stock market since the occurrence of several high-profile accounting scandals starting with the fall of *Enron* in 2001, the topic of price crashes has largely been ignored in the REIT literature (see Habib, Hasan and Jiang, 2018, for a review of literature on the general stock market). A positive exception is the study of An, Wu and Wu (2016), who examine the stock price crash risk of US equity REITs, dependent on their institutional investor base. The authors find that a large share of transient investors increases the susceptibility of a REIT to stock price crashes. However, the dependency of crash risk on the business model of equity REITs has so far not been subject to analysis. In light of the prevailing low-interest environment and the resulting struggle for equity REITs to find sound investment opportunities, a closer view on the implications of REIT property-development seems especially useful. The goal of this paper is thus to bring together two under-researched subjects by analyzing the effect of development activity on the stock price crash risk of equity REITs. This combination is particularly relevant, because there is an argument to be made both for a positive as well as a negative influence.

The theoretical model used in this paper to explain stock price crashes was suggested by Jin and Myers (2006), and has generally been the standard explanatory model in the empirical crash risk literature. The core of the model is information asymmetry between firm management and shareholders. Management usually receives news on the firm’s prospects before other shareholder groups. While there should be no meaningful explanation for the hoarding of positive news, there is an array of potential personal reasons for the management to hold back negative news. Career concerns (Kothari, Shu and Wysocki, 2009), or maximization of short-term equity-based compensation (Kim, Li and Zhang, 2011a) are just two of them. Following Jin and Myers' (2006) model, bad news can only be hoarded for a certain period of time, and consequently, if it is not offset by subsequent good news, the accumulated bad news tends to enter the market all at once. This sudden revelation of a bulk of negative information on the firm then results in an adverse stock price reaction or even a price crash.

The tendency to hoard bad news is of course dependent on the degree of short-term orientation of the firm's management. Development projects create sharply increased potential for the occurrence of potentially crash-triggering bad news compared to the conventional management of income-producing properties. However, the (voluntary) realization of such projects can itself be considered a sign of a long-term management orientation. This argument becomes evident when considering the definition of managerial myopia suggested by Stein (1989) of "inflating current earnings at the expense of longer term benefits". Real estate development activity is inevitably associated with the sacrifice of short-term in favor of expected long-term profits. It is thus possible that property-developing REITs exhibit reduced stock price crash risk due to management's lack of disposition to hoard bad news.

The analysis of what effect dominates the relationship between development activity and crash risk has some practical implications for REIT investors. Given the technical capability for in-house execution of complex projects, it would be desirable for REITs to perform property development if their management displays a tendency for long-termism and transparency regarding the publication of relevant news. If management teams of property-developing REITs on the other hand exhibit myopic behavior and are inclined to hoard bad news from their investors, REIT development may have adverse effects on shareholder value. This paper therefore contributes to answering the question posed by Brounen and Eichholtz (2004) as to whether regulators should restrict equity REITs from participating in property development.

The study finds that development activity is significantly negatively related to future stock price crash risk. The relationship remains strong even after performing an array of robustness checks, including variation of the used crash risk proxy, the calculation method of market-adjusted returns used in the crash measures, and winsorization of regression inputs. The findings suggest that REIT willingness to engage in property development can be interpreted as a proxy for long-termism of the REIT executives.

As an additional test for this hypothesis, the tendency of an equity REIT to manage up earnings dependent on its development share, is analyzed. The regressions reveal that property-developing REITs exhibit a lower tendency to engage in positive earnings management than their non-developing counterparts. Altogether, the evidence is strong that equity REITs act with a non-myopic, long-term orientation when performing real estate development. From a crash risk perspective, additional regulation of equity REITs thus seems unnecessary.

The paper proceeds as follows: The next section introduces the relevant REIT property development and stock price crash risk literature, and derives the research questions and hypotheses for this study. Section 4.4 provides information on the construction of the sample and the regression inputs. Section 4.5 presents the empirical design and the results for the regression analyses on the relationship of crash risk and development activity. The section furthermore contains a number of robustness checks. Section 4.6 provides additional evidence supporting the tendency of property-developing REITs towards long-termism, by analyzing their propensity for earnings management. Section 4.7 concludes.

4.3 Literature Review and Research Motivation

With only a few publications, the subject of **REIT and property company development activity** is heavily under-researched. The existing studies focus mainly on the first and second moment of the return distribution of property-developing REITs. Gerbich, Levis and Venmore-Rowland (1999) analyze post IPO and post rights issue performance of UK property management and development firms. They find that issuing development firms significantly underperform their non-issuing peers and furthermore perform relatively worse than issuing property management firms. The authors attribute this finding to the higher pricing uncertainty facing development firms compared to management firms. In another early study, Brounen, Kanters and Eichholtz (2000) analyze the impact of property development activity on the performance of US equity REITs. The authors find that development projects are mainly undertaken by larger REITs and to a greater extent by REITs specializing in retail properties. Brounen, Kanters and Eichholtz (2000) furthermore observe a positive relationship between development activity and unadjusted return. However, the higher returns vanish after risk adjustment. The authors then decompose risk into a systematic and an unsystematic component, and find that development activity increases a REIT's systematic risk and beta, and, surprisingly, lowers its unsystematic risk. Brounen, Kanters and Eichholtz attribute this finding to the size difference between developing and non-developing REITs, and the greater ability of large REITs to diversify risk. In a subsequent cross-sectional analysis they find that smaller developing REITs indeed exhibit a slight increase in unsystematic risk, as predicted. In an international study, Brounen and Eichholtz (2004) compare real estate companies engaging in property development from five different countries in terms of their risk and return characteristics. Their findings suggest that both risk and return is increased for property-developing firms. The authors then examine operational performance of developing firms and find that they again demonstrate both higher profitability as well as more volatility. In a recent study, Geltner, Kumar and van de Minne (2019) examine the net present value of development projects undertaken by US REITs and find a large positive added value for such projects during their study period between 1998 and 2018.

The research field of **stock price crashes** has received a fair amount of attention regarding the general stock market in recent years. Several theories capable of explaining drops in individual stock values have been suggested (see for example Cao, Coval and Hirshleifer, 2002; or Chen, Hong and Stein, 2001 and Hong and Stein, 2003).[36] In the empirical literature, particularly the theory of Jin and Myers (2006) has been used to explain crashes. Jin and Myers base their model on the existence of asymmetric information between firm executives and outsiders, resulting from opaqueness in the firms' operations and reporting. The informational advantage of executives enables them to hide bad news for an extended time period, hoping that it will ultimately be offset by future good news. The motivation behind this behavior of hoarding bad news can be manifold. Kothari, Shu and Wysocki (2009) identify career concerns of managers, Ball (2009) the preservation of peer esteem, as well as that of coworkers or the general public. But also attempts to maximize short-term equity-based compensation (Kim, Li and Zhang, 2011a) can create incentives to hold back

[36] For a summary of alternative stock price crash risk theories and a comprehensive literature review, see Habib, Hasan and Jiang (2018). The summary of relevant literature in this paper focuses on the most influential publications and does not intend to be exhaustive.

adverse information. The hoarding of bad news however, is limited by responsible executives' lack of willingness or ability to continue such behavior. When the accumulated bad news reaches an unsustainable level, the incentives to hold it back break down and the news is released all at once into the market. The sudden arrival of a large body of bad news can then result in a crash of the respective company's stock price. Under the assumption that firm executives have little or no incentives to likewise hold back positive information, stock returns should be left-skewed, and price plunges more frequent than price jumps.

A wide range of publications has been devoted to the empirical examination of Jin and Myers' theory, most of which covers not REITs but the general stock market. Accruals- and real earnings management (Hutton, Marcus and Tehranian, 2009; Francis, Hasan and Li, 2016), sensitivity of CFOs' option portfolio value to stock price (Kim, Li and Zhang, 2011a), tax avoidance (Kim, Li and Zhang, 2011b), short interest (Callen and Fang, 2015), and CEO overconfidence (Kim, Wang and Zhang, 2016) have been shown to be positively related to future crash risk. According to the literature, corporate social responsibility disclosures (Kim, Li and Li, 2014), conservative accounting (Kim and Zhang, 2016) and auditor tenure (Callen and Fang, 2017) are negatively associated with future crash risk.

In recent years, the focus has shifted towards the role of a firm's investor base with regard to its crash risk. In particular the monitoring engagement and trading behavior of institutional investors have received attention. Findings show that holdings by dedicated institutional investors (An and Zhang, 2013) and a stable institutional investor base (Callen and Fang, 2013) reduce future crash risk. In the only study so far on crash risk in the REIT sector, An, Wu and Wu (2016) analyze the role of institutional investors concerning the crash susceptibility of US equity REITs. The authors find that pension fund (bank trust) holdings decrease (increase) crash risk, and that in more recent years, the trading activity of investment companies has had a positive effect on crash risk. An, Wu and Wu then address the impact of the behavior of institutional investors on crash risk and show that transient investors attempting to maximize short-term gains elevate crash risk. In a related study on the general stock market, Chang, Chen and Zolotoy (2017) broaden the picture and examine the role of stock liquidity in crash risk. They find that higher liquidity leads to an increase in the tendency of managers to withhold bad news, resulting from a fear of transient investors selling after the release of bad news into the market.

Apart from their empirical analysis, Chang, Chen and Zolotoy (2017) also provide a decomposition of the possible drivers of stock price crash risk. The authors differentiate between (1) the likelihood of bad news formation, (2) the extent of managerial bad news hoarding and (3) the stock market's reaction to a release of accumulated bad news. This distinction can be helpful in order to hypothesize on the relationship between the property-development activity of REITs and their crash risk.

In terms of (1) **bad news formation**, development activity should exert a reinforcing effect. As stochastic processes whose features are inconstant over time (Byrne, 1996), property developments impose a considerable amount of additional unpredictability on REIT cash flows. Compared to an equity REIT's fairly stable default business activity of owning and operating income-producing real estate, a property-developing REIT's management has to handle unique, non-standardizable ventures. Using Jin and Myers' (2006) terminology, development projects should moreover be accompanied by an increase in the respective REIT's opaqueness, at least to some extent. Alongside conventional risks (construction cost and time overruns, financing gaps etc.) Loizou and French (2012) also emphasize the risks resulting from highly complex human relationships that are necessary to successfully execute development projects. The likelihood of bad news occurrence with the subsequent potential for managerial bad news hoarding within property-developing REITs should thus exceed that of their non-developing counterparts.

(2) **Managerial bad news hoarding tendency** on the other hand might be negatively related to development activity. The reasoning behind this hypothesis is that voluntary engagement in property development projects can be interpreted as a proxy for REIT management long-termism. According to this perspective, the development expenditure of a REIT is analogous to the R&D expenditure of firms outside the REIT sector. Initially triggered by Porter's (1992) criticism of the tendency of US firms to neglect R&D as a response to myopic pressure from highly liquid capital markets, R&D reduction has developed a long history as a proxy for managerial short-termism (see e.g. Bushee, 1998; or Samuel, 2000). The argument can be made that entering development projects, which generally take years to complete and provide little or no return in the short run, requires a long-term perspective on the REIT's business activity (comparable to R&D projects of firms in other industries). Considering that the alternative, i.e. the acquisition of completed properties, does not necessarily require that long-term perspective, development projects might exercise a disciplinary effect

on firm management by shifting focus from instant- to delayed gratification. It then seems contradictory to assume myopic hoarding behavior of bad news towards investors from management teams demonstrating such long-term orientation in their time preferences.

Opposing that view, there is the possibility that a REIT's management overestimates its own capabilities when engaging in development projects. Developments lie beyond the customary competences of equity REITs and require a set of construction-related and coordination skills unequalled by the tasks required for operating stable properties. If a property-developing REIT's management regularly acts overconfidently, a positive association between development activity and news hoarding behavior could be possible, given the empirical evidence on managerial overconfidence provided by Kim, Wang and Zhang (2016) and the tendency of developers to underestimate project risk (MacFarlane, 1995).

Unlike in Chang, Chen and Zolotoy's (2017) analysis of capital market liquidity, development activity does not itself suggest assumptions on the (3) **market response to the release of hoarded bad news**. However, to incorporate the effect of tradability of the respective REIT's stock after such releases, this paper controls for stock liquidity in the empirical models.

Considering the existence of theoretical arguments for either a positive or a negative relationship between development activity and stock price crash risk, the question of the direction of that relationship has to be answered empirically.

4.4 Data Sources and Variable Construction

4.4.1 Sample Selection

The paper makes use of the *S&P Global Market Intelligence* (former *SNL Financial*) database to retrieve property type and development data for listed US equity REITs. Accounting data for the REITs is obtained from *Compustat*, and stock market data from the *Center for Research in Security Prices* (CRSP). All US real estate companies in the S&P database form the starting point for the sample generation. To eliminate non-REITs, the retrieved companies are compared to the yearly constituent lists of the *SNL US Equity REIT Index*. Companies that are not part of the index are removed from the sample for the respective years. Stock years with a duration other than 12 months, a year-end closing price below one dollar, less than 100 days with at least one trade occurring, negative book value of equity or incomplete data from one of the three databases, are excluded from the sample.[37] The *SNL US Equity REIT Index* is only available from 2001, which thus constitutes the first year included in the panel. The final sample consists of 1,704 REIT stock-year observations from the fiscal years 2001 to 2017.

4.4.2 Development Activity

To measure a REIT's real estate development exposure, this paper follows Geltner, Kumar and van de Minne (2019) in selecting relevant components. DL, defined as the book value of non-depreciable land held for the purpose of future development sites is used to account for planned future development potential. CIP is defined as the book value of non-depreciable real estate for which construction is in progress and accounts for ongoing development activity. To obtain a complete picture of REIT development, this paper's development measure takes the sum of DL and CIP for a respective REIT i and fiscal year t and divides it by its book value of total assets (TA):[38]

$$DEV_{i,t} = \frac{DL_{i,t} + CIP_{i,t}}{TA_{i,t}}. \tag{4.1}$$

Geltner, Kumar and van de Minne (2019) furthermore make use of a variable for measuring the book value of the unfunded development pipeline of a REIT, in order to measure its exposure to projects which are not yet expanded, for one of their development measures. S&P's sample for the pipeline variable is greatly reduced, and this paper's author also argues that a focus on actual balance sheet development exposure might result in more reliable measures than the incorporation of future exposures. In this paper, unfunded pipeline is thus not added to the development measure.

[37] Following several other publications (see e.g. Kim, Li and Zhang, 2011a; or Chang, Chen and Zolotoy, 2017), this paper excludes "penny stocks", i.e. those valued below 1 USD at fiscal-year end. Running the empirical analysis without removing such stocks however, leads to similar results.
[38] The use of book value figures for measuring development activity is not ideal, as a market-oriented view could provide a more accurate picture of a REIT's development exposure. However, due to the unavailability of market value figures and the broad use of book measures in the literature (see e.g. Brounen, Kanters and Eichholtz, 2000; Brounen and Eichholtz, 2004; or Geltner, Kumar and van de Minne, 2019), the author of this paper argues that the compromising use of fiscal year end book-measures as market value proxies is reasonable.

4.4.3 Crash Risk Measures

For the construction of the stock price crash risk measures employed in the study, weekly total returns for the REITs in the sample are calculated. The use of weekly instead of daily returns minimizes problems with illiquid REITs with a low trading frequency. In order to obtain measures of idiosyncratic risk instead of merely capturing swings of the general stock- or REIT market, an adjustment of returns used for the construction of crash risk measures is necessary. For this reason, the weekly REIT returns are regressed against weekly returns of the CRSP market index, as well as the *SNL US Equity REIT Total Return Index*. This paper follows the methodology of An, Wu and Wu (2016) and uses the expanded market model denoted in Equation (4.2) to filter out the idiosyncratic component of the weekly returns for REIT i in fiscal year t:

$$R_{i,w} = \beta_0 + \beta_1 R_{c,w} + \beta_2 R_{s,w} + \varepsilon_{i,w} \,, \tag{4.2}$$

with $R_{i,w}$ representing the return of REIT i in week w, $R_{c,w}$ the CRSP value-weighted market return in week w, and $R_{s,w}$ the return on the value-weighted *SNL US Equity REIT Total Return Index* in week w. The residuals $\varepsilon_{i,w}$ of the market model in Equation (4.2) are log transformed and used to construct the crash measures for this study, i.e. the firm-specific weekly return for REIT i in week w is calculated as:

$$W_{i,w} = \ln\left(1 + \varepsilon_{i,w}\right) . \tag{4.3}$$

For the main part of the empirical analysis, the measures for stock price crash risk initially suggested in an early study by Chen, Hong and Stein (2001) are employed. The first measure represents the negative conditional skewness (NCSKEW) of returns, defined as the third moment of the firm-specific weekly returns, divided by the standard deviation of the firm-specific weekly returns raised to the third power. The resulting skewness measure is multiplied by minus 1 to yield a positive relationship to crash risk:

$$NCSKEW_{i,t} = -\frac{n(n-1)^{\frac{3}{2}} \sum W_{i,w}^{3}}{(n-1)(n-2)\left(\sum W_{i,w}^{2}\right)^{\frac{3}{2}}} \,, \tag{4.4}$$

where $W_{i,w}$ denotes the firm-specific weekly return of REIT i in week w, and n the number of weekly returns in the respective fiscal year.

The second measure used by Chen, Hong and Stein (2001) is DUVOL, which compares the standard deviation of returns below and above the fiscal-year mean return of a stock. Specifically, DUVOL is calculated as:

$$DUVOL_{i,t} = ln\left(\left((n_u - 1)\sum_{DOWN} W_{i,w}^2\right) \Big/ \left((n_d - 1)\sum_{UP} W_{i,w}^2\right)\right), \quad (4.5)$$

where n_u and n_d denote the number of firm-specific weekly returns ($W_{i,w}$) above and below the fiscal-year mean, respectively. Following Chen, Hong and Stein, the natural logarithm of the quotient is taken to constitute DUVOL. The measure should be less susceptible to the influence of return outliers than NCSKEW, as third moments are not involved in its calculation.

4.4.4 Control Variables

The regression analysis employs a number of control variables whose influence on crash risk has been proposed in the literature. Following the rationale behind the early study of Chen, Hong and Stein (2001), this paper controls for detrended turnover (DTURN). DTURN represents the difference between the fiscal-year mean monthly stock turnover of a REIT in fiscal year t and fiscal year t-1. Monthly stock turnover is calculated as the monthly number of shares traded, divided by the average number of shares outstanding for the REIT in the respective month. The measure proxies for the difference of opinions among investors, which, following Chen, Hong and Stein's reasoning, increases crash risk. In accordance with the existing literature, controls for past stock return (RET), return volatility (SD), firm size (SIZ), market-to-book ratio (MTB), leverage (LEV), as well as contemporaneous profitability (ROA), are added to the regression analyses (see for example Hutton, Marcus and Tehranian, 2009; Kim and Zhang, 2016; or An, Wu and Wu, 2016). RET is defined as the fiscal-year mean of a REIT's weekly excess returns over the weekly returns of the *SNL US Equity REIT Index*, multiplied by 100. SD is the standard deviation of a REIT's weekly excess return over the weekly returns of the SNL index in a given fiscal year. The log of each REIT's fiscal-year-end market-capitalization (in million USD) proxies for firm size (SIZ), the ratio of the market value of equity divided by the book value of equity constitutes the market-to-book ratio (MTB). Hutton, Marcus and Tehranian (2009) document a negative relationship between crash risk and leverage, as well as contemporaneous operating profitability. Thus LEV, defined as long-term debt divided by total assets, and ROA, defined as contemporaneous net income divided by total assets, are added into the equations. In order to account for the persistence of return skewness over time (Chen, Hong and Stein, 2001), this study includes a lagged version of NCSKEW in the regressions. To furthermore incorporate the more recent findings of An, Wu and Wu (2016) and Chang, Chen and Zolotoy (2017), the influence of transient institutional investors on the respective REIT's crash risk has to be controlled for. Thus, to account for the attractiveness of the REIT to such investors, the yearly average of daily bid-ask-spreads (SPR) is added to the regression equations as a proxy for liquidity. The daily spreads are calculated as the difference between the quoted closing bid and ask price of the respective REIT, divided by its quote midpoint. Chung and Zhang (2014) show that this low-frequency estimation is highly correlated with spreads calculated on the basis of intraday data (as used, for example, in Chang, Chen and Zolotoy, 2017), and thus serves well as a proxy.

4.4.5 Summary Statistics and Correlations

Table 4.1 provides summary statistics (Panel A) and correlations (Panel B) on the crash risk-, development-, and control variables. The share of development activity for the sample is on average 3.4%. While, in a considerable share of REIT stock years, no property development projects are undertaken, a DEV maximum value of 50.1% suggests that the disparity in terms of development activity in the sample is quite large. Note also that the market-to-book ratio (MTB) exhibits a high degree of volatility and, with a ratio of about 641, an exceptionally high maximum. This peculiarity is caused by a few REITs with very low, almost negative book value of equity, inflating MTB values severely. In the robustness test section, the effect of these potentially biasing data points is removed by winsorization. For the main analysis, the MTB values are left unaltered. The correlation matrix suggests possible multicollinearity issues for the variable pair of bid-ask-spread (SPR) and market capitalization (SIZ). The correlation of -0.605 is however, comparable to the correlation reported by Chang, Chen and Zolotoy (2017).

Table 4.1: Summary Statistics and Correlation Matrix

Panel A: Summary Statistics

Statistic	Mean	St. Dev.	Min	25th	75th	Max
NCSKEW	0.066	0.797	-5.378	-0.354	0.413	5.033
DUVOL	0.024	0.356	-1.927	-0.201	0.235	1.643
DEV	0.034	0.051	0.000	0.000	0.048	0.501
SPR	0.006	0.012	0.0001	0.001	0.005	0.163
ROA	0.023	0.044	-0.295	0.008	0.038	0.904
MTB	2.922	16.414	0.106	1.224	2.392	640.965
LEV	0.468	0.142	0.000	0.395	0.555	0.887
SIZ	7.037	1.509	1.029	6.201	8.021	10.713
DTURN	0.006	0.079	-0.754	-0.015	0.026	0.527
RET	0.039	0.438	-2.199	-0.185	0.230	3.976
SD	0.028	0.021	0.008	0.017	0.031	0.245

Panel B: Correlation Matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	$NCSKEW_t$	1											
(2)	$DUVOL_t$	0.951	1										
(3)	DEV_{t-1}	-0.056	-0.067	1									
(4)	SPR_{t-1}	-0.086	-0.087	-0.121	1								
(5)	ROA_t	-0.031	-0.016	-0.027	-0.151	1							
(6)	MTB_{t-1}	-0.017	-0.015	0.039	-0.019	0.031	1						
(7)	LEV_{t-1}	-0.003	-0.012	0.085	-0.006	-0.239	0.130	1					
(8)	SIZ_{t-1}	0.067	0.079	0.205	-0.605	0.171	0.016	-0.017	1				
(9)	$DTURN_{t-1}$	-0.001	-0.010	0.061	0.010	-0.019	-0.007	0.042	-0.059	1			
(10)	RET_{t-1}	0.014	0.026	-0.017	-0.004	0.095	0.018	-0.025	0.045	0.037	1		
(11)	SD_{t-1}	-0.059	-0.072	-0.053	0.446	-0.241	-0.027	0.021	-0.495	0.244	0.165	1	
(12)	$NCSKEW_{t-1}$	0.080	0.066	-0.048	-0.089	-0.072	-0.018	0.015	0.034	0.036	-0.294	-0.016	1

Notes: Panel A of Table 4.1 reports summary statistics for the variables used in the regression analyses, Panel B reports correlations. Correlations are calculated using the lag structure applicable for the regressions, i.e. all variables except the dependent variables NCSKEW and DUVOL, as well as return on assets (ROA) enter the calculation lagged once. The sample contains all listed US equity REITs covered in the *SNL US Equity REIT Index* for the fiscal years 2001 to 2017, with accounting data from *Compustat*, stock data from CRSP, and real estate development data from *S&P Global Market Intelligence*.

4.5 REIT Property Development and Crash Risk: Empirical Analysis

4.5.1 Research Design and Results

In terms of empirical design, this paper follows the literature (see e.g. An, Wu and Wu, 2016; Chang, Chen and Zolotoy, 2017) and uses a panel structure with leading dependent variables:

$$\begin{aligned} \begin{matrix} NCSKEW_{i,t} \\ DUVOL_{i,t} \end{matrix} &= \beta_0 + \beta_1 DEV_{i,t-1} + \beta_2 SPR_{i,t-1} + \beta_3 ROA_{i,t} \\ &+ \beta_4 MTB_{i,t-1} + \beta_5 LEV_{i,t-1} + \beta_6 SIZ_{i,t-1} \\ &+ \beta_7 DTURN_{i,t-1} + \beta_8 RET_{i,t-1} + \beta_9 SD_{i,t-1} \\ &+ \beta_{10} NCSKEW_{i,t-1} + YEAR_t + PROPTYPE_i + \varepsilon_{i,t} \,. \end{aligned} \tag{4.6}$$

Regressions incorporate fiscal-year ($YEAR_t$) and property-type ($PROPTYPE_i$) fixed-effects and either NCSKEW or DUVOL as the dependent variable.[39] DEV constitutes the independent variable measuring a REIT's development activity in a respective fiscal year. All control variables are included in both models. The regressions are estimated via OLS, heteroscedasticity-robust standard errors are adjusted for clustering at the REIT level.

The results for the control variables in Table 4.2 are in line with the findings in the literature. A strong, significantly negative relationship between liquidity (SPR) and crash risk can be observed. Considering that the construction of SPR implies an inverse relationship to liquidity (i.e. higher spreads imply lower liquidity), this result confirms Chang, Chen and Zolotoy's (2017) finding of a crash-inducing effect of high liquidity for the REIT market, and indirectly An, Wu and Wu's (2016) results, by providing evidence for the influence of transient investors on crash risk. Such investors are attracted by high trading liquidity and cause crashes through their cut-and-run investment behavior after the release of negative news. The finding persists when SIZ is excluded from the regression, so as to eliminate potential multicollinearity problems.

[39] An, Wu and Wu (2016) do not include property-type fixed-effects in their models. However, with development activity as the independent variable in this paper, controlling for each REIT's primary property-type focus is necessary in order to avoid endogeneity issues resulting from the diverging development exposure prevailing in different property-type subsectors (consider in this context for example the findings of Brounen, Kanters and Eichholtz, 2000).

ROA likewise exhibits a significantly negative relationship to crash risk, meaning that firms with strong operating performance have a reduced propensity for stock price crashes. The coefficients of MTB and LEV are both close to zero and insignificant, which corresponds to the results of An, Wu and Wu (2016). Concerning SIZ, An, Wu and Wu find a significantly positive relationship to crash risk for most of their models, indicating that larger REITs are more crash-prone. Although the coefficients of SIZ are positive, they are insignificant for the purposes of this paper. This difference might be driven by the early- and mid-1990s years, which are included in An, Wu and Wu's study and were characterized by significantly smaller average market capitalizations in the US REIT sector. DTURN is negatively associated with crash risk, but insignificant. In most of their models, An, Wu and Wu find detrended turnover to have a significantly negative impact on crash risk, opposing the findings of Chen, Hong and Stein (2001). When truncating the study period used in this paper to match the end of An, Wu and Wu's sample, DTURN turns significantly negative for both NCSKEW and DUVOL, replicating An, Wu and Wu's result. The relationship between turnover and crash risk thus might have weakened over the more recent years. Like An, Wu and Wu, this paper finds RET to be positively, and SD negatively associated with crash risk. Both controls however, are insignificant in An, Wu and Wu's regressions, while for the DUVOL specification, RET is significant in this paper. Finally, this study confirms the finding in the literature that skewness in firm-specific returns is persistent over time, i.e. lagged NCSKEW is linked significantly positively to both NCSKEW and DUVOL.

With regard to the variable of interest *development activity*, the regressions exhibit insightful results. DEV displays a strong negative relationship to stock price crash risk both for the NCSKEW and the DUVOL model, with each of the coefficients being significant at the 1% level. For the NCSKEW specification, a one standard deviation increase in the development share of a REIT leads to a 4.8% decrease in future crash risk for the sample ($0.051 * -0.950 \approx -0.048$). To put this value into perspective, note that a change in contemporaneous profitability (ROA) of one standard deviation leads to a decrease in crash risk of 4.9% for the sample ($0.044 * -1.124 \approx -0.049$). Given that profitability has been shown to constitute a strong contrarian predictor for crash risk (see e.g. Hutton, Marcus and Tehranian, 2009; Kim, Li and Zhang, 2011a; or Chang, Chen and Zolotoy, 2017), the comparable magnitude of development activity seems to be not only significant, but also economically meaningful.

Table 4.2: REIT Stock Price Crash Risk - Development Activity (Main Model)

	Dependent variable:	
	$NCSKEW_t$	$DUVOL_t$
DEV_{t-1}	-0.950***	-0.495***
	(0.353)	(0.163)
SPR_{t-1}	-9.101**	-3.683**
	(3.947)	(1.504)
ROA_t	-1.124**	-0.413**
	(0.449)	(0.193)
MTB_{t-1}	-0.0002	-0.00001
	(0.001)	(0.0003)
LEV_{t-1}	-0.010	-0.020
	(0.141)	(0.062)
RET_{t-1}	0.072	0.036*
	(0.047)	(0.019)
SD_{t-1}	-1.550	-0.742
	(1.618)	(0.672)
SIZ_{t-1}	0.003	0.006
	(0.023)	(0.009)
$DTURN_{t-1}$	-0.057	-0.054
	(0.289)	(0.134)
$NCSKEW_{t-1}$	0.064*	0.023**
	(0.033)	(0.012)
Property-type fixed-effects	YES	YES
Fiscal-year fixed-effects	YES	YES
Observations	1,704	1,704
R^2	0.039	0.039
Significance levels:		**p<0.1; **p<0.05; ***p<0.01*

Notes: Table 4.2 reports the results for the main regressions of stock price crash risk on property-development activity for the sample of listed US equity REITs from the fiscal year 2001 to 2017. NCSKEW is the conditional skewness of firm-specific weekly returns multiplied by minus 1 for each stock year. DUVOL is the natural logarithm of the volatility of the below-fiscal-year mean firm-specific weekly returns divided by the volatility of the above-fiscal-year mean firm-specific weekly returns for each stock year. DEV denotes the sum of development land and construction in progress, divided by total assets for each stock year. Regressions contain property-type- and fiscal-year fixed-effects. Heteroscedasticity-robust standard errors (in parenthesis) are adjusted for clustering at the firm level.

The findings can be interpreted as evidence supporting the hypothesis that property-developing REITs tend to act in a long-term manner, and that their managements are less susceptible to myopic bad news hoarding. The long-termism effect seems to outweigh the potentially higher likelihood of bad news occurrence and opaqueness for developing REITs. In this context it is important to bear in mind that the arrival of bad news itself does not necessarily increase the probability of stock price crashes, given the willingness of management to provide transparency and full news disclosure to shareholders. Overconfidence of developing REITs' managers furthermore does not seem to be a channel that induces a critical effect on the relationship between development activity and crash risk.

An identification of the deeper psychological causes of the supposed long-termism of property-developing REITs' executives is speculative at this point. A possible access point to approach the observed relationship from a psychological perspective could be the concept of *density*. Density originally describes an individual's assessment of its future in terms of the anticipated number of personal prospects (Kastenbaum, 1961). In an organizational context, the concept denotes the feeling of the firm management that the future either will hold an array of potentially profitable possibilities (high density), or that it will lead to struggles for the firm to even hold its current position (low density), following Laverty (2004). Managerial density has been shown to be negatively related to an under-valuation of the long term (Laverty, 2004). One could argue that voluntary performance of property developments necessarily requires a high density assessment of the firm's prospects. Their positive appraisal of the firm's density might direct the focus of developing REITs' managements on long-term goals and away from the chase of myopic success.

As a personal trait of the respective firm executives, long-termism itself is unobservable. A possible remedy to elaborate on its effects could be an analysis of the relationship between development activity and alternative, more obvious long-termism proxies. More precisely, if officials of property-developing REITs indeed behave less myopic than those of non-developers, they should also display a lower tendency for other manifestations of short-termism. Arguably the most direct indication of managerial short-termism is the manipulation of firm earnings. Stein's (1989) definition of myopia as the inflation of current earnings implies that, given their long-term orientation holds true, property-developing REITs should show little or no inclination to perform earnings manipulation. The section "REIT Development

Activity: Relationship to Managerial Myopia" follows up on that thought and presents further analysis on the long-termism hypothesis for developing REITs. First, however, several tests for the robustness of the empirical results of this section are provided.

4.5.2 Robustness Tests

To ensure the robustness of the results from the previous section, several checks are performed. First, the dependence of the relationship between crash risk and development activity on the choice of crash risk measure is analyzed. Instead of using the rather indirect representation of crashes through NCSKEW and DUVOL, regressions are rerun with a more intuitive crash variable, first used by Jin and Myers (2006). For the COUNT variable, a crash is defined as an extremely negative return event, surpassing a certain standard deviation threshold. Specifically, this paper follows the literature (e.g. the REIT study of An, Wu and Wu, 2016) and sets the threshold for the classification of a firm-specific weekly return as a crash at 3.09 standard deviations below the stock-year mean weekly return. This results in a crash probability of 0.1% under the assumption of normally distributed returns. To prevent COUNT from merely capturing the volatility of firm-specific returns, Jin and Myers (2006) suggest that the number of upward return jumps in the respective stock year (i.e. weekly firm-specific returns 3.09 standard deviations or more above the yearly mean) must be subtracted from the number of crashes in that stock year. As the threshold of 3.09 standard deviations for the basic measure (COUNT_M) is chosen arbitrarily, alternative thresholds of 2.5 (COUNT_L) and 3.5 (COUNT_H) standard deviations are considered additionally. Results on the regressions using COUNT as the measure for crash risk can be found in Table 4.3.

Table 4.3: Robustness Tests, REIT Stock Price Crash Risk - Development Activity

	Dependent variable:		
	$COUNT\_L_t$	$COUNT\_M_t$	$COUNT\_H_t$
DEV_{t-1}	-1.144***	-0.697*	-0.391*
	(0.426)	(0.371)	(0.206)
SPR_{t-1}	-4.238	-5.862*	-4.033**
	(3.491)	(3.358)	(1.642)
ROA_t	-0.528	-0.568	-0.327
	(0.469)	(0.380)	(0.271)

(Table continues on next page.)

Table 4.3: Robustness Tests, REIT Stock Price Crash Risk - Development Activity (cont.)

MTB_{t-1}	-0.002**	-0.001	0.0001
	(0.001)	(0.001)	(0.0002)
LEV_{t-1}	0.083	0.020	-0.003
	(0.150)	(0.121)	(0.076)
RET_{t-1}	0.072	0.015	0.020
	(0.054)	(0.040)	(0.032)
SD_{t-1}	-1.938	-0.484	-0.409
	(1.841)	(1.415)	(0.834)
SIZ_{t-1}	0.009	0.005	-0.006
	(0.022)	(0.017)	(0.010)
$DTURN_{t-1}$	-0.194	-0.021	0.031
	(0.421)	(0.280)	(0.189)
$NCSKEW_{t-1}$	0.055*	0.047**	0.031*
	(0.028)	(0.022)	(0.019)
Property-type fixed-effects	YES	YES	YES
Fiscal-year fixed-effects	YES	YES	YES
Observations	1,704	1,704	1,704
R^2	0.023	0.031	0.023

Significance levels: **p<0.1; **p<0.05; ***p<0.01*

Notes: Table 4.3 reports the results for the robustness test regressions of stock price crash risk on development activity for the sample of listed US equity REITs from the fiscal year 2001 to 2017. The COUNT figures denote an alternative crash risk measure, taking the difference of extremely negative firm-specific weekly returns (crashes) and extremely positive firm-specific weekly returns (jumps) for a stock year. Three different standard deviation thresholds for the calculation of crashes and jumps constituting COUNT are used in the regressions. 3.09 (COUNT_M) is the standard threshold used in the crash risk literature (see e.g. An, Wu and Wu, 2016), thresholds of 2.5 (COUNT_L) and 3.5 (COUNT_H) are additionally tested to evaluate the robustness, independently of the arbitrary choice of the standard threshold value. DEV again denotes the sum of development land and construction in progress, divided by total assets for each stock year. The regressions contain property-type- and fiscal-year fixed-effects. Heteroscedasticity-robust standard errors (in parenthesis) are adjusted for clustering at the firm level.

While there are some changes in significance, the signs of the control variable coefficients in the COUNT regressions remain vastly similar to the ones in the regressions from the main section. The relevant coefficients on the DEV variables are negative and significant for all three models, indicating a reduced crash propensity for REITs with higher property-development share. The coefficients of the COUNT_M and COUNT_H model are only significant at the 10% level. Considering the relatively low volatility of equity REITs, these stricter measures define only a small number of weekly firm-specific returns as crashes or jumps, which might reduce their explanatory power. Specifically, for all weekly firm-specific returns in the 1,704 stock years of the sample, the threshold of 3.09 (3.5) standard deviations leads to only 357 (208) crashes and 289 (131) jumps, while in the model specification using a threshold of 2.5, the number of crashes is 946 and the number of jumps 809. Nonetheless, the relationship between development activity and stock price crash risk appears to be robust in the choice of the crash measure.

Considering the vast economic distortions caused by the global financial crisis (GFC) of 2007-08, the relationship between development activity, managerial myopia and stock price crash risk might have changed in the aftermath of the crisis. Perhaps, the tendency of property-developing REITs to exhibit a long-term perspective on investing only emerged after the conclusions of the GFC had been drawn. To test for this possibility, the study's sample is split into a pre-crisis and a post-crisis subsample and the relationship between property-development and crash risk re-evaluated for both time periods separately. The first subsample covers the fiscal years 2001 to 2006, the second subsample the fiscal years 2009 to 2017. DEV coefficients both for the NCSKEW as well as the DUVOL model specification are significantly negative for the two subsamples.[40] The relationship thus seems to be robust with respect to the specification of the sample period and not merely attributable to a shift in conduct of REIT officials in response to the GFC.

As a further robustness check, a lead and a lag weekly return are added to the market model regression of Equation (4.2) for both the CRSP as well as the SNL index, in order to account for non-synchronous trading. This adjustment is frequently used in the crash risk literature (see e.g. Jin and Myers, 2006; Hutton, Marcus and Tehranian, 2009; or Chang, Chen and Zolotoy, 2017) and serves as a bias correction of market model betas of infrequently traded shares (Dimson, 1979). With leads and lags included, the coefficients of development activity decrease slightly, but retain their significance.[41]

[40] For the early subsample specification (fiscal years 2001-2006), the coefficient of DEV using NCSKEW (DUVOL) as the dependent variable is -1.021 (-0.574) and significant at the 10% (1%) level. The NCSKEW (DUVOL) coefficient for the late subsample specification (fiscal years 2009-2017) is -0.990 (-0.479) and significant at the 5% (5%) level.

[41] The coefficient of DEV in the model using NCSKEW (DUVOL) as the dependent variable is -0.692 (-0.388) and significant at the 5% (5%) level.

Chang, Chen and Zolotoy (2017) winsorize all non-dummy variable regression inputs both at the 1st and 99th percentiles, to reduce the impact of outliers on the results. A test of this practice seems particularly sensible for this paper, as a few data points of the control variable market-to-book ratio (MTB) are substantial outliers. Regressions are thus reestimated with dependent, independent, and control variables winsorized at the 1st and 99th percentiles. This modification leaves the negative relationship of development activity and crash risk unchanged, or if anything, slightly stronger.[42]

A concern regarding the eligibility of the development activity variable (DEV) used in this paper lies in its denominator. The use of book value of total assets can potentially bias the DEV figures of long-existing REITs with a largely written-off asset base. For such REITs, the denominator of DEV might be underestimated, which would result in an upward biased development share. For this reason, total accumulated depreciation is obtained from the *S&P Global Market Intelligence* database and added back to the book value of total assets in the denominator of DEV. The results including this correction are again basically unchanged, or even marginally stronger than the default setup.[43]

The use of a development measure including development land furthermore entails the possibility that its relationship to crash risk is majorly driven by the land holdings of a REIT. To rule out this possibility, development activity is recalculated without adding DL to the numerator. This variation again leaves the coefficients for development activity significant and even slightly higher negative, compared to the standard specification.[44]

A critical element in the crash risk literature lies in the methodology used to determine firm-specific weekly returns. Running market model regressions with relatively few observations per stock year implies a vast impact of return outliers. Idiosyncratic crash weeks might themselves lead to a shifted estimate of the regression hyperplane and by so doing, severely change the magnitude of the regression residuals used to calculate

[42] The coefficient of DEV in the model using NCSKEW (DUVOL) as the dependent variable is -1.021 (-0.545) and significant at the 1% (1%) level.
[43] S&P defines the "Total Accumulated Depreciation"-variable as "depletion, depreciation, or amortization on all properties, plant and equipment subject to depletion, depreciation, or amortization". Note that two stock-year observations are lost in the depreciation-adjusted regressions, due to missing S&P data, leaving a sample of 1,702 stock years. The coefficient of DEV in the model using NCSKEW (DUVOL) as the dependent variable is -1.045 (-0.544) and significant at the 1% (1%) level for the depreciation-adjusted model.
[44] When DL is dropped from the numerator, the coefficient of DEV in the model using NCSKEW (DUVOL) as the dependent variable is -1.048 (-0.550) and significant at the 1% (1%) level.

firm-specific weekly returns. Thus, as a further test, firm-specific weekly return outliers for each stock year are identified with Cook's (1977) measure for determining influential observations. For this purpose, the market models are first run according to Equation (4.2) without any modifications. In the next step, Cook's distance is calculated for each week w for the observations in a respective stock year as:

$$D_w = \frac{\sum_{j=1}^{n} (\hat{y}_j - \hat{y}_{j(w)})^2}{p * \frac{RSS}{n-p}}, \tag{4.7}$$

where $\hat{y}_{j(w)}$ denotes the fitted response value when week w is excluded from the stock-year regression, and $\hat{y}_j$ denotes the response value for the regression containing all weekly observations. p is the number of coefficients in the model, RSS the residual sum of squares, and n the number of observation weeks. This paper adopts the threshold suggested by Cook and Weisberg (1982) and classifies observations with a Cook's distance above 1 as outliers. The market model regressions are then rerun without the outliers. Using outlier-adjusted residuals from the market models for the calculation significantly changes the crash measures for the affected stock years. Nevertheless, the relationship between development activity and crash risk remains negative and significant for both measures.[45]

4.6 REIT Development Activity: Relationship to Managerial Myopia

The interpretation of development activity as a proxy for managerial long-term orientation should imply that property-developing REITs also exhibit a reduced tendency toward other variants of short-term behavior. A well-analyzed field in the literature is the tendency of management to manipulate earnings in order to show the firm in a better light.[46] In their survey on earnings reporting and disclosure decisions among company officials, Graham, Harvey and Rajgopal (2005) state that "many executives feel that they are choosing the lesser evil by sacrificing long-term value to

[45] To illustrate the magnitude of the differences between the standard sample and the outlier-adjusted sample, quintile correlations for NCSKEW are calculated. To this end, NCSKEW values resulting from the outlier-adjusted regressions are sorted into five quantiles. The correlations between the adjusted and unadjusted NCSKEW values within the quantiles are (in ascending order) 0.936, 0.816, 0.749, 0.807, 0.954. The coefficient of DEV in the model using NCSKEW (DUVOL) as the dependent variable is -0.962 (-0.491) and significant at the 1% (1%) level for the outlier-adjusted model.

[46] Examples of studies using earnings management as a manifestation of managerial short-termism include Stein (1989), Bhojraj and Libby (2005), Edmans (2009), or Chen, Rhee, Veeraraghavan and Zolotoy (2015).

avoid short-term turmoil”, suggesting a widespread willingness to participate in such behavior. Both the creation of annual report opacity through accruals management (Hutton, Marcus and Tehranian, 2009) and the inflation of real earnings (Francis, Hasan and Li, 2016) by firm officials have been shown to lead to an increase in future crash risk. Following the findings of Cohen, Dey and Lys (2008), who state that after the passing of the *Sarbanes-Oxley Act* in 2002, real earnings management increased sharply compared to accruals management, this paper utilizes real earnings management as a proxy for managerial short-termism. More precisely, two earnings management proxies suggested by Ambrose and Bian (2010) in a study on positive and negative REIT earnings management are used to construct a joint measure. The two proxies from Ambrose and Bian are based on the premise that managers can carry out earnings management by manipulating revenue and cost. Normal revenue- and cost levels are estimated via OLS for each property type and fiscal year. The first model, estimating the normal level of revenue is constructed as:

$$\frac{REV_{i,t}}{A_{i,t-1}} = \beta_0 \frac{1}{A_{i,t-1}} + \beta_1 \frac{REV_{i,t-1}}{A_{i,t-1}} + \beta_2 \frac{\Delta REV_{i,t-1}}{A_{i,t-1}} + \varepsilon_{i,t} \,, \qquad (4.8)$$

with $REV_{i,t}$ denoting the revenue of REIT i in fiscal year t, $REV_{i,t-1}$ the previous fiscal year’s revenue of REIT i, and $\Delta REV_{i,t-1}$ the change in revenue from the previous fiscal year of REIT i. All regression variables including the intercept are scaled by the previous fiscal year’s total assets of REIT i ($A_{i,t-1}$).

Following an analogous approach, cost of goods sold is estimated as a function of revenue:

$$\frac{COGS_{i,t}}{A_{i,t-1}} = \beta_0 \frac{1}{A_{i,t-1}} + \beta_1 \frac{REV_{i,t}}{A_{i,t-1}} + \beta_2 \frac{\Delta REV_{i,t}}{A_{i,t-1}} + \beta_3 \frac{\Delta REV_{i,t-1}}{A_{i,t-1}} + \varepsilon_{i,t} \,, \quad (4.9)$$

with $COGS_{i,t}$ denoting cost of goods sold of REIT i in fiscal year t and all other variables defined similarly to Equation (4.8). Again, all variables are scaled by lagged total assets.

Equations (4.8) and (4.9) are estimated individually for each fiscal-year and REIT property type, while a required minimum number of 10 observations per property type and year is set so as to obtain meaningful estimates of normal revenue- and cost levels. The residuals $\varepsilon_{i,t}$ of Equations (4.8) and (4.9) constitute the proxies for abnormal revenue and cost, respectively. In order to create a joint measure for real earnings management, this paper follows common practice (see Cohen and Zarowin, 2010; or Chen, Rhee, Veeraraghavan and Zolotoy, 2015) and multiplies the residual of Equation (4.9) for a stock year by minus 1 and adds the result to the residual of Equation (4.8). The created earnings management measure EM takes high (low) values for REIT stock years with abnormally high (low) combined values for revenue and negative cost.

To test for a relationship between development activity and EM, this paper again uses Ambrose and Bian's (2010) approach as a starting point. The applied model includes fiscal-year and property-type fixed-effects and can be written as:

$$\begin{aligned} EM_{i,t} = \beta_0 + \beta_1 \frac{DEV_{i,t-1}}{CON_{i,t-1}} + \beta_2 COMP_{i,t} + \beta_3 AGR_{i,t} + \beta_4 ROA_{i,t} \\ + \beta_5 TOQ_{i,t-1} + \beta_6 LEV_{i,t-1} + \beta_7 SIZ_{i,t-1} + \beta_8 \Delta DS_{i,t} \\ + \beta_9 VCF_{i,t-1} + YEAR_t + PROPTYPE_i + \varepsilon_{i,t} \,. \end{aligned} \tag{4.10}$$

EM for REIT i in fiscal year t is regressed against the development activity measure DEV of i at the end of the previous fiscal year. As a robustness check, EM is also regressed against the development measure calculated without the inclusion of development land (DL) in its numerator. This additional specification (CON) is performed to rule out the possibility that REITs with a large amount of development land might have reduced potential to conduct earnings management and consequently cause a spurious, negative relationship with EM. Following Ambrose and Bian, the models control for firm size (SIZ), leverage (LEV), return on assets (ROA), Tobin's Q (TOQ), change in diluted shares (ΔDS), and cash flow volatility (VCF). The first three control variables are constructed similarly to those in the development activity – crash risk models. Tobin's Q (TOQ) is the ratio of market value of assets to book value of assets.[47] SIZ, LEV, and TOQ enter the regressions lagged once. ΔDS is the percentage change in fully diluted shares between the current and the previous fiscal-year end. VCF is the standard deviation of the previous three fiscal years' operating

[47] Following common practice, the book value of debt is used as a proxy for market value of debt.

cash flow.[48] Ambrose and Bian remove the between-subject variation in their models by using REIT fixed-effects, which is unsuitable for the purposes of this paper. To keep a possible impact of endogeneity issues low, besides the addition of property-type dummies, controls for stock compensation (COMP) and asset growth (AGR) are added into the equations. COMP are the stock-compensation expenses of REIT i in year t, scaled by lagged total assets. The intuition behind this control variable is to incorporate the potential positive relationship between the degree of performance-based compensation of a REIT and the tendency of its management to manipulate earnings. AGR is the percentage change in total assets for REIT i from year t-1 to year t. Internally- or externally-growing REITs might experience higher costs and lower earnings in the medium term, due to adaption processes (even without any earnings manipulation). This effect might lead to a spurious negative relationship between earnings management, as measured with Equations (4.8) and (4.9), and property development, when not controlled for. Residuals from Equation (4.8) and (4.9) used to calculate EM and the set of control variables are winsorized at the tails of 0.5% and 99.5% to limit the influence of outliers, as suggested by Ambrose and Bian. The results for the development activity – earnings management regressions can be found in Table 4.4.

The availability of stock compensation data and the threshold of 10 REITs per property type and fiscal year reduce the sample size to 1,403 for the DEV model and 1,429 for the CON model. The signs of the coefficients of the control variables are completely in line with Ambrose and Bian's findings for their separate regressions for cost- and revenue management. The controls not utilized by Ambrose and Bian COMP and AGR furthermore show the expected relationship to earnings management. While the compensation coefficient (COMP) is positive but insignificant, asset growth (AGR) can significantly explain part of the estimated earnings-management variation. It thus seems to serve as a corrective for the potential shortcoming of the earnings management models of not taking into account the effect of firm growth on revenue and cost.

[48] Due to the limited space, this paper refers to Ambrose and Bian (2010) for a detailed discussion of the reasoning behind the choice of the control variables.

Table 4.4: REIT Earnings Management - Development Activity

	Dependent variable:	
	EM_t	
DEV_{t-1}	-0.061**	
	(0.027)	
CON_{t-1}		-0.071**
		(0.030)
$COMP_t$	0.555	0.553
	(1.023)	(1.030)
AGR_t	0.055***	0.055***
	(0.007)	(0.007)
ROA_t	0.279***	0.278***
	(0.080)	(0.079)
TOQ_{t-1}	-0.002	-0.001
	(0.007)	(0.007)
LEV_{t-1}	0.025*	0.025*
	(0.014)	(0.014)
SIZ_{t-1}	0.002	0.002
	(0.002)	(0.002)
ΔDS_t	0.010	0.009
	(0.008)	(0.008)
VCF_{t-1}	-0.00003	-0.00003
	(0.0001)	(0.0001)
Property-type fixed-effects	YES	YES
Fiscal-year fixed-effects	YES	YES
Observations	1,403	1,429
R^2	0.168	0.167
Significance levels:		*$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$*

Notes: Table 4.4 reports the results for the regressions of upward earnings management on development activity for the sample of listed US equity REITs from the fiscal year 2001 to 2017. EM denotes a measure for earnings management constructed as a combined figure of revenue-upward and cost-downward management for a particular stock year. The respective revenue and cost measures are calculated as the residuals of yearly regressions for each REIT property type estimating the normal level of revenue and cost for each stock year. DEV denotes the sum of development land and construction in progress, divided by total assets for a stock year, CON is DEV without the inclusion of development land in the numerator. Regressions contain property-type- and fiscal-year fixed-effects, as well as heteroscedasticity robust standard errors (in parenthesis) adjusted for clustering at the REIT level.

The key finding from the earnings management regressions is the negative effect of DEV and CON on EM. DEV and CON are both significant at the 5% level. Without winsorization, the development activity coefficients barely change (DEV: -0.061; CON: -0.072) and retain their significance. REITs engaging in property development seem to show a lower tendency to manage their earnings upward, compared to their non-developing peers. Furthermore, the effect cannot be attributed solely to the holding of development land. Altogether, this finding can be seen as an additional piece of evidence to support the hypothesis that development activity is interpretable as a proxy for long-termism of the respective REIT's management. With reference to the findings of Graham, Harvey and Rajgopal (2005), property-developing REITs seem to accept short-term turmoil in exchange for the prospect of long-term value.

4.7 Conclusion

Real estate development can be an attractive way for equity REITs to bypass the historically high multiples demanded for acquiring completed properties. Arguments both for and against the development involvement of REITs can be proposed. Proponents usually argue that particularly during up-markets, access to attractive properties can only be achieved through in-house development. Critics on the other hand stress the additional risks resulting from deviation from the standard business profile of an equity REIT. Under the assumption of availability of the required capacities to execute in-house development projects, the question of whether such projects on average are beneficial for shareholders has to be answered empirically. An analysis of the stock price crash risk of property-developing REITs can be one aspect of that evaluation. Considering that property developments involve a substantial amount of risk and uncertainty, their execution should intuitively lead to an increase in crash propensity for the respective REIT. On the other hand, the performance of development projects is voluntary for equity REITs. This implies that management commitment to postponing relatively safe short-term revenue generated through the ownership and operation of properties, in favor of uncertain future benefits from development endeavors, could itself be interpreted as a sign of long-termism and thus decrease crash risk.

The paper finds that an increase in development activity on average results in a reduced stock price crash risk for the sample REITs. This finding is robust in terms of the specification of the market model used to calculate firm-specific returns, the measurement of crash risk, the exclusion of return-outliers, construction of the development-activity variable, splitting of the sample into a pre- and post-crisis subperiod, and winsorization of regression inputs. Further analysis shows that property-developing REITs in the sample also exhibit a reduced tendency toward upward earnings management, a common proxy for short-termism. Altogether, the findings suggest that equity REITs developing own properties engage in less short-term behavior of holding back negative news and instead act transparently towards their shareholders. There are certain implications for REIT investors resulting from these findings. Although the reduced crash propensity does not allow for any conclusions to be drawn on the success of equity REITs in performing on-balance development projects, it is revealing to find that REITs engaging in real estate development are not inclined to perform such projects in a myopic fashion. Quite the contrary seems to be true, as property-developing REITs' managements do not seem to exploit the opacity resulting from development projects and rather tend to disclose emerging bad news transparently to their shareholders. Value-destroying stock price crashes thus appear to be a subpar problem for investors in property-developing REITs.

This paper was intended as a starting point for more research on the relationship between an equity REIT's business model and its stock price crash risk. For instance, the diversification in terms of property types or investment regions could be further characteristics affecting the crash propensity of a REIT, which are worthy of empirical analysis. It also seems plausible that REITs with a stable portfolio differ in their crash behavior from REITs with significant property turnover. Properties which have been held for long time periods might be well understood by shareholders, as a result reduce firm opacity and consequently the potential for bad news hoarding. If this argument holds, stable portfolios might reduce crash risk. A field so far completely neglected in the literature are the consequences of crashes for REITs. In this context, an examination of which firms are capable of recovering quickly from crashes and which are not seems particularly worthwhile.

4.8 References

Ambrose, B. W. and X. Bian, Stock market information and REIT earnings management, Journal of Real Estate Research, 2010, 32:1, 101–38.

An, H., Q. Wu and Z. Wu, REIT crash risk and institutional investors, The Journal of Real Estate Finance and Economics, 2016, 53:4, 527–58.

An, H. and T. Zhang, Stock price synchronicity, crash risk, and institutional investors, Journal of Corporate Finance, 2013, 21, 1–15.

Ball, R., Market and political/regulatory perspectives on the recent accounting scandals, Journal of Accounting Research, 2009, 47:2, 277–323.

Bhojraj, S. and R. Libby, Capital market pressure, disclosure frequency-induced earnings/cash flow conflict, and managerial myopia, The Accounting Review, 2005, 80:1, 1–20.

Brounen, D. and P. M. A. Eichholtz, Development involvement and property share performance: International evidence, The Journal of Real Estate Finance and Economics, 2004, 29:1, 79–97.

Brounen, D., P. M. Kanters and P. M. A. Eichholtz, The effects of property development activities on the performance of REITs, Real Estate Finance, 2000, 16:4, 17–22.

Bushee, B. J., The influence of institutional investors on myopic R&D investment behavior, The Accounting Review, 1998, 73:3, 305–33.

Byrne, P., Risk, uncertainty and decision-making in property development, 2nd edition, Spon - Taylor & Francis: London, 1996.

Callen, J. L. and X. Fang, Institutional investor stability and crash risk: Monitoring versus short-termism?, Journal of Banking & Finance, 2013, 37:8, 3047–63.

Callen, J. L. and X. Fang, Short interest and stock price crash risk, Journal of Banking & Finance, 2015, 60, 181–94.

Callen, J. L. and X. Fang, Crash risk and the auditor-client relationship, Contemporary Accounting Research, 2017, 34:3, 1715–50.

Cao, H. H., J. D. Coval and D. Hirshleifer, Sidelined investors, trading-generated news, and security returns, The Review of Financial Studies, 2002, 15:2, 615–48.

Chang, X., Y. Chen and L. Zolotoy, Stock liquidity and stock price crash risk, Journal of Financial and Quantitative Analysis, 2017, 52:4, 1605–37.

Chen, J., H. Hong and J. C. Stein, Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices, Journal of Financial Economics, 2001, 61:3, 345–81.

Chen, Y., S. G. Rhee, M. Veeraraghavan and L. Zolotoy, Stock liquidity and managerial short-termism, Journal of Banking & Finance, 2015, 60, 44–59.

Chung, K. H. and H. Zhang, A simple approximation of intraday spreads using daily data, Journal of Financial Markets, 2014, 17, 94–120.

Cohen, D. A., A. Dey and T. Z. Lys, Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods, The Accounting Review, 2008, 83:3, 757–87.

Cohen, D. A. and P. Zarowin, Accrual-based and real earnings management activities around seasoned equity offerings, Journal of Accounting and Economics, 2010, 50:1, 2–19.

Cook, R. D., Detection of influential observation in linear regression, Technometrics, 1977, 19:1, 15–8.

Cook, R. D. and S. Weisberg, Residuals and influence in regression, Chapman & Hall: New York, 1982.

Dimson, E., Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics, 1979, 7:2, 197–226.

Edmans, A., Blockholder trading, market efficiency, and managerial myopia, The Journal of Finance, 2009, 64:6, 2481–513.

Francis, B., I. Hasan and L. Li, Abnormal real operations, real earnings management, and subsequent crashes in stock prices, Review of Quantitative Finance and Accounting, 2016, 46:2, 217–60.

Geltner, D., A. Kumar and A. van de Minne, Super-normal profit in real estate development, University of Connecticut School of Business Research Paper, 2019, No. 19-14 (2019/09/05).

Gerbich, M., M. Levis and P. Venmore-Rowland, Property investment and property development firm performance around initial public offerings and rights offerings: U.K. evidence, The Journal of Real Estate Finance and Economics, 1999, 18:2, 207–38.

Graham, J. R., C. R. Harvey and S. Rajgopal, The economic implications of corporate financial reporting, Journal of Accounting and Economics, 2005, 40:1-3, 3–73.

Habib, A., M. M. Hasan and H. Jiang, Stock price crash risk: Review of the empirical literature, Accounting & Finance, 2018, 58:S1, 211–51.

Hong, H. and J. C. Stein, Differences of opinion, short-sales constraints and market crashes, Review of Financial Studies, 2003, 16:2, 487–525.

Hutton, A. P., A. J. Marcus and H. Tehranian, Opaque financial reports, R^2, and crash risk, Journal of Financial Economics, 2009, 94:1, 67–86.

Jin, L. and S. C. Myers, R^2 around the world: New theory and new tests, Journal of Financial Economics, 2006, 79:2, 257–92.

Kastenbaum, R., The dimensions of future time perspective, an experimental analysis, The Journal of General Psychology, 1961, 65, 203–18.

Kim, J.-B., Y. Li and L. Zhang, CFOs versus CEOs: Equity incentives and crashes, Journal of Financial Economics, 2011a, 101:3, 713–30.

Kim, J.-B., Y. Li and L. Zhang, Corporate tax avoidance and stock price crash risk: Firm-level analysis, Journal of Financial Economics, 2011b, 100:3, 639–62.

Kim, J.-B., Z. Wang and L. Zhang, CEO overconfidence and stock price crash risk, Contemporary Accounting Research, 2016, 33:4, 1720–49.

Kim, J.-B. and L. Zhang, Accounting conservatism and stock price crash risk: Firm-level evidence, Contemporary Accounting Research, 2016, 33:1, 412–41.

Kim, Y., H. Li and S. Li, Corporate social responsibility and stock price crash risk, Journal of Banking & Finance, 2014, 43, 1–13.

Kothari, S. P., S. Shu and P. D. Wysocki, Do managers withhold bad news?, Journal of Accounting Research, 2009, 47:1, 241–76.

Laverty, K. J., Managerial myopia or systemic short-termism? The importance of managerial systems in valuing the long term, Management Decision, 2004, 42:8, 949–62.

Loizou, P. and N. French, Risk and uncertainty in development: A critical evaluation of using the Monte Carlo simulation method as a decision tool in real estate development projects, Journal of Property Investment & Finance, 2012, 30:2, 198–210.

MacFarlane, J., The use of simulation in property investment analysis, Journal of Property Valuation and Investment, 1995, 13:4, 25–38.

Porter, M. E., Capital choices: Changing the way America invests in industry, Journal of Applied Corporate Finance, 1992, 5:2, 4–16.

Samuel, C., Does shareholder myopia lead to managerial myopia? A first look, Applied Financial Economics, 2000, 10:5, 493–505.

Stein, J. C., Efficient capital markets, inefficient firms: A model of myopic corporate behavior, The Quarterly Journal of Economics, 1989, 104:4, 655–69.

5 Conclusion

5.1 Executive Summary

In an attempt to contribute to a better understanding of real estate markets in terms of their informational efficiency, this thesis focuses on two problem areas: (1) The effect of noise on the market, and (2) the effect of information withheld from the market. The contribution to the first problem field consists of two papers, one analyzing the predictability of direct property market returns with an estimation of market participants' sentiment, and another scrutinizing the effects of sentiment on direct property market liquidity. The third paper is devoted to the information-hoarding tendency of property-developing equity REIT officials, and the propensity of such REITs to experience stock price crashes as a consequence of the hoarding. The following sections provide a brief summary of the three contributions, an outlook for future research, and some concluding remarks.

Summary, Paper 1:
"On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach"

In an attempt to overcome several problems associated with existing market sentiment measures, the paper develops a novel text-based indicator. The indicator utilizes an ANN in order to determine the polarity of 66,070 news texts provided by the *S&P Global Market Intelligence database* for the US real estate market between 2006 and 2018. Instead of using headlines (as, for example in Hausler, Ruscheinsky and Lang, 2018) or abstracts (see e.g. Beracha, Lang and Hausler, 2019), the paper considers full articles and thereby incorporates the full range of views and underlying intentions of the authors. ANNs do not require a predefined set of rules to indicate news sentiment and thus potentially allow for the extraction of a richer information structure, compared, for example, to the use of sentiment dictionaries. The training of the ANN is executed through distant supervision. Specifically, 17,822 distant supervision-labeled long- and short ideas from the online investment advisory platform *Seeking Alpha*, are collected. The texts form an almost optimal training corpus for the sentiment training of the ANN, as they include either a distinct positive or a distinct negative sentiment towards the asset or market in question. The word universe that is relevant for the *Seeking Alpha* investment ideas is furthermore, thematically closely related to that of the S&P news articles, thus minimizing potential semantic problems due to a different term usage in the two corpora. The outlined distant supervision training process precludes bias that could otherwise arise from the subjective classification of a training sample through a limited number of human annotators. After the ANN training process, the S&P articles are assigned a sentiment score, ranging between 0 (complete negativity) and 1 (complete positivity) and averaged on a monthly basis so as to obtain a market-wide sentiment indication. The resulting monthly sentiment index is then used to attempt predictions of the development of the *CoStar Commercial Repeat-Sale Index* (CCRSI). The results from a vector autoregressive framework show that the constructed sentiment index is significantly positively related to future returns of the CCRSI. Additionally, the application of Markov-switching models exhibits a varying influence of the sentiment indicator on returns during up- and down-market periods. The relationship seems to be stronger during down-market periods, while being insignificant (but nevertheless positive) for up-market periods. The sentiment indicator is then tested in a logit forecasting framework. In this context, it demonstrates predictive potential, but struggles with short-term market fluctuations. Nevertheless, with a novel approach to sentiment estimation involving several benefits in terms of ease of construction and flexibility, the paper lays the foundation for more research in the field of real estate market forecasting via deep-learning-based sentiment analysis.

Summary, Paper 2:

"Artificial Intelligence, News Sentiment, and Property Market Liquidity"

Liquidity is a fundamental requirement for information-efficient trading. Fully rational markets should not be affected by the actions of sentiment traders and exhibit a high degree of liquidity at all times. Nevertheless, theoretical papers like Kyle (1985) in combination with De Long, Shleifer, Summers and Waldmann (1990), or Baker and Stein (2004) suggest that a departure from the assumption of efficient markets can explain a potential positive relationship between sentiment and liquidity. This relationship has been confirmed empirically for the general stock market by Liu (2015) and for the direct real estate market by Freybote and Seagraves (2018). A consideration so far mostly neglected in the sentiment literature on direct real estate markets is the large time lag between the initiation of a property transaction and its completion (see e.g. Devaney and Scofield, 2015). Assuming that the sentiment of market participants might serve as a trigger for transactions, the effect of sentiment on liquidity should manifest itself only slowly over the months following the initial transaction impulse. This paper thus makes use of an approach which takes into account the slow pace of direct property markets, i.e. both contemporary as well as lagged sentiment is utilized as an explanatory variable. The paper again employs the sentiment measurement approach developed for Paper 1, as its advantages of short aggregation periods and a direct link to the US property market imply an accurate and timely estimation of market participant sentiment. The findings of the paper suggest that sentiment in fact exerts a lagged effect on the liquidity of the direct US property market for the study period between 2006 and 2018. Price-impact-of-trades, proxying for the depth- and resilience dimensions of liquidity, exhibits a negative relationship to sentiment. This implies that higher sentiment reduces the impact of transactions on market prices, and therefore increases liquidity. In accordance with that outcome, transaction-volume, proxying for the breadth of the market, displays a positive relationship with sentiment. Regressions making use of the price-impact measure of liquidity exhibit a significant relationship for up to 8 months in the past, confirming the lagged influence of market sentiment on liquidity measures. Sentiment therefore seems to work as a trigger for transactions. Due to the long transaction periods in direct real estate markets, the effects only manifest themselves during the following months in the common measures for market liquidity. Thus, instead of focusing on ‘stale’ contemporary liquidity measures, direct property market participants should keep market sentiment in view, so as to draw conclusions on the prevailing liquidity situation.

Summary, Paper 3:
“REIT Property Development and Stock Price Crash Risk”

The third paper departs from the area of market sentiment and is instead dedicated to the dispersion of relevant firm information. The paper deals with the question of whether equity REITs transparently disclose negative information to their shareholders. In other words, rather than analyzing the problem of noise penetrating into the market, the paper looks at the problem of information that is deliberately withheld from the market. Jin and Myers’ (2006) theory on the formation of stock price crashes suggests that the hoarding of bad news about the firm through firm executives can result in a pileup of adverse firm information, which can lead to a price crash after its disclosure. An informational advantage compared to outsiders enables executives to display such news-hoarding behavior, and the personal reasons can be manifold. Career concerns (Kothari, Shu and Wysocki, 2009), or the desire of executives to maximize their own compensation (Kim, Li and Zhang, 2011) are two examples. Stock price crash literature can be roughly split up into two areas. The first entails research considering the reaction of capital markets to the disclosure of hoarded bad news. An, Wu and Wu (2016) analyze this question for the US equity REIT market and find, among other things, that a high share of transient institutional investors increases the risk of REITs experiencing stock price crashes. The second area of crash research is represented by the analysis of the contribution of firm-internal factors to the crash risk of a company. This field has so far not received any attention in the REIT literature. The aim of the paper was thus to focus on one aspect of the equity REIT business model, namely property development activity, and to open the field for research on internal crash risk factors of equity REITs. The aspect of development activity is particularly interesting in this respect, as Jin and Myers' (2006) theory can explain both a positive as well as a negative relationship between stock price crash risk and the property development activity of REITs. On the one hand, developments are opaque ventures involving vast potential for the occurrence of bad news, which form the basis for a subsequent price crash. On the other hand, developments as long-term endeavors require management dedication and consequently a long-term view on the company's success. To carry out property developments, short-term interests have to be sacrificed in the interest of long-term goals. Executive short-term behavior in the form of bad news hoarding thus seems contradictory from this point of view, for developing REITs. The paper finds that a higher share of property development indeed leads to a decrease in stock price crash risk for a sample of US equity REITs between 2001 and 2017. As further analysis shows, property-developing REITs in the sample also exhibit a lower tendency for upward earnings-management, a common proxy for managerial short-termism in the literature. Altogether, the evidence suggests that development projects of REITs are not performed myopically, and that respective managers seems to refrain from short-termism practices like bad news hoarding.

5.2 Final Remarks and Opportunities for Further Research

Successful discrimination between noise and information is a pivotal task for financial market participants. With that in mind, economic research should attempt to contribute to the informational improvement of markets by fostering the reduction of noise and an increase in actual information. The aim of this thesis was therefore to single out two small areas for such improvements. First, the work reveals the effects of investor sentiment on direct property markets. Disclosing such effects might help rational market participants to circumvent problems associated with market exuberance, or at least to improve their awareness in this regard. Awareness on the part of an increasing share of market participants of the effects of sentiment on the market contributes to the minimization of its adverse impact. Second, efficient markets must provide a steady flow of relevant firm-specific information. Behaviors like bad news hoarding and other principal-agent problems must therefore be minimized in order to allow prices to reflect all relevant information at all times.

A logical next step would involve a combination of both research areas. Building on Jin and Myers' (2006) theory, arguments can be made for a time-varying effect of market sentiment on the bad news hoarding tendency of firm executives. Dependent on the prevailing market sentiment, managers might adjust their willingness to disclose new internal information. In fear of a strong adverse reaction of sentiment-driven investors during periods of negative market sentiment, the tendency of managers to hoard bad news might increase, and vice versa for positive sentiment periods. The result would be an increased number of idiosyncratic stock price crashes during negative-sentiment periods. Additionally, recent years have witnessed a steady increase in available online text data, relative ease in collecting articles from a multitude of such data sources via web crawling, and the development of powerful deep-learning algorithms. The medium-term goal should thus be to raise the analysis to a firm-specific level and test the predictability of idiosyncratic crashes on the basis of a firm-level sentiment estimate.

Nowadays, few scholars rely on the premise of fully efficient financial markets. The occurrence of flash crashes on the one hand and of exuberant rallies on the other, make it increasingly difficult to adhere to the assumption of prices which reflect all relevant information at all times, even for the comparatively efficient stock market. Nevertheless, despite perfectly efficient markets remaining illusionary, the objective of taking marginal but constant steps in their direction is certainly worthwhile. Recent financial crises, and particularly the market turmoil of 2007-08, which originated in the US real estate sector, have brought forth increasingly louder calls for the impairment of frictionless trading through transaction taxes or even trade prohibitions in certain fields. The role of economic scholars in this context is surely to constantly question simple solutions and to 'drill' for the underlying relationships. Outlining ways to minimize agency-problems and to create an awareness of human susceptibility to sentiment-driven decision making are two approaches which cannot claim to be easily achievable, but they can claim to be real solutions. The aim of this thesis was accordingly to shed some light on those fields. Discovering inefficiencies and creating an awareness of their existence is the only way to overcome them.

5.3 References

An, H., Q. Wu and Z. Wu, REIT crash risk and institutional investors, The Journal of Real Estate Finance and Economics, 2016, 53:4, 527–58.

Baker, M. and J. C. Stein, Market liquidity as a sentiment indicator, Journal of Financial Markets, 2004, 7:3, 271–99.

Beracha, E., M. Lang and J. Hausler, On the relationship between market sentiment and commercial real estate performance - A textual analysis examination, Journal of Real Estate Research, 2019, 41:4, 605–37.

De Long, J. B., A. Shleifer, L. H. Summers and R. J. Waldmann, Noise trader risk in financial markets, Journal of Political Economy, 1990, 98:4, 703–38.

Devaney, S. and D. Scofield, Liquidity and the drivers of search, due diligence and transaction times for UK commercial real estate investments, Journal of Property Research, 2015, 32:4, 362–83.

Freybote, J. and P. A. Seagraves, The impact of investor sentiment on commercial real estate market liquidity, Journal of Real Estate Research, 2018, 40:4, 597–627.

Hausler, J., J. Ruscheinsky and M. Lang, News-based sentiment analysis in real estate: A machine learning approach, Journal of Property Research, 2018, 35:4, 344–71.

Jin, L. and S. C. Myers, R^2 around the world: New theory and new tests, Journal of Financial Economics, 2006, 79:2, 257–92.

Kim, J.-B., Y. Li and L. Zhang, CFOs versus CEOs: Equity incentives and crashes, Journal of Financial Economics, 2011, 101:3, 713–30.

Kothari, S. P., S. Shu and P. D. Wysocki, Do managers withhold bad news?, Journal of Accounting Research, 2009, 47:1, 241–76.

Kyle, A. S., Continuous auctions and insider trading, Econometrica, 1985, 53:6, 1315–35.

Liu, S., Investor sentiment and stock market liquidity, Journal of Behavioral Finance, 2015, 16:1, 51–67.