Essays on the Influence of Textual Sentiment in Real Estate Markets

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

vorgelegt von:
JOCHEN HAUSLER

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

Tag der Disputation: 27. November 2019

Essays on the
Influence of Textual Sentiment
in Real Estate Markets

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 General Motivation and Theoretical Foundation

The analysis of a potential influence of sentiment on asset markets in general and real estate markets in particular rests on two crucial assumptions, controversially discussed by scholars of behavioral finance and market efficiency supporters. The first assumption is that of investors subject to sentiment, which contrasts the standard finance model of unemotional and rational investors. Hereby, sentiment is commonly referred to as the existence of beliefs about cash flows and risks not explained by fundamentals (Baker and Wurgler, 2007). This assumption was pioneered for financial markets by Long *et al.* (1990) emphasizing that models incorporating „noise traders" in the spirit of Kyle (1985) and Black (1986) are able to explain financial anomalies. The second essential assumption is that of limited possibilities for arbitrage (Shleifer and Vishny, 1997). When betting against sentiment-driven investors is risky and costly, prices are prevented from being aggressively forced back to fundamentals by arbitrageurs (Baker and Wurgler, 2007).

Considering the primary case of application of these models are highly efficient stock markets, real estate as an "imperfect", alternative asset class does not evidently provide a good fit. However, through its inherent imperfections, the nature of "real" assets abets these two aforementioned assumptions in property markets as well. Buildings are immobile, heterogeneous and of large investment volume, which leads to prolonged transaction periods and benefits local agents acting in segmented and informationally inefficient regional markets. Driven by asymmetric information between buyer and seller, sentiment-induced trading behavior is promoted. This notion is supported by findings of Gallimore and Gray (2002) which stress the high importance of investor sentiment in property investment decision-making. Yet these findings are not limited to transaction activities. With respect to property valuations in the UK, Crosby *et al.*

(2010) also demonstrate the significant role of client influence on appraisal outcomes. Additionally, non-existing short sale opportunities for direct real estate limit the possibilities to eliminate mispricing. Correspondingly, Clayton *et al.* (2009) were able to provide evidence that investor sentiment also impacts commercial real estate pricing in the US.

With a growing body of literature on sentiment in finance and real estate markets, research focus gradually shifted away from the question of relative importance and existence of sentiment to the question of appropriate measurement of sentiment and the quantification of its influence. Textual sentiment analysis, i.e. the attempt to extract evaluations, attitudes and emotions from text corpora (Liu, 2012, preface), is in fact a more recent approach. More traditional alternatives are surveys and market-based sentiment proxies such as NAV discounts, mortgage fund flows, property index transaction frequencies, past returns and buy-sell imbalances (see e.g. Lin *et al.*, 2009; Clayton *et al.*, 2009; Freybote and Seagraves, 2017).

Despite their frequent application in real estate research (see e.g. Clayton *et al.*, 2009; Das *et al.*, 2015; Freybote, 2016), surveys such as the *Real Estate Research Corporation* sentiment indicator are, by their very nature, associated with material disadvantages. They are time-consuming, costly and might also reflect false sentiment when survey participants are wrongly incentivized or do intentionally provide wrong answers. Additionally, due to their usually low frequency as well as time-lag bias, they are less useful for time series analysis. In contrast, market-based proxies are fundamentally incentivized as they proxy the behavior of market participants but also have respective drawbacks. Besides being highly dependent on the underlying theory, they might also ignore unperceived, but valuable factors of decision-making, which are not captured by mere quantitative market measures.

However, due to an increasing number of digital text documents, a rise in computational power and the development of new classification techniques, textual sentiment analysis gained more and more attention in more recent years. When using real estate related textual documents as “sentiment provider”, the measures are not only directly linked to the asset class but also allow for short and flexible sentiment aggregation periods. As the availability of text-based sentiment measures solely depends on the frequency of publication of the underlying text corpus, these indicators not only surpass many other sentiment measures with regard to actuality but also enable an investigation of whether news sentiment leads market movements or vice versa (see Tetlock, 2007 for an in-depth discussion of the lead-lag theory).

Especially sentiment dictionaries allow for a straightforward and transparent application because of readily available software solutions (Kearney and Liu, 2014). However, in direct comparison, the performance of more technically advanced machine-learning approaches is usually higher (Li, 2010). Furthermore, artificial neural networks, as part of so-called deep-learning classifiers, have the potential to extract a much richer information structure from textual documents. With more and more data available for training, they provide a better scalability and are predestined for real time analytics and big data applications, which further deems them superior to traditional indicators.

In spite of these theoretical advantages, the potential of textual sentiment indicators in real estate has not been explored in practice. Although there is some related literature, studies such as Soo (2015) and Walker (2014) are – except of Ruscheinsky *et al.* (2018) – mostly limited to the housing market and rely on a dictionary-based approach. With respect to textual sentiment in real estate, machine- and deep-learning classifiers have been completely ignored. Accordingly, the following four studies bridge the gap and shed light on the potential of such textual sentiment indicators within the neglected area of commercial real estate. For the first time, the capabilities of a machine- and a deep-learning classifier for predicting direct and securitized market returns and liquidity within the US are assessed. Additionally, the relationship with up- and down-market periods, market regimes and out-of-sample forecasting performance are studied. Overall, this should answer the question whether real estate news analytics by means of textual sentiment classifiers in general and machine- and deep-learning algorithms in particular can be perceived as a valuable and innovative source of market sentiment and is able to provide researchers and practioneers with a reliable leading market indicator.

Accordingly, paper 1 kicks off the series with an attempt to predict private commercial real estate market returns by analyzing abstracts from the *Wall Street Journal* through the application of a sentiment dictionary. Subsequently, paper 2 refines the approach in several dimensions: The study is conducted on a monthly instead of quarterly basis, relies on a support vector machine (SVM) i.e. a machine-learning approach for sentiment classification and also includes the securitized real estate market in the United States. With paper 3 and paper 4, research delves into the sphere of deep-learning by facilitating an artificial neural network (ANN) for sentiment extraction. Accounting for different market regimes, the created indicator is once more related to market returns. Additionally, in- and out-of-sample forecasting performance in up- and down-market periods and the link to market liquidity is evaluated. Using a distant supervision-labelled dataset for training eliminate the need for manual classification and thus represents an additional innovation.

Table 1.1 highlights the main features of the four research studies. Subsequently, Section 1.2 elaborates the research questions examined. Section 1.3 provides an overview on submissions and conference presentations before Sections 2 to 5 present the studies in their entirety. Section 6 concludes and discusses limitations and future research opportunities.

Table 1.1: Course of Analysis

	Paper 1	**Paper 2**	**Paper 3**	**Paper 4**
Text corpus	Abstracts *Wall Street Journal* (35,398 abstracts)	Headlines *S&P Global Market Intelligence Database* (54,530 headlines)	Full articles *S&P Global Market Intelligence Database & Seeking Alpha* (66,070 articles) (17,822 investment ideas)	Full articles *S&P Global Market Intelligence Database & Seeking Alpha* (66,070 articles) (17,822 investment ideas)
Classifier	Sentiment dictionary	Machine-learning (SVM)	Deep-learning (ANN)	Deep-learning (ANN)
Market	Direct real estate, US	Direct & indirect real estate, US	Direct real estate, US	Direct real estate, US
Frequency	Quarterly	Monthly	Monthly	Monthly
Sample Period	2001:Q1 to 2016:Q4	2005:M01 to 2016:M12	2006:M01 to 2018:M12	2006:M01 to 2018:M12
Research focus	Market returns	Market returns	Market returns	Market liquidity

1.2 Research Questions

While all four research articles revolve around the primary question of the influence of textual sentiment in real estate markets, each study concentrates on partial aspects which – at the end – are intended to condense to an overall picture. As initial research, paper 1 and 2 take advantage of a set of predefined hypotheses or research questions. However, paper 3 and paper 4 attempt to provide a more “out-of-the-box”-approach and explore the general nature of a deep-learning sentiment classifier with respect to market liquidity and returns. In more detail, the following aspects are addressed in the respective papers.

Paper 1 | On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination

- ***Research question 1:***
  Does real estate market sentiment extracted by means of a sentiment dictionary predict future returns of the private commercial real estate market in the US?

- ***Research question 2:***
  Is there evidence of a one- or of a bi-directional relationship? More formally, does media-expressed sentiment predict future returns of private commercial real estate, while returns on private commercial real estate do not predict future media-expressed sentiment?

- ***Research question 3:***
  When considering appreciation instead of total returns, do results with respect to the influence of media-expressed sentiment still hold?

- ***Research question 4:***
  Is sentiment-based predictability of total returns asymmetric i.e. is there higher predictability power during periods of slower market growth?

Paper 2 | News-Based Sentiment Analysis in Real Estate:
A Machine-Learning Approach

- ***Research question 1:***
  Can sentiment measures created via machine-learning predict the securitized commercial real estate market?

- ***Research question 2:***
  Is the predictive power different for the direct real estate market?

- ***Research question 3:***
  How do the created sentiment indicators perform in addition to established sentiment measures?

- ***Research question 4:***
  Is there evidence that market participants react differently to negative news in contrast to positive ones?

Paper 3 | On the Predictive Potential of Investor Sentiment:
A Deep-Learning Approach

- ***Research question 1:***
  Is textual sentiment extracted from news articles with the help of an artificial neural network able to explain direct real estate market returns?

- ***Research question 2:***
  Is there any evidence of a non-linear relationship between sentiment and market returns? Should econometric models account for different market regimes?

- ***Research question 3:***
  Does the sentiment indicator show some binary return forecast potential? Thus, is textual sentiment capable of forecasting up- and down-market periods?

Paper 4 | Artificial Intelligence, News Sentiment and Property Market Liquidity

- ***Research question 1:***
  Do results provide any evidence of explanatory power of the sentiment indicator with respect to changes in market liquidity?

- ***Research question 2:***
  Do results differ with respect to the depth, resilience and breadth dimensions of market liquidity? Hence, can one find evidence of Baker and Stein's (2004) hypotheses of a negative relationship between sentiment and price impact as well as of a positive relationship of sentiment and trading volume?

- ***Research question 3:***
  Do results still hold when using other measures of market liquidity as alternatives to Amihud's (2002) price impact measure as well as transaction volume?

1.3 Submissions and Conference Presentations

While the main purpose and study design of the four presented research articles has been highlighted in Sections 1.1 and 1.2, Section 1.3 complements the previous sections with details regarding submission to journals, publication status and conference presentations.

Paper 1 | On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination

Authors:
Eli Beracha, Jochen Hausler and Marcel Lang

Submission:

Journal:	Journal of Real Estate Research
Submission date:	04/02/2018
Current Status:	Accepted (11/30/2018)

Conference presentations:

The paper was presented at the 34th Annual Conference of the American Real Estate Society (ARES) in Bonita Springs, US.

Paper 2 | News-Based Sentiment Analysis in Real Estate: A Machine-Learning Approach

Authors:

Jochen Hausler, Jessica Ruscheinsky, Marcel Lang

Submission:

Journal:	Journal of Property Research
Submission date:	03/15/2018
Current Status:	Accepted (11/19/2018)

Conference presentations:

In 2018, this paper was presented at the 24th Annual Conference of the European Real Estate Society (ERES) in Delft, Netherlands as well as at the 34th Annual Conference of the American Real Estate Society (ARES) in Bonita Springs, US. Furthermore, the published version was presented at the 35th Annual Conference of the American Real Estate Society (ARES) in Paradise Valley, US.

Paper 3 | On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach

Authors:

Jochen Hausler, Johannes Braun, Wolfgang Schäfers

Submission:

Journal:	Journal of Real Estate Research
Submission date:	08/08/2019
Current Status:	Under review

Conference presentations:

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

Paper 4 | Artificial Intelligence, News Sentiment and Property Market Liquidity

Authors:

Johannes Braun, Jochen Hausler, Wolfgang Schäfers

Submission:

Journal:	Journal of Property Investment & Finance
Submission date:	08/01/2019
Current Status:	Under review

1.4 References

Amihud, Y. (2002), "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects", *Journal of Financial Markets*, Vol. 5 No. 1, pp. 31–56.

Baker, M. and Wurgler, J. (2007), "Investor Sentiment in the Stock Market", *Journal of Economic Perspectives*, Vol. 21 No. 2, pp. 129–152.

Black, F. (1986), "Noise", *The Journal of Finance*, Vol. 41 No. 3, pp. 528–543.

Clayton, J., Ling, D. C. and Naranjo, A. (2009), "Commercial Real Estate Valuation: Fundamentals Versus Investor Sentiment", *Journal of Real Estate Finance and Economics*, Vol. 38 No. 1, pp. 5–37.

Crosby, N., Lizieri, C. and McAllister, P. (2010), "Means, Motive and Opportunity? Disentangling Client Influence on Performance Measurement Appraisals", *Journal of Property Research*, Vol. 27 No. 2, pp. 181–201.

Das, P. K., Freybote, J. and Marcato, G. (2015), "An Investigation into Sentiment-Induced Institutional Trading Behavior and Asset Pricing in the REIT Market", *The Journal of Real Estate Finance and Economics*, Vol. 51 No. 2, pp. 160–189.

Freybote, J. (2016), "Real Estate Sentiment as Information for REIT Bond Pricing", *Journal of Property Research*, Vol. 33 No. 1, pp. 18–36.

Freybote, J. and Seagraves, P. A. (2017), "Heterogeneous Investor Sentiment and Institutional Real Estate Investments", *Real Estate Economics*, Vol. 45 No. 1, pp. 154–176.

Gallimore, P. and Gray, A. (2002), "The Role of Investor Sentiment in Property Investment Decisions", *Journal of Property Research*, Vol. 19 No. 2, pp. 111–120.

Kearney, C. and Liu, S. (2014), "Textual Sentiment in Finance: A Survey of Methods and Models", *International Review of Financial Analysis*, Vol. 33, pp. 171–185.

Kyle, A. S. (1985), “Continuous Auctions and Insider Trading”, *Econometrica*, Vol. 53 No. 6, p. 1315.

Li, F. (2010), “The Information Content of Forward-Looking Statements in Corporate Filings-A Naïve Bayesian Machine Learning Approach”, *Journal of Accounting Research*, Vol. 48 No. 5, pp. 1049–1102.

Lin, C. Y., Rahman, H. and Yung, K. (2009), “Investor Sentiment and REIT Returns”, *The Journal of Real Estate Finance and Economics*, Vol. 39 No. 4, pp. 450–471.

Liu, B. (2012), *Sentiment Analysis and Opinion Mining, Synthesis lectures on human language technologies*, #16, Morgan & Claypool Publishers, San Rafael.

Long, J. B. de, Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990), “Noise Trader Risk in Financial Markets”, *Journal of Political Economy*, Vol. 98 No. 4, pp. 703–738.

Ruscheinsky, J. R., Lang, M. and Schäfers, W. (2018), “Real Estate Media Sentiment through Textual Analysis”, *Journal of Property Investment & Finance*, Vol. 36 No. 5, pp. 410–428.

Shleifer, A. and Vishny, R. W. (1997), “The Limits of Arbitrage”, *The Journal of Finance*, Vol. 52 No. 1, pp. 35–55.

Soo, C. K. (2015), “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, 2015.

Tetlock, P. C. (2007), “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”, *The Journal of Finance*, Vol. 62 No. 3, pp. 1139–1168.

Walker, C. B. (2014), “Housing Booms and Media Coverage”, *Applied Economics*, Vol. 46 No. 32, pp. 3954–3967.

2 On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination

2.1 Abstract

We examine whether and the extent to which news-based sentiment, captured by textual analysis, can predict the performance of the private commercial real estate market in the United States. Our results show that sentiment reflected in news abstracts of the *Wall Street Journal* predicts returns of commercial real estate up to four quarters in advance. These findings are statistically significant and persist even when controlling for other related factors. This suggests that news-based sentiment can serve as an early market indicator. This paper is the first to examine the bi-directional relationship between sentiment, measured by textual analysis, and the performance of the private US commercial real estate market. The findings presented in this paper not only contribute to the academic literature, but also carry practical implications for real estate professionals.

Keywords: Sentiment, Textual Analysis, News analytics, Forecasting, Commercial Real Estate

2.2 Introduction

Empirical evidence by Baker and Wurgler (2007) as well as Seiler *et al.* (2012b) suggests that real estate investors bear not only economic, but also emotional factors in mind when making real estate investment decisions. A variety of other studies also show that economic fundamentals do not account for all observed price changes in commercial or residential real estate markets and much of the expectations about future cash flow are tied to information that is related to other factors (e.g. see Shiller, 2007; Lin *et al.*, 2009; Ling *et al.*, 2014). That said, only limited academic research directly investigates the role of sentiment in the commercial real estate (CRE) markets. In this study, we look to address this underexplored topic and examine the bi-directional relationship between sentiment and market returns of the private CRE in the US. We do so by analyzing real estate sentiment gathered from news data of a leading financial newspaper in the US, which is a new source of sentiment to be used for this type of analysis.

The private CRE suffers from several obvious market inefficiencies. Compared with the securitized real estate markets, the transparency of the private CRE market is limited, causing asymmetric information situations to be more frequent. All else equal, asymmetric information leads to high information and transaction costs, which results in a less efficient market, overall. The heterogeneity of properties provides additional challenges to real estate appraisers and lengthens investors' decision-making and transaction processes. Therefore, it is reasonable to expect that investors and appraisers in private CRE market are especially vulnerable to the influence of sentiments and opinions expressed in the news items that they consume. The tendency of the private CRE to adjust slower to new information and its vulnerability to non-economic fundamentals makes it particularly worth examining under the light of textual sentiment analysis.

In this study we gather more than 35,000 real-estate related news articles from *The Wall Street Journal (WSJ),* spanning the 2001 through 2016 time period, and analyze them in order to detect real-estate related sentiment. Specifically, a dictionary-based textual analysis approach is used to quantify the level of optimism and pessimism expressed through the abstracts of these articles. The intertemporal links between this sentiment and the private CRE market over the 16-year sample period are then examined to determine whether media-expressed real estate sentiment can help predict private CRE returns.

Our findings indeed suggest that sentiment reflected in news articles can help predict returns on the private CRE market in the US even after controlling for other macroeconomic factors. On average, our measure for media-expressed sentiment leads total returns on private CRE properties up to four quarters. Additionally, we do not find evidence for a feedback loop, where information on the performance of private CRE is reflected in future media-expressed sentiment although this could be expected.

Following prospect theory (see e.g. Kahneman and Tversky, 1979; Tversky and Kahneman, 1991), which advocates the maximization of an S-shaped value function by market participants and therefore loss aversion as a stable preference[1] – we further investigate the relevance of text-based sentiment measures during decelerating and accelerating market phases by splitting the sample accordingly. The results indeed show a higher relevance of sentiment indicators when markets are slowing down, which is consistent with previous findings in literature.

This study contributes to the existing literature by being the first to employ a real estate specific word dictionary to construct a real estate sentiment measure and determine whether and the extent to which such measure can help predict private CRE returns. More broadly, the results reported in this paper can be generalized to other less efficient investment asset classes.

The rest of this paper is organized as follows. In Section 2.3 we discuss the importance of sentiment in CRE markets and review relevant literature on investors' sentiment and textual analysis in the realm of real estate research. Section 2.4 presents the data set employed in this paper as well as a description of the sentiment-extraction procedure. In Section 2.5 we detail the methodology used for the analysis and present the hypotheses. Sections 2.6 and 2.7 reports the results and assess their robustness, while Section 2.8 concludes and discusses the implications of the findings.

[1] Bokhari and Geltner (2011) provide an excellent discussion of prospect theory, its three essential features – (1) evaluation of gains and losses relative to a reference point, (2) a steeper value function for losses than for equal-size gains and (3) a diminishing marginal value of gains/losses with size – as well as of the application of the theory in empirical studies when examining loss aversion and anchoring in commercial real estate pricing.

2.3 Literature Review

This study relates to two separate streams of literature. The first stream is the role of investors' sentiment with respect to the commercial real estate markets and its performance. The second stream refers to the textual analysis methodology used in this paper and the most recent developments in text-based sentiment measures in the realm of real estate.

2.3.1 Investors' Sentiment and Commercial Real Estate

Investors' sentiment is often measured directly or indirectly using two types of proxies. The most common direct measure approach is survey-based, such as the *Real Estate Research Corporation* sentiment measure that is employed in a few recent studies (Clayton *et al.*, 2009; Das *et al.*, 2015b; Freybote, 2016). While claiming to capture investors' sentiment directly, survey-based indicators, by their very nature, are associated with several material disadvantages. The surveys are not only costly and time consuming, but are also subject to the possibility that the answers provided by the respondents do not reflect their true sentiment. This might be due to the fact that respondents are not incentivized to take the surveys seriously or intentionally do not provide accurate and honest answers.

Indirect sentiment measures do not usually suffer from the disadvantages associated with the direct measures, because they are proxied by the actual behavior of market participants, which is fundamentally incentivized. These measures include, for example, closed-end fund discounts (Barkham and Ward, 1999; Clayton and MacKinnon, 2003; Lin *et al.*, 2009), buy-sell-imbalances (Freybote and Seagraves, 2017), mortgage fund flows (Clayton *et al.*, 2009; Ling *et al.*, 2014), search engine volumes and trends (Beracha and Wintoki, 2013; Das *et al.*, 2015a).

While many studies have examined the role of sentiment with relation to the residential real estate market, only a few studies have sought to investigate how investors' sentiment is related to the performance of private CRE. At least five recent studies that identify the relationship between sentiment and CRE performance in the US are closely related to this study. Clayton *et al.* (2009) analyze the impact of fundamentals and their sentiment index – constructed from sentiment-related proxies – on CRE values over the 1997-2007 period. Their results suggest that investors' sentiment does play a role in CRE pricing at the national as well as MSA-level and is robust to relevant macroeconomic factors. Ling *et al.* (2009) investigate the role of capital flows and turnover rates on returns of the UK private CRE market in the United Kingdom. Using a panel VAR approach, they do not find evidence for "price pressure" effects on capital flows, but for an information effect on turnover rates. Although not directly facilitating sentiment measures, the examined causal relationships (return chasing, joint dependency and information cascades) can be interpreted as expressions of investor sentiment, making the study worthwhile in a real estate sentiment context. Similarly, Ling *et al.* (2014) examined the relationship between investor sentiment – measured via direct and indirect real estate sentiment measures – and private as well as public CRE market returns over the 1992-2009 period. Using VAR models, the authors provide evidence for a positive relation between investor sentiment and private market performance in subsequent quarters. However, the relationship between investor sentiment and public real estate market returns in subsequent periods was negative. The authors support their findings with the argument that, in the short term, sentiment drives prices away from fundamentals, i.e. causes sentiment-induced mispricing. Furthermore, assessing various survey-based sentiment measures, their study concludes that real-estate-specific sentiment measures are of high importance, when quantifying the influence of sentiment on real estate. Another related study is by Tsolacos *et al.* (2014). Their paper deploys a probit and Markov-switching model to predict rental growth in CRE and apartment rent series in the US. The authors illustrate the prediction power of several sentiment-based leading indicators on commercial rent price movements. Finally, Marcato and Nanda (2016) assess whether survey-based sentiment indices help predict changes in quarterly US commercial and residential real estate returns. Using a VAR approach, their findings suggest significant effects of sentiment on the residential, but not the CRE, market over the period 1988-2010. Moreover, their results reveal that real estate specific sentiment indicators are more suited in explaining real estate markets than general business indicators.

Each of the above-mentioned studies contributes to our knowledge on investors' sentiment and CRE performance, but is also associated with its respective drawbacks. Specifically, these studies ignore the impact of other unperceived, but valuable, factors on investors' decision-making processes. For example, Price *et al.* (2017) show that executive emotions during earnings conference calls are positively related to investors' initial reactions. Analyzing the vocal cues of managers with a voice analysis software revealed that investors do indeed react to this emotionally charged information. Similarly, professional news outlets publish daily thousands of news articles on the real estate market. These publications range from reports and opinions to views and perspectives and are likely to, consciously or unconsciously, influence investors' action and, by extension, CRE performance.

In this study, we exploit this valuable source of information by applying textual analysis to published real estate news articles. This approach has already been applied in mainstream finance, but should be even more relevant to the private CRE market, which is arguably less efficient compared with the public market for common stocks. Section 2.3.2 provides a concise overview of related research using textual analysis conducted to date.

2.3.2 Sentiment Measure Using Textual Analysis

In the finance literature, Tetlock (2007) is regarded as one of the pioneers in applying textual analysis in order to capture market sentiment. Tetlock employs a sentiment dictionary on the "Abreast of the Market" column of the *Wall Street Journal* and successfully shows a relationship between pessimism reflected in news items and price changes of the Dow Jones Industrial Average Index, as well as its trading volume. A few other studies followed with a similar methodology and facilitated dictionary-based approaches using sentiment-annotated word-lists in order to extract sentiment from news items (see, for example, Henry and Leone, 2016; Feldman *et al.*, 2010; Davis *et al.*, 2012). While Tetlock (2007) use the Harvard GI word list from the field of psychology, Loughran and McDonald (2011) set a further cornerstone by highlighting the importance of a domain-specific dictionary. The authors develop a dictionary relevant to financial text corpora, which Boudoukh *et al.* (2013) and Heston and Sinha (2016) successfully utilize in their research.

Recently, a few studies examine the impact of sentiment extracted from text corpora in the context of real estate. Soo (2015) investigates the sentiment expressed in 37,500 local housing news articles of 34 US cities in order to predict future house prices. The author finds that the measured sentiment has predictability power and leads housing price movements by more than two years. Walker (2014) illustrates a material positive relationship between newspaper articles in the *Financial Times* and returns of listed companies engaged in the UK housing market. In accordance with his earlier findings, Walker (2016) subsequently analyzes the private housing market in the UK, and ascertains that news media granger-caused real house price changes from 1993 to 2008.

This paper aims to fill a gap in the literature and examines the relationship between textual based sentiment and the performance of private CRE in the US rather than the housing market or foreign publicly traded real estate firms. Investigating sentiment in the context of the private CRE market, which is expected to be less efficient than the public market, provides a meaningful contribution to the literature and the results can be generalized to other less efficient markets.

2.4 Data

The dataset complied for the empirical analysis conducted in this study is based on three main sources: (1) a news media corpus to extract sentiment, (2) a measure of private US commercial real estate market performance and (3) general macroeconomic factors.

2.4.1 News Data

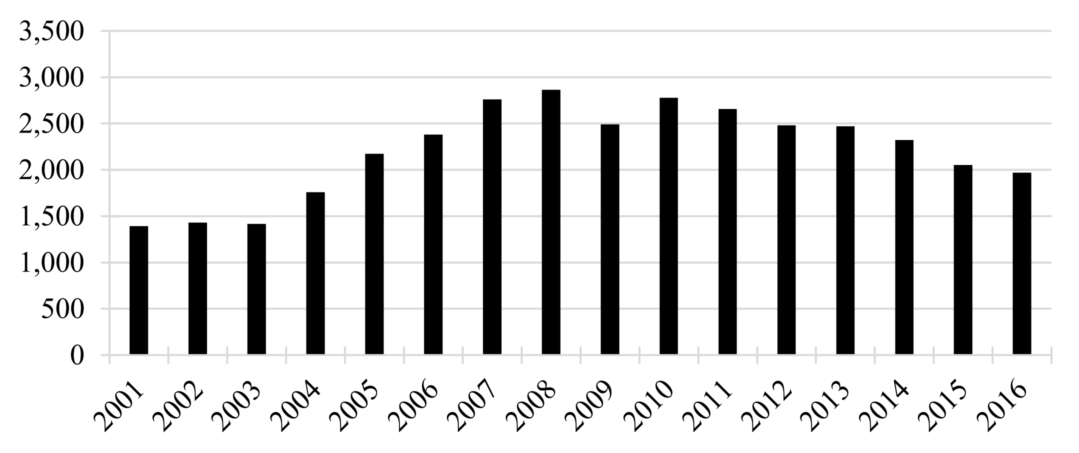
Our news data source used for the analysis in this study is *The Wall Street Journal (WSJ)*. Founded in New York City in 1889, the *WSJ* is nowadays the largest newspaper in the US in terms of its daily circulation.[2] Nationally and internationally, the *WSJ* is considered by many as one of the leading sources of business and financial news and it includes a dedicated real estate section. The *WSJ* has a broad readership, ranging from retail to institutional investors as well as managers and real estate professionals. Given its corporate news, political and economic reporting as well as its financial and real estate market coverage, the *WSJ* is of great importance to the CRE market. Although Tetlock (2007) pioneered textual analysis based on the "Abreast of the Market" column of the *WSJ* in mainstream finance, the real estate literature still lacks an attempt to capture its sentiment.

[2] According to the *WSJ*'s June 2017 10-K Filing, it had a paid circulation of more than 2.2 million subscribers whereof more than 50% were digital subscriptions.

Considering the aforementioned aspects, we use news items from the *WSJ* to capture and quantify media-expressed sentiment concerning the private CRE market. Specifically, via *ProQuest* (www.proquest.com), we accessed *WSJ*'s digital archive of the period that spans January 2001 until December 2016 and retrieved articles containing either the keywords "real estate" or "REIT". This 16-year period is a representative and worthwhile time span as it contains the real estate boom market phase until 2007, the real estate bust and the global financial crisis (GFC) from 2007 to 2010, as well as the subsequent recovery market phase from 2011. We further limited the data queries geographically to the US and to news reported in the English language. Over the sample period, the *WSJ* published 35,398 unique real estate-related news, which – on average – translates to more than 550 news items per calendar quarter. It is worth mentioning, that we exclusively analyze the abstracts of the newspaper articles. We assume, that these abstracts contain all relevant information of the articles themselves, but exclude noise in terms of irrelevant words and additional information, which are not necessary in order to capture the "tone" or sentiment expressed.

Figure 2.1 shows the annual number of real estate-related news published by the *WSJ* over the sample's 16-year time period that spans 2001 to 2016. The graph depicts a significant increase in news coverage during the boom market phase starting with around 1,759 news in 2004 and ending with 2,762 articles in 2007. During the real estate bust period, the number of articles reached its peak with 2,863 news items released in 2008 and then gradually declined. 1,970 news items were included in 2016, which is slightly above the average number of articles during the pre-crisis period. This general increase in real estate news coverage may suggests an overall rise of attention for real estate as an asset class.

Figure 2.1: WSJ Real Estate News Coverage, 2001 – 2016



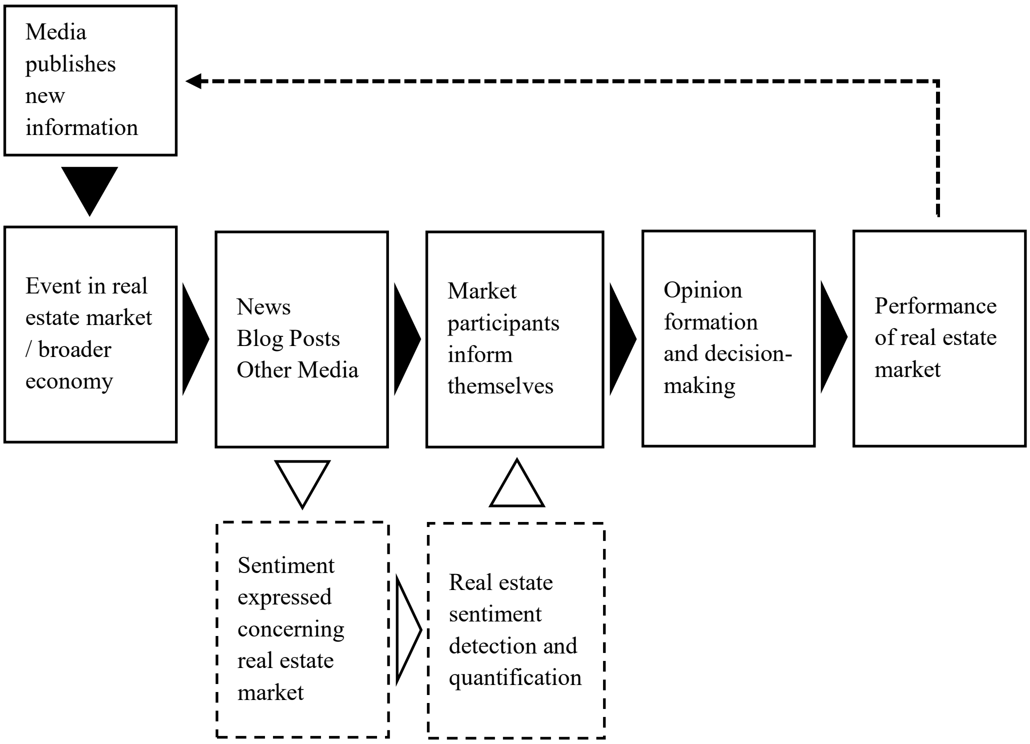
Notes: Figure 2.1 plots the sample distribution with respect to the number of real-estate related news published by the *WSJ* per annum. All *WSJ* news were retrieved using ProQuest; all articles contain either the keyword "real estate" or "REIT". The sample period is 2001:Q1 to 2016:Q4

2.4.2 Sentiment Measure Construction

To illustrate the theoretical background of our sentiment extraction procedure, we refer to the News-Impact-Model (Figure 2.2) of Lang (2018, p. 2). Accordingly, different news outlets report on certain events in the broader economy or on real estate markets. We assume that when real estate investors and appraisers inform themselves, the news to which they are exposed to – consciously or unconsciously – affect their opinion-formation and decision-making processes. Hence, the news-based sentiment is assumed to affect their individual sentiment. Thus, market participants' actions in aggregation are based on certain expectations and are in turn able to influence the performance of the commercial real estate markets. Consequently, from a total return perspective, we expect news-based sentiment to affect the appreciation returns since real estate investors and appraisers adjust their willingness to pay and valuations, respectively, upon their expectations and beliefs of future market developments.[3] Ultimately, the resulting events and market performance might be newsworthy and reported on again. Accordingly, this research paper yields to detect and quantify the sentiment expressed in real estate news abstracts published by the *WSJ*.

[3] With respect to the income component of total returns, rent prices for CRE are typically contractual and expected to be less dynamic. Therefore, short-term income returns are rather unlikely to be impacted by news-based sentiment. However, this is further examined in the robustness section.

Figure 2.2: News-Impact-Model



Based on this idea, we use a dictionary-based sentiment classifier to extract the sentiment from news abstracts, which could influence market participants during their opinion-formation and decision-making processes. Hence, we employ a pre-defined sentiment dictionary i.e. a word list annotated by sentiments such as positive or negative to every single news item and aggregate the sentiment of the identified words. This allows us to measure the overall "tone" of the abstracts.

Following Loughran and McDonald (2011) we apply a domain-specific dictionary by extending their pure finance dictionary to real estate specific terms. Our word list contains 408 positive and 2,455 negative terms. To ease the process of sentiment extraction, words in the dictionary and in the news abstracts are preprocessed, i.e. converted in well-defined sequences of linguistically meaningful units following Uysal and Gunal (2014).[4]

For every abstract, we count positive and negative words. Hereby, each positive word is counted as a "+1" and each negative word as a "-1". Because the sentiment

[4] For more details on this process please see sections 2.9.1 and 2.9.2.

dictionary does not consist of an equal number of positive and negative words, positive scores are multiplied by the inverse of the total number of positive terms divided by the total number of negative words in the dictionary. This calibrates the likelihood of that positive and negative words have similar impact on total count. This procedure allows us to calculate the overall sentiment score of each abstract by addition of the numeric values from the positive and negative words. An abstract can be viewed as positive, if the sentiment score is greater than 0, negative if the sentiment score is smaller than 0, and neutral if it is 0.

Subsequently, all positive, negative and neutral abstracts are added up for a defined period in order to arrive with a total periodic score of the positive, negative and neutral categories, respectively. This value is calculated on an absolute or weighted basis. The absolute basis only considers the raw number of positive and negative news items. For example, if there are 56 positive abstracts published during a given period, the positive periodic score for that period would simply be 56. On the other hand, the weighted approach uses the actual sentiment scores assigned to every abstract. This means that two negative abstracts with a score of “-5” and “-2” are added up for a score of “-7”.

This periodical aggregation of sentiment scores further allows us to generate a final combined sentiment measure for each period by calculating a so-called Positive-Negative-Ratio (PNR). This ratio expresses the amount of positive sentiment relative to total amount of negative sentiment. A higher ratio suggests a more positive sentiment and a lower ratio suggests a more negative sentiment with respect to the commercial real estate market. More formally, the PNR is calculated as the following:

$$PNR_t = \frac{\sum_1^I Positive\ Sentiment\ Score_{i,t}}{\left|\sum_1^J Negative\ Sentiment\ Score_{i,t}\right|}, \tag{2.1}$$

where i and j represent the abstracts with positive and negative scores, respectively, and t is the time period during which the published abstracts are accounted for. Because the category scores are measured either on an absolute or weighted basis so are the PNR-ratios.[5] For further details and a numeric example of the overall PNR calculation

[5] While we acknowledge that there is heterogeneity across locations, especially in the United States, we assume that institutional investors and decision-makers act from a portfolio-perspective. Thus, we deem one overall-market sentiment measures to be appropriate since we assess its relationship with overall CRE market performance.

process, please refer to the "Quantifying News-Based Sentiment" of the appendix (Section 2.9.3).

While some scholars such as Ling *et al.* (2014) or Marcato and Nanda (2016) orthogonalize their sentiment proxies against a set of macroeconomic controls, others such as Freybote and Seagraves (2017) and Das *et al.* (2015b) do not. As dictionary-based approaches solely rely on opinionated word lists to proxy sentiment, one could argue that orthogonalizing is not as important as it would be for survey-based measures. However, it can also be stated that every sentiment indicator as a proxy of market perception should most likely be influenced by facts and sentiment at the same time and that this should be accounted for. Therefore, the fact that we do not orthogonalize our sentiment measure can be interpreted as a possible shortcoming of this study and should be a subject of future research.

2.4.3 Other Data

The data on the performance of the private CRE market in the US used in this paper is the NPI series extracted from the *National Council of Real Estate Investment Fiduciaries (NCREIF)*. The NPI is an unleveraged total return index for private CRE properties held by contributing institutional investors. Published with quarterly frequency since 1977, the NPI is an appraisal-based index where each property's performance is weighted by its market value. Though it is available for different property types, we use the national composite NPI to measure total returns of the private US CRE market, incorporating the major property types i.e. apartments, hotels, industrial, office and retail. For our analysis, we are using total returns as well as capital-appreciation returns only as we expect news-based sentiment to especially affect appreciation returns.

In order to control for economic factors that are likely to affect CRE returns, we follow Clayton *et al.* (2009) and Ling *et al.* (2014) and include in our dataset macroeconomic variables proven to affect CRE returns. These variables include: the term structure of interest rates (defined as the spread between the ten-year US Treasury Constant Maturity rate and the 3-Month Treasury Bill yield), the percentage change in the Consumer Price Index (CPI) and the spread between Baa- and Aaa-rated corporate bonds yields. We obtain these economic variables from the *Federal Reserve Bank of St. Louis* with quarterly frequency.

Table 2.1 provides descriptive statistics about the quarterly NPI total returns (*NPI*), quarterly NPI capital-appreciation returns (*NPI_CR*), absolute and weighted Positive-Negative-Ratios (*PNR_A* and *PNR_W*) and our macroeconomic control variables. For each variable, we report the mean, median, standard deviation (SD), minimum (Min) and maximum (Max). The average quarterly total returns of the private CRE during our sample period is 2.19% and ranges between -8.40% and 5.49%, given the high volatility during the boom and bust phases that are part of our sample period. Capital-appreciation returns are associated with lower quarterly values, ranging between -9.66% and 3.89% with a mean (median) of 0.65% (1.27%). The average *PNR_W* value (7.60) is more than three times of the *PNR_A* value (2.27), which depicts the importance of distinguishing between the two measures and sheds light on the strength of the respective sentiment. The average quarterly *INFLATION* during the sample period was 0.52%, while *TERM* and *SPREAD* float around 1.1% and 2.1%, on average.

Table 2.1: Descriptive Statistics

Statistic	Mean	Median	SD	Min	Max
NPI (%)	2.191	2.687	2.550	-8.399	5.490
NPI_CR (%)	0.648	1.268	2.545	-9.655	3.889
PNR_A	2.272	1.592	1.142	0.912	4.814
PNR_W	7.601	5.619	4.453	2.057	17.530
INFLATION (%)	0.518	0.584	1.021	-3.910	2.476
TERM (%)	2.103	2.240	1.021	-0.380	3.580
SPREAD (%)	1.104	0.975	0.453	0.550	3.380

Notes: Table 2.1 reports summary statistics of variables used in the analysis on a quarterly basis. *NPI* is the total return of the NPI and *NPI_CR* is the capital appreciation return. *PNR_A* and *PNR_W* are the absolute and weighted Positive-Negative-Ratio sentiment measures, respectively. *INFLATION* is the percentage change of the Consumer Price Index (CPI). *TERM* is the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields. *SPREAD* is the spread between Baa- and Aaa-rated corporate bonds yields. The sample period is 2001:Q1 to 2016:Q4.

2.5 Methodology and Hypothesis Formation

2.5.1 Visual and Correlation Analysis

As our preliminary visual analysis, we plot the media-expressed sentiment measures against the returns of the private CRE market. Specifically, we plot the deviation of the sentiment measure from its 1-year moving average relative to the quarterly CRE total returns. This type of plot would illustrate the general relationship between changes in market sentiment and CRE returns and highlights whether market sentiment leads or lags returns. Additionally, we calculate the respective correlations between our quarterly sentiment values and CRE quarterly returns.

2.5.2 Regression Analysis

We begin our empirical analysis by investigating the ability of real-estate related sentiment, expressed in the news, to predict total returns on the private CRE market in the US. To do so, we regress the NPI total return on the lagged absolute or weighted Positive-Negative-Ratios. By regressing CRE returns on our lagged media-expressed sentiment values, we test the hypothesis that market sentiment predicts future returns of the private CRE market.

Hypothesis 1: Real estate market sentiment predicts future returns of the private CRE market.

In addition to lagged media-expressed real estate sentiment, the regression specifications also control for other relevant macroeconomic variables proven to affect CRE market returns, (see e.g. Clayton *et al.*, 2009 and Ling *et al.*, 2014). Controlling for the term structure of interest rates is relevant because it is related to commercial real estate financing cost and expectations of future economic developments. Accounting for the percentage changes in the Consumer Price Index (CPI) is important because many commercial rental contracts are linked to inflation and therefore affect future returns. The spread between Baa- and Aaa-rated corporate bonds yields reflects the overall business conditions and general default risk in the economy. Finally, we include a dummy variable to control for any factors associated with the global financial crisis (GFC) from 2007:Q3 to 2008:Q4. Autocorrelation and heteroscedasticity issues are accounted for by using Newey and West (1987) robust standard errors.

Formally, we estimate the following equation:

$$\Delta NPI_t = c + \sum_{i=1}^{i=5} \propto_i (\Delta PNR_{t-i}) + \beta_1(\Delta INFL_t) + \beta_2(\Delta TERM_t) + \beta_3(\Delta SPREAD_t) + GFC_t + \varepsilon_t, \tag{2.2}$$

where NPI_t is the total return during quarter t; PNR_{t-i} is the Positive-Negative-Ratio to measure media-expressed sentiment with i quarterly lags; $INFL_t$ is the inflation rate, $TERM_t$ the interest term ensure structure and $SPREAD_t$ the spread between Baa- and Aaa-rated corporate bonds. GFC is a dummy variable to indicate the global financial crisis and ε_t represents the error term. Except of the crisis dummy, all variables are applied in first differences to stationarity.[6]

2.5.3 Vector Autoregressive Analysis

The multiple linear regression model described above estimates the value of the dependent variable (*NPI*) using several, supposedly independent, variables. However, it could be presumed that our media-expressed sentiment measures also contain information about past CRE market performance as indicated by the proposed News-Impact-Model of Section 2.4.2. Consequently, we examine the bi-directional relationship between media-expressed sentiment and the performance of the private US CRE market using a Vector Autoregressive (VAR) framework. According to this model, each variable is a linear function of lags of itself and lags of other variables. Hence, the VAR model allows us to estimate the intertemporal links between media-expressed sentiment and the private CRE market and address the potential endogeneity problem. Furthermore, the VAR model enables us to analyze whether the media-expressed sentiment predicts returns on private CRE, even when controlling for the lags of the NPI itself, which is shown to contain momentum (Beracha and Downs, 2015). Formally, the VAR model used in our analysis is specified as the following:

[6] For results of the augmented Dickey-Fuller tests for the presence of unit roots, i.e. non-stationarity please refer to section 2.9.4 in the appendix.

$$\Delta NPI_t = \alpha_{10} + \sum_{i=1}^{i=5} \beta_{1i}(\Delta NPI_{t-i}) + \sum_{i=1}^{i=5} \gamma_{1i}(\Delta PNR_{t-i})$$

$$+ \delta_1(\Delta Exog_t) + \varepsilon_{1t}$$

$$\tag{2.3}$$

$$\Delta PNR_t = \alpha_{20} + \sum_{i=1}^{i=5} \beta_{2i}(\Delta PNR_{t-i}) + \sum_{i=1}^{i=5} \gamma_{2i}(\Delta NPI_{t-i})$$

$$+\delta_2(\Delta Exog_t) + \varepsilon_{2t}.$$

The variables are as described above and defined in equation (2.2). Note that, for brevity, the control variables ($INFL_t$, $TERM_t$ and $SPREAD_t$) are summarized in $Exog_t$[7]. ε_{1t} and ε_{2t} are the error terms. The endogenous variables are quarterly NPI returns (NPI_{t-i}) and the media-expressed sentiment (*PNR_A* or *PNR_W*). We include lags up to t-5 based on the Akaike Information Criteria (AIC) for various choices of the lag length *p*. Applying the Augmented-Dickey-Fuller unit root test (see Dickey and Fuller, 1979; Said and Dickey, 1984) suggests using first differences of all variables to ensure stationarity.

2.5.4 Granger Causality Tests

We further examine the bi-directional relationship between media-expressed sentiment and CRE returns, by conducting pairwise Granger causality tests (Granger, 1969). This type of analysis helps us better understand the lead-lag relationships between sentiment in real estate related news and the private CRE market. We hypothesize that media-expressed sentiment drives total returns of the private CRE market, but not the other way around. We base our hypothesis on evidence from the literature that the CRE market is not fully efficient and is slow to react to new market information. Formally, our hypothesis is stated as the following:

Hypothesis 2: Media-expressed sentiment predicts future returns of private commercial real estate, but returns on private commercial real estate do not predict future media-expressed sentiment.

[7] Note that when the crisis dummy is included, results are similar with respect to the sign, size and the statistical significance of the *PNR_A* and *PNR_W* coefficients.

Formally, the model for testing Granger causality between real estate market sentiment and returns is defined as follows:

$$\Delta NPI_t = \propto_0 + \sum_{i=1}^{i=5} \beta_i(\Delta NPI_{t-i}) + \sum_{i=1}^{i=5} \gamma_i(\Delta PNR_{t-i}) \\ + \delta_1(Exog_t) + \varepsilon_t \tag{2.4}$$

$$\Delta PNR_t = \propto_0 + \sum_{i=1}^{i=5} \beta_i(\Delta PNR_{t-i}) + \sum_{i=1}^{i=5} \gamma_i(\Delta NPI_{t-i}) \\ + \delta_1(Exog_t) + \varepsilon_t. \tag{2.5}$$

The variables included in equations (2.4) and (2.5) are as described and defined earlier in the text. Consistent with our previous models, we conduct the tests for 1 to 5 lags and report the X^2 (Wald) statistics for the joint significance of each of the other lagged endogenous variables in both equations. The null hypothesis is that *ΔPNR* does not Granger-cause *ΔNPI* in equation (2.4) and vice versa in equation (2.5).

2.6 Results

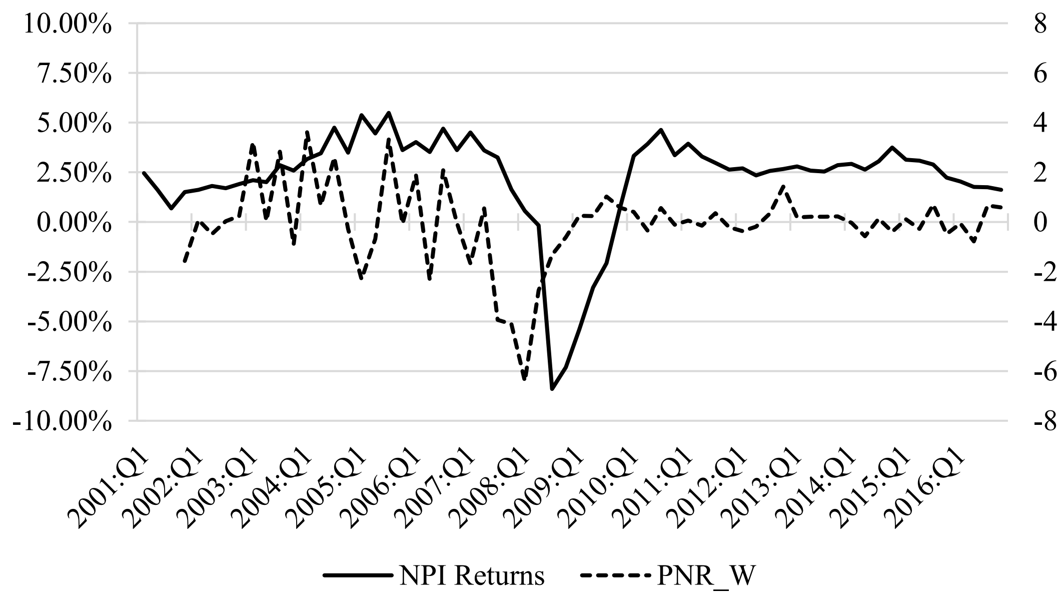
2.6.1 Visual and Correlation Results

Figure 2.3 provides visual illustration of the relationship between our weighted media-expressed sentiment measure (*PNR_W*) and the returns on the private CRE market.[8] A glance at the figure reveals that the two variables are correlated and that *PNR_W* seems to lead the private CRE market returns. For example, a substantial drop in sentiment occurred late 2007 and early 2008 and was followed by meaningful negative returns in the CRE market two quarters later. More specifically, the *PNR_W* drops from 0.54 in 2007:Q2 to -6.41 in 2008:Q1 and NPI total return bottomed in 2008:Q3 (-8.40%). Similarly, the sentiment seems to also be a leading indicator in periods of recovery and expansion. Following the drop in real estate market sentiment the measure improved from 2008:Q1 to 2009:Q3 while returns on the private CRE market gradually

[8] A figure using the absolute sentiment measure *PNR_A* was also conducted and appears qualitatively similar. However, because the absolute measure only accounts for general optimism and pessimism in news abstracts, but not the respective magnitude, up and downs are less pronounced. This figure is omitted from this version of the paper for brevity.

recovered from 2008:Q3 to 2010:Q3 with most of the recovery taking place before 2010:Q1. That said, the relationship in pre-crisis years is less clear as the sentiment measures show a high level of fluctuation relative to the performance of the CRE market.

Figure 2.3: Commercial Real Estate Returns and Media-Expressed Sentiment



Notes: Figure 2.3 plots levels of real estate media-expressed sentiment and the total returns on the CRE market. The media-expressed sentiment is quantified using the weighted Positive-Negative-Ratio (PNR_W) measure as described in the text. The sentiment is plotted based on the difference between current PNR (PNR_W_t) and the simple average of the weighted PNR of the last 4 quarters (PNR_{t-1} to PNR_{t-4}). The sample period is 2001:Q1 to 2016:Q4.

Table 2.2 presents the correlations between the level and change in media-expressed sentiment (*PNR_A* and *PNR_W*) and private CRE returns (*NPI* and *NPI_CR*). Returns based on the NPI are calculated on a quarterly basis. When the level of media-expressed sentiment is considered, the correlations between the *PNR_A* and *PNR_W* and the quarterly *NPI* are positive with the 1st quarterly lag and gradually dissipate through the 5th lag.[9] The correlation results for the capital appreciation returns behave in a very similar manner, which is expected given the fact that the correlation between the two return measures is 0.9936.[10] When the change in media-expressed sentiment

[9] Note that this behavior does not continue beyond the 5th lag.

[10] Note that income returns were quite stable over the sample period of 2004:Q1 to 2016:Q4. They deviated only between 1.14% and 2.14% with an average value of 1.56% and a standard deviation of only 0.29%.

is considered, the correlations of the *PNR_A* and *PNR_W* and the quarterly *NPI* are mostly positive in the early lags, but volatile. Overall, these results suggest that returns on CRE (total returns as well as capital appreciation returns) are correlated with the level and the change in the level of past media expressed real estate sentiment.

Table 2.2: Correlations: Sentiment and NPI Total Returns

	NPI Total Return (quarterly)		**NPI Capital Return (quarterly)**	
	Level	**Change in level**	**Level**	**Change in level**
PNR_A_{t-1}	0.41	0.02	0.41	0.03
PNR_A_{t-2}	0.39	0.49	0.39	0.49
PNR_A_{t-3}	0.26	0.06	0.26	0.08
PNR_A_{t-4}	0.11	0.14	0.11	0.15
PNR_A_{t-5}	-0.08	-0.27	-0.08	-0.26
PNR_W_{t-1}	0.45	-0.08	0.45	-0.07
PNR_W_{t-2}	0.44	0.43	0.44	0.44
PNR_W_{t-3}	0.32	-0.07	0.32	-0.06
PNR_W_{t-4}	0.21	0.47	0.21	0.47
PNR_W_{t-5}	-0.03	-0.45	-0.03	-0.44

Notes: Table 2.2 reports the correlations between the level and change in level for lags 1 to 5 of the absolute and weighted Positive-Negative-Ratio (*PNR_A* and *PNR_W*) and the quarterly CRE returns (*NPI* and *NPI_CR*). The sample period is 2001:Q1 to 2016:Q4.

2.6.2 Regression Analysis Results

Table 2.3 presents the results of several regressions specifications as per equation (2.2). Specifications (I) and (II) examine the ability of our absolute media-expressed sentiment measure to predict quarterly CRE returns with and without our macroeconomic control variables, respectively. When the control variables are excluded, the coefficient of the 2nd lag of the sentiment measure is positive and statistically significant at the 1% level. The coefficients then turn insignificant for the following lags until the 5th one, which has a negative sign and is significant at a 10% level. When the control variables are included, the 2nd and 5th sentiment measure lags still have the same sign and similar size but only the 2nd on remains significant, while the 5th one is no longer statistical significance at traditional threshold levels. The absolute sentiment measure leads total returns by two quarters corresponding to findings in Table 2.3.

Table 2.3: MLR Results: Quarterly NPI Returns and Media-Expressed Sentiment

	Regressand: NPI (quarterly)							
	(I)		**(II)**		**(III)**		**(IV)**	
	Absolute		Absolute		Weighted		Weighted	
PNR_{t-1}	0.0020		0.0016		0.0007		0.0015	
PNR_{t-2}	0.0079	***	0.0066	**	0.0067	***	0.0065	*
PNR_{t-3}	0.0030		0.0017		0.0066	**	0.0058	
PNR_{t-4}	0.0013		0.0003		0.0043	*	0.0043	*
PNR_{t-5}	-0.0036	*	-0.0035		-0.0060	**	-0.0056	**
INFLATION			0.2596				0.1208	
TERM			0.0398				0.0173	
SPREAD			0.3558				-0.1124	
GFC			-0.0077				0.0004	
INTERCEPT	0.0003		0.0010		0.0001		0.0001	
Adj. R^2	0.29		0.31		0.46		0.44	
AIC	-5.87		-5.83		-6.14		-6.04	

Notes: Table 2.3 reports the coefficients of the estimated MLR (multiple linear regression) models with quarterly NPI returns as the dependent variable on the lagged media-expressed sentiment (*PNR*) as well as macroeconomic control variables. The set of control variables in our regression are the CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*), the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*) and a dummy variable that captures the effect of the great financial crisis (*GFC*), which is set to 1 during the 2007:Q3 to 2008:Q4 time period and 0 otherwise. We use Newey and West (1987) standard errors that are robust to heteroscedasticity and autocorrelation. We transformed all variables to their first differences. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

Specifications (III) and (IV) repeat the analysis from specifications (I) and (II), but with our weighted rather than absolute media-expressed sentiment measure. Findings are similar to models (I) and (II) and even more pronounced, which is expected when considering that the weighted sentiment measure not only captures the raw existence of sentiment in abstracts, but also its magnitude in contrast to the *PNR_A* indicator. Moreover, the adjusted R^2 for specifications (III) and (IV) are materially larger than the R^2 in the specifications where the absolute measure is employed (44% and 46 % compared to about 30%). Except of the 3rd lag of model (IV), the 2nd through 5th lags are significant and the sign is positive for lag 2 through 4 and only the last lag (t-5) turns negative. This implies that when taking the "strength" of sentiment expressed in news abstracts into account, the relationship between sentiment and return is more pronounced. However, in terms of magnitude of specific lags, the results are quite similar. A change of ΔPNR_A_{t-2} (ΔPNR_W_{t-2}) by one standard deviation in model (II) and (IV) leads, ceteris paribus, to an increase of *ΔNPI* by 0.66 and 0.65 percentage points, respectively.[11]

The negative sign of the 5th lag in specifications (I) to (IV) may indicate a potential reversal or correction effect of the media-expressed sentiment. Other researchers such as Tetlock (2007) and Antweiler and Frank (2006) found similar evidence with respect to the general stock market. It is also important to note that the negative coefficient of the 5th lag does not eliminate the positive impact of *ΔPNR_A* or *ΔPNR_W* on *ΔNPI* over the previous four lags. When looking at impulse-response-functions, the influence of a one standard deviation innovation of *ΔPNR_A* or *ΔPNR_W* persists over time.[12] This is also in line with findings of Ling, Naranjo and Scheick (2014) with respect to the influence of investor sentiment on private real estate markets returns.

It is also worth mentioning, that all models show comparable dissipation of the size of the coefficient (with or without statistical significance) from the 2nd to 5th lag as was evident from the correlation analysis with the *NPI* in Table 2.2. Thus, for example, an increase of the 2nd lag of *ΔPNR_W* by one standard deviation leads, ceteris paribus, to a positive increase of 0.67 percentage points in the *NPI* in model (III) while the impact of the 3rd and 4th lag is smaller at 0.66 and 0.43 percentage points, respectively. These results suggest that real estate sentiment predicts future returns of private CRE market and therefore provide support to hypothesis 1.

2.6.3 Vector Autoregressive Analysis Results

Table 2.4 reports the VAR estimation outputs as per equation (2.3). Like in the previous table, columns (I) and (II) presents the estimation results using the absolute media-expressed sentiment measure and columns (III) and (IV) using the weighted measure. The purpose of this analysis is to examine the ability of media-expressed real

[11] Note that both sentiment measures (ΔPNR_A and ΔPNR_W) are scaled to unit variance.

[12] Impulse response figures are available upon request and omitted from this version of the paper for brevity.

estate sentiment to predict the returns of private CRE while controlling for possible momentum behavior embedded within CRE returns. The VAR framework also allows controlling for a possible feedback loop as previously stated by the News-Impact-Model of Section 2.4.3. Overall, the results presented in Table 2.4 are consistent with the results presented in the previous tables and provide support to hypothesis 1, showing that real estate sentiment helps predict the returns of the CRE market and that the results of our prior regression models hold within the VAR framework. Again, the first 4 lags are positive and the 2nd (and 4th with the weighted measure) lag is statistically significant; the coefficients dissipate from the 2nd to the 4th lag and turn negative for the last lag (t-5), which is only significant for the *PNR_W*. In terms of size, the coefficients of Table 2.3 and Table 2.4 are quite similar. Moreover, the results again suggest that our weighted sentiment measure is better suited, compared with the absolute sentiment measure, as a predictor. Aside from the statistical significance of the lagged coefficients, the adjusted R^2 and AIC values in these VAR specifications are materially higher with the weighted compared to when the absolute sentiment measure is used.

Table 2.4: VAR Results: Quarterly NPI Returns and Media-Expressed Sentiment

	Dependent variable: NPI (quarterly)							
	(I) PNR_A		**(II) PNR_A**		**(III) PNR_W**		**(IV) PNR_W**	
	w/o CV		*w/ CV*		*w/o CV*		*w/CV*	
NPI_{t-1}	-0.0771		-0.1119		0.1283		0.2105	
NPI_{t-2}	0.0700		0.1607		-0.0431		0.0010	
NPI_{t-3}	0.0114		0.0167		0.0740		0.0356	
NPI_{t-4}	-0.0227		-0.0855		-0.0716		-0.1416	
NPI_{t-5}	-0.2089	***	-0.1832	**	-0.1899	**	-0.1816	*
PNR_{t-1}	0.0019		0.0025		0.0005		0.0013	
PNR_{t-2}	0.0078	**	0.0073	**	0.0061	**	0.0060	**
PNR_{t-3}	0.0039		0.0033		0.0060		0.0056	
PNR_{t-4}	0.0014		0.0005		0.0050	*	0.0053	*
PNR_{t-5}	-0.0031		-0.0027		-0.0063	*	-0.0059	
INTERCEPT	0.0003		0.0003		0.0001		0.0002	

(Table continues on the following page.)

Table 2.4: VAR Results: Quarterly NPI Returns and Media-Expressed Sentiment (continued)

Adj. R^2	0.28	0.32	0.47	0.46
AIC	-5.78	-5.80	-6.09	-6.03

Notes: Table 2.4 reports the estimated coefficients from the VAR (vector autoregression) models with quarterly NPI total returns (*NPI*) and Positive-Negative-Ratio (*PNR*) as endogenous variables. The lag length of the VAR is based on the Hannan-Quinn criterion. The set of the macroeconomic control variables (*CV*) in our regression are the CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*), the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*). We use Newey and West (1987) standard errors that are robust to heteroscedasticity and autocorrelation. We transformed all variables to their first differences. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

2.6.4 Granger Causality Test Results

Table 2.5 presents the results of our Granger causality tests conducted in order to examine the causal relationship of the PNR measures and NPI returns as proposed by the dynamic nature of news-impact process. Note that the null hypothesis assumes media-expressed sentiment does not cause CRE returns and vice versa. Specifically, columns (I) and (II) test whether the coefficients of the lagged media-expressed sentiment variables are equal to zero. Conversely, columns (III) and (IV) are based on the null hypothesis that the coefficients of the lagged CRE return variables do not influence changes in future sentiment measures.

Table 2.5: Granger Causality Test Results

	H_0: Media-expressed sentiment does not cause NPI		**H_0: NPI does not cause Media-expressed sentiment**	
	(I)	**(II)**	**(III)**	**(IV)**
	Absolute PNR	Weighted PNR	Absolute PNR	Weighted PNR
X^2 *(w/o CV)*	23.67 ***	49.03 ***	2.17	1.82
X^2 *(w CV)*	20.24 ***	37.16 ***	5.48	1.68

Notes: Table 2.5 reports the Granger causality results of the estimated VAR (vector autoregression) models of specifications (I) to (IV) of Table 2.4. Granger causality results test the joint significance of all lags for a given variable. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

The results presented in columns (I) and (II) reject the null hypothesis and suggest that there is a statistically significant causality running from media-expressed real estate sentiment (*PNR*) to private CRE returns (*NPI*). This relationship is statistically significant at a 1% level and holds when control variables are included, regardless of the sentiment measure (*PNR_A* or *PNR_W*). Column (III) and (IV) suggest that the CRE market does not drive media-expressed sentiment when proxied using the PNR sentiment measures. Hence, it can be argued that – based on findings in our sample – there is no feedback loop between media-expressed sentiment and CRE returns. Overall, real estate related news contains new information and sentiment that affects CRE returns. However, a reverse causation could not be shown. Note that this does not mean, that past return movements are not reflected in news at all. Nevertheless, these results imply that these movements do not influence opinion building and decision-making. It is possible that other aspects of news are of higher importance to real estate market participants when forming their expectations about the future.

2.7 Robustness Checks

When we introduce the News-Impact-Model in Section 2.4.3, we state that we expect our news-based sentiment measures to especially affect the capital appreciation component of the NPI. This is due to the fact that income returns are more stable, in general. For example, during the 16-year period examined in this paper the standard deviation of the income return was only 0.29%. Table 2.6 reports the results of the regressions from Table 2.3 when using the *NPI_CR* instead of the *NPI* as measure of market returns as per equation (2.2). When comparing the results of Table 2.3 with Table 2.6 one can see that the main findings persist, but the results deviate slightly in terms of size and significance of the *PNR*. While only the 2nd lag is significant in model (I) and (II), the weighted *PNR* explains future returns up to four quarters in the future. Adjusted R^2 are also materially larger for the weighted *PNR* measure and models (I), (III) and (IV) show a significant reversion in the 5th lag. For the sake of brevity and due to the high level of similarity we refrain from showing further corresponding VAR results within the paper, but include these results in the appendix (Section 0). Nevertheless, we want to state that both tables provide convincing evidence that, as expected, text-based sentiment indicators indeed affect returns via capital appreciation.

Table 2.6: MLR Results: Quarterly Appreciation Returns and Media-Expressed Sentiment

	Regressand: NPI capital return (quarterly)			
	(I)	**(II)**	**(III)**	**(IV)**
	Absolute	Absolute	Weighted	Weighted
PNR_{t-1}	0.0020	0.0016	0.0008	0.0014
PNR_{t-2}	0.0079 ***	0.0066 **	0.0067 ***	0.0064 *
PNR_{t-3}	0.0032	0.0019	0.0067 **	0.0059
PNR_{t-4}	0.0015	0.0005	0.0045 *	0.0045 *
PNR_{t-5}	-0.0034 *	-0.0034	-0.0058 *	-0.0053 *
INFLATION		0.2491		0.1115
TERM		0.0304		0.0116
SPREAD		0.3571		-0.1046
GFC		-0.0076		0.0002
INTERCEPT	0.0004	0.0011	0.0003	0.0002
Adj. R^2	0.30	0.31	0.46	0.44
AIC	-5.90	-5.85	-6.17	-6.06

Notes**:** Table 2.6 reports the coefficients of the estimated MLR (multiple linear regression) models with quarterly capital returns of the NPI as the dependent variable on the lagged media-expressed sentiment (*PNR*) as well as macroeconomic control variables. The set of control variables in our regression are the CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*), the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*) and a dummy variable that captures the effect of the great financial crisis (*GFC*), which is set to 1 during the 2007:Q3 to 2008:Q4 time period and 0 otherwise. We use Newey and West (1987) standard errors that are robust to heteroscedasticity and autocorrelation. We transformed all variables to their first differences. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

The prospect theory presented by Kahneman and Tversky (1979) posits that individuals are more prone to avoid losses than to achieve economic gains. In accordance, empirical evidence in real estate research suggests that investors yield to avoid experiencing regret by deviating from rational behavior. On the residential real estate side, for example, Seiler *et al.* (2012a) found that the willingness of investors to sell a residential property increases most when their investment breaks even. Following a similar logic, it can also be argued that private CRE markets are particularly prone to media-expressed sentiment during downward markets and recessions.

To investigate whether the sentiment-based predictability of NPI total returns is asymmetric, i.e. higher predictability power during periods of slower market growth, we run two separate regressions (based on equation 2.2) using the portion of the sample when the market is accelerating or decelerating.[13] Hence, the samples include only the quarters of growing and shrinking total NPI returns, respectively. Table 2.7 presents the results of the *up-market* vs. *down-market* trend analysis in panel A and B, respectively.

Table 2.7: MLR Results: Sentiment in Accelerating vs. Slowing-Down Markets

Panel A: Up-market trend

	Regressand: NPI (quarterly)			
	(I)	**(II)**	**(III)**	**(IV)**
	Absolute	Absolute	Weighted	Weighted
PNR_{t-1}	0.0024	0.0048	0.0030	0.0066 **
PNR_{t-2}	0.0028 **	0.0034 **	0.0044 ***	0.0073 ***
PNR_{t-3}	-0.0009	-0.0021	0.0001	0.0008
PNR_{t-4}	-0.0006	-0.0005	-0.0008	0.0004
PNR_{t-5}	-0.0029	-0.0005	-0.0027	-0.0008
INFLATION		-0.0787		-0.1004
TERM		-0.2602		0.0134
SPREAD		-1.2034		-0.9795 *
GFC		0.0111		0.0165 **
INTERCEPT	0.0045 **	0.0046 *	0.0043 *	0.0038
Adj. R^2	0.17	0.13	0.15	0.16
AIC	-6.66	-6.52	-6.64	-6.56

(Table continues on the following page.)

The results reported in Panel A indicate that our absolute real estate sentiment measure (columns I and II) has lower predictive power with respect to future CRE performance in up-market phases. Compared to Table 2.3, the 2nd lag is still significant but at a weaker level and coefficients are smaller in size. It carries the anticipated sign with a statistical significance at the 5% level. While there is still a correction effect in sign of coefficients for the following lags, the results are not statistically significant. When

[13] Because reverse-causation was not found in our VAR models and results are similar in Table 2.3 and Table 2.4, we stick to MLR regression for robustness checks. The regression models with capital appreciation returns are also available from the authors upon request.

our weighted real estate sentiment measure is considered (columns III and IV) the overall predictability power improves somewhat as coefficients of the 2nd lag are now significant at a 1% level and of greater magnitude. However, for both measures – the *PNR_A* and the *PNR_W* – only the 2nd lag is still significant, while the later ones are not. Additionally, overall, adjusted R²s are roughly half of size of the ones reported in Table 2.3, supporting a lower predictive or explanatory power of our sentiment measures when the CRE market accelerates.

Table 2.7: MLR Results: Sentiment in Accelerating vs. Slowing-Down Markets (continued)

Panel B: Down-market trend

	Regressand: NPI (quarterly)			
	(V)	**(VI)**	**(VII)**	**(VIII)**
	Absolute	Absolute	Weighted	Weighted
PNR_{t-1}	0.0063 **	0.0049 **	0.0024	0.0008
PNR_{t-2}	0.0135 ***	0.0100 ***	0.0155 ***	0.0116 ***
PNR_{t-3}	0.0080 ***	0.0048 *	0.0114 ***	0.0073 **
PNR_{t-4}	0.0044	0.0044	0.0088 ***	0.0072 ***
PNR_{t-5}	0.0001	0.0017	-0.0006	0.0010
INFLATION		0.2952 *		0.1354
TERM		-0.3441		-0.0480
SPREAD		-1.2979 *		-0.9772
GFC		-0.0152		-0.0183 *
INTERCEPT	-0.0051 *	-0.0032	-0.0024	-0.0013
Adj. R^2	0.59	0.69	0.71	0.75
AIC	-5.98	-6.18	-6.34	-6.39

Notes: Table 2.7 reports the coefficients of the estimated MLR (multiple linear regression) models with quarterly NPI returns as the dependent variable on media-expressed sentiment (*PNR*) and a set of macroeconomics control variables: CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*), the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*) and a dummy variable that captures the effect of the great financial crisis (*GFC*), which is set to 1 during the 2007:Q4 to 2009:Q1 time period and 0 otherwise. Panel A presents the results using the up-market trend portion of the sample and Panel B for the down-market trend portion of the sample. We use Newey and West (1987) standard errors that are robust to heteroscedasticity and autocorrelation. We transformed all variables to their first differences. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

In comparison, the results reported for the down-market trend portion of the sample suggest that our sentiment real estate measure has high level of predictability during these periods. The magnitude of the lagged coefficients is larger and carry higher statistical significance compared with the up-market specifications. Similarly, the adjusted R^2 for down-market specifications are also materially higher compared with the up-market specifications. For example, a one standard deviation increase of ΔPNR_A_{t-2} in model (V) leads, ceteris paribus, to a 1.35 percentage point increase of ΔNPI. The respective impact is 0.79 (0.28) percentage points in model (I) of Table 2.3 (Table 2.7, Panel A). In all specifications (columns V though VIII) the coefficients of the 1st, through the 4th quarter lag of the real estate sentiment measure (absolute or weighted) are positive and statistically significant at the 1% or 5% level (with the exceptions of the 4th lag in columns V and VI). These findings are consistent with our expectations and suggest that commercial real estate prices are better predicted by sentiment during periods of decelerating markets compared with up-market trends. Furthermore, our results are also consistent with the findings of Beracha and Wintoki (2013) for the role of sentiment in the residential properties during up versus down markets.

2.8 Conclusion

The existing literature provides evidence that sentiment plays an important role in pricing different asset classes, independent of fundamentals. In this paper, we employ a real estate specific sentiment measure that is based on news articles in order to determine the extent to which media-expressed sentiment can help predict private CRE returns in the US. The results of our analysis show that media-expressed sentiment predicts returns of commercial real estate up to four quarters in advance. These results are robust when macroeconomic factors are accounted for. When analyzing the bi-directional relationship between media-expressed sentiment and CRE returns, our results show clear evidence that information is flowing from the media-expressed sentiment to the private CRE market, but not vice versa. Moreover, our results show that the predictability of the media-expressed sentiment is especially pronounced for down-market trends, rather than up-market trends, which is also consistent with the related literature.

Our findings contribute to the literature on market sentiment and private CRE performance and should be of interest to academics as well as real estate professionals. Analyzing textual documents about real estate markets can provide investors with valuable information about the future performance of the CRE market. Specifically, the results of this paper highlight the fact that news serve as a leading indicator and can help real estate investors predict price movements in the commercial real estate market up to 4 quarters in advance.

This study also set the foundation for future research on advanced methods of textual analysis and machine-learning algorithms with respect to investments, in general, and CRE, in particular. The results presented in this paper show that the text-based sentiment indicators are valuable to commercial real estate investors in the US and opens the door to research on other asset classes and/or locations. More generally, applications of textual analysis and machine-learning algorithms with respect to investment is still in its infancy. Therefore, we expect many future studies to build and improve upon our methodologies and results.

2.9 Appendix

2.9.1 Creation of a Real-Estate Specific Dictionary

While different sentiment related word lists and dictionaries are available, this paper follows Loughran and McDonald (2011) who found that sentiment dictionaries should be domain-specific in order to classify text corpora adequately. Thus, as a starting point, we deploy their finance dictionary based on the assumption that the terminology in the realm of real estate should be linked to vocabulary used in finance. Albeit the lexicon distinguishes between the sentiment categories of positive, negative, uncertain, litigious, constraining, superfluous, interesting and modal terms, only the first two categories of positive (354 words) and negative (2,355 words) are used. In the second step, this basic finance dictionary is adapted to real estate. More specifically, we perform the following tasks: First, the dictionary is revised in terms of its accuracy in a real estate related context. If a word's classification as positive or negative is ambiguous, it was removed, leading to the elimination of 43 words. We continue and manually analyze over 10,000 real estate-related headlines of a second news source – the *Financial Times* – regarding real-estate specific words indicating sentiment. As a result, 190 words are added to the dictionary, whereof 61 are positive and the remaining 136 are negative. For example, the terms “bubble”, “crisis” and “crash” were included in the real estate dictionary as they were missing in the finance dictionary but can be considered highly relevant in the context of real estate. In the end, the final real estate dictionary consists of 408 positive and 2,455 negative words and is slightly larger than the finance dictionary.

2.9.2 Text Pre-Processing

Given the abstracts from the *WSJ*, we pre-process the text of each abstract and convert them into well-defined sequences of linguistically meaningful units. This procedure is done in order to ensure that the computer can “understand” the language input for the following steps of the analysis and improves the quality of the dictionary-based approach. Following Uysal and Gunal (2014), the pre-processing procedure consists of four steps: lowercase conversion, stop-word removal, stemming and tokenization. Additionally, numbers and punctuations were eliminated. Stop-words removal is concerned with words such as “and”, “in” and “the”, which are usually conjunctions, prepositions, articles etc. and considered irrelevant to text classification. Stemming replaces each word within a sentence by its stem or root form as derived word forms should typically have a similar semantic meaning as their original root.[14] Finally, tokenization segments the text into smaller meaningful units called tokens. Note, that the real estate dictionary must be pre-processed accordingly to allow a comparison of news abstracts and dictionary terms. This leads to a reduced form of 959 negative and 189 positive tokens as some words in the original list are stemmed to the same root.

The example below illustrates each of the text pre-processing tasks.

We begin with the following sentence:

“Sales of US homes show a 2.7% rise.”

Eliminating numbers and punctuation leaves us with:

“Sales of US homes show a rise”.

[14] We follow Porter (1980) by using a suffix stripping algorithm, which is widely used for analyzing text corpora in English.

Stop-word removal and lowercase conversion reduce the sentence to:

"sales us homes show rise",

which can be stemmed and tokenized to the final version of:

"sale - us – home – show – rise".

Every single token of this string is then compared to the terms included in the reduced form real estate dictionary in order to measure the sentence's tone or attitude as described in Section 2.4.3.

2.9.3 Quantifying News-Based Sentiment

By applying the dictionary-based approach, we are able to transform qualitative information into quantitative data. More specifically, each positive word in a news abstract is counted as "+1" and each negative word is counted as "-1". Subsequently, this allows us to calculate a sentiment score for each abstract based on raw word counts and a multiplication factor for positive words as described in Section 3.2. This factor accounts for the fact that the sentiment dictionary does not include an equal number of positive and negative words. To avoid negatively biased results, positive words are "over-weighted" by the inverse of the total number of positive terms divided by the total number of negative words in the dictionary. Furthermore, two variations of the Positive-Negative-Ratio measures are used as sentiment indicators. The following simplified example illustrates the differences between the *absolute* and *weighted* Positive-Negative-Ratios:

Table 2.8: Calculating the PNR for Three Exemplary Abstracts A, B, C

Date	News Abstract	# of positive words	# of negative words	Sentiment score
2004:Q1	Abstract A	2	0	10.15[15]
2004:Q1	Abstract B	0	2	-2[16]
2004:Q1	Abstract C	4	2	18.30[17]

Notes: Sentiment scores are calculated based on the amount of identified sentiment words within an abstract and a multiplication factor of 1 for negative words and 5.074 for positive ones. This factor is calculated by dividing the number of negative words (959) by the number of positive ones (189) in our stemmed real estate dictionary.

When aggregating the news-based sentiment for the first quarter of 2004, the absolute Positive-Negative-Ratio (*PNR_A*) has a value of 2[18] according to equation 2.1 (two positive abstracts divided by the absolute number of one negative abstract), while the weighted Positive-Negative-Ratio (*PNR_W*) is 14.23[19]. Hence, the *PNR_A* only accounts for the raw number of abstracts with an overall positive or negative sentiment. However, the *PNR_W* takes into account the actual sentiment score that is assigned to each news abstract. This is also reflected in Table 2.1 as the *PNR_W* has higher minimum and maximum as well as standard deviation. In other words, while the *PNR_A* only accounts for the occurrence of overall optimism and pessimism in the news, the weighted Positive-Negative-Ratio accounts for the actual intensity of the respective sentiment. Therefore, the *PNR_W* is expected to be a more appropriate and precise measure of news-based sentiment.

[15] $2 \times 5.074 = 10.15$
[16] $2 \times \ -1 = \ -2$
[17] $4 \times 5.074 + 2 \times -1 = 18.30$
[18] $2/1 \ = \ 2$
[19] $10.15 + 18.30 \ / \ |-2| = 14.23$

2.9.4 Testing for Stationarity – Unit Root Test Results

As stationarity is a required assumption for time series regression techniques, we run Augmented-Dickey-Fuller tests to check for the existence of unit roots i.e. non-stationarity of variables. In accordance with our findings below, all variables were differenced once although some could be used in level form. This was done in order to be as consistent as possible and to ease interpretation of regression results.

Table 2.9: Unit Root Tests: Augmented Dickey-Fuller Test Results

Levels	**None**		**Intercept**		**Trend and Intercept**	
NPI	8.06%	*	19.98%		47.77%	
NPI_CR	3.15%	**	20,40%		48.19%	
PNR_A	13.06%		30.31%		46.41%	
PNR_W	30.40%		52.39%		50.43%	
INFLATION	6.98%	*	0.74%	***	<1.00E-04	***
TERM	42.80%		15.77%		40.97%	
SPREAD	24.97%		0.39%	***	2.11%	**
First difference	**None**		**Intercept**		**Trend and Intercept**	
Δ *NPI*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *NPI_CR*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *PNR_A*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *PNR_W*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *INFLATION*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *TERM*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***
Δ *SPREAD*	<1.00E-04	***	<1.00E-04	***	<1.00E-04	***

***Notes*:** The table above reports the findings of Augmented-Dickey-Fuller tests of all variables in levels (upper panel) and in first differences (bottom panel). The null hypothesis is presence of a unit root in a specific time series. A trend and/or an intercept can be included in the test equations to fit the time series of a variable more appropriately. Although all models are reported, numbers marked in grey indicate less appropriate models based on graphical inference. *NPI* are quarterly total returns of the NPI, and *NPI_CR* are quarterly capital appreciation returns, *PNR_A* and *PNR_W* the absolute and weighted media-expressed sentiment measures. The macroeconomic control variables are CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*) and the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*). * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

2.9.5 VAR Results with Capital Appreciation Returns Only

Table 2.10 and Table 2.11 below show the results of the analysis when using capital appreciation returns instead of NPI total returns. These results therefore correspond with Table 2.4 and Table 2.5 in the main section.

Table 2.10: VAR Results: Quarterly Appreciation Returns and Media-Expressed Sentiment

	Dependent variable: NPI capital return (quarterly)							
	(I) PNR_A		**(II) PNR_A**		**(III) PNR_W**		**(IV) PNR_W**	
	w/o CV		w/ CV		w/o CV		w/CV	
NPI_CR_{t-1}	-0.0889		-0.1282		0.1137		0.1922	
NPI_CR_{t-2}	0.0607		0.1470		-0.0480		-0.0089	
NPI_CR_{t-3}	0.0036		0.0071		0.0663		0.0273	
NPI_CR_{t-4}	-0.0276		-0.0833		-0.0742		-0.1392	
NPI_CR_{t-5}	-0.2107	***	-0.1885	**	-0.1931	**	-0.1861	*
PNR_{t-1}	0.0020		0.0025		0.0005		0.0013	
PNR_{t-2}	0.0078	**	0.0074	**	0.0061	**	0.0060	**
PNR_{t-3}	0.0042		0.0037		0.0062	*	0.0058	
PNR_{t-4}	0.0017		0.0008		0.0053	*	0.0056	*
PNR_{t-5}	-0.0028		-0.0024		-0.0059	*	-0.0056	
INTERCEPT	0.0006		0.0005		0.0003		0.0004	
Adj. R^2	0.29		0.32		0.47		0.46	
AIC	-5.81		-5.82		-6.11		-6.05	

Notes: The table above reports the estimated coefficients from the VAR (vector autoregression) models with quarterly NPI capital returns (*NPI*) and Positive-Negative-Ratio (*PNR*) as endogenous variables. The lag length of the VAR is based on the Hannan-Quinn criterion. The set of the macroeconomic control variables (*CV*) in our regression are the CPI growth (*INFLATION*), the spread between the ten-year US Treasury Bond and the 3-Month Treasury Bill yields (*TERM*), the spread between Baa- and Aaa-rated corporate bonds yields (*SPREAD*). We use Newey and West (1987) standard errors that are robust to heteroscedasticity and autocorrelation. We transformed all variables to their first differences. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

Table 2.11: Granger Causality Test Results with Capital Returns

	H_0: Media-expressed sentiment does not cause NPI_CR		**H_0: NPI_CR does not cause Media-expressed sentiment**	
	(I)	**(II)**	**(III)**	**(IV)**
	Absolute PNR	Weighted PNR	Absolute PNR	Weighted PNR
X^2 *(w/o CV)*	24.47 ***	49.54 ***	2.17	1.82
X^2 *(w CV)*	20.64 ***	37.25 ***	5.54	1.66

***Notes*:** The table above reports the Granger causality test results of the estimated VAR models of specifications (I) to (IV) of Table 2.10. Granger causality results test the joint significance of all lags for a given variable. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. The sample period is 2001:Q1 to 2016:Q4.

2.10 References

Antweiler, W. and Frank, M. Z. (2006), “Do US Stock Markets Typically Overreact to Corporate News Stories?”, *SSRN Electronic Journal*.

Baker, M. and Wurgler, J. (2007), “Investor Sentiment in the Stock Market”, *Journal of Economic Perspectives*, Vol. 21 No. 2, pp. 129–151.

Barkham, R. J. and Ward, C. W.R. (1999), “Investor Sentiment and Noise Traders: Discount to Net Asset Value in Listed Property Companies in the U.K.”, *Journal of Real Estate Research*, Vol. 18 No. 2, 291–312.

Beracha, E. and Downs, D. H. (2015), “Value and Momentum in Commercial Real Estate: A Market-Level Analysis”, *Journal of Portfolio Management*, Vol. 41 No. 6, pp. 48–61.

Beracha, E. and Wintoki, M. B. (2013), “Forecasting Residential Real Estate Price Changes from Online Search Activity”, *Journal of Real Estate Research*, Vol. 35 No. 3, pp. 283–312.

Bokhari, S. and Geltner, D. (2011), “Loss Aversion and Anchoring in Commercial Real Estate Pricing: Empirical Evidence and Price Index Implications”, *Real Estate Economics*, Vol. 39 No. 4, pp. 635–670.

Boudoukh, J., Feldman, R., Kogan, S. and Richardson, M. (2013), “Which News Moves Stock Prices? A Textual Analysis”, NBER Working Papers 18725, National Bureau of Economic Research, Cambridge, MA, 2013.

Clayton, J., Ling, D. C. and Naranjo, A. (2009), “Commercial Real Estate Valuation - Fundamentals Versus Investor Sentiment”, *The Journal of Real Estate Finance and Economics*, Vol. 38 No. 1, pp. 5–37.

Clayton, J. and MacKinnon, G. (2003), “The Relative Importance of Stock, Bond and Real Estate Factors in Explaining REIT Returns”, *The Journal of Real Estate Finance and Economics*, Vol. 27 No. 1, pp. 39–60.

Das, P., Ziobrowski, A. and Coulson, N. E. (2015a), "Online Information Search, Market Fundamentals and Apartment Real Estate", *The Journal of Real Estate Finance and Economics*, Vol. 51 No. 4, pp. 480–502.

Das, P. K., Freybote, J. and Marcato, G. (2015b), "An Investigation into Sentiment-Induced Institutional Trading Behavior and Asset Pricing in the REIT Market", *The Journal of Real Estate Finance and Economics*, Vol. 51 No. 2, pp. 160–189.

Davis, A. K., Piger, J. M. and Sedor, L. M. (2012), "Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language*", *Contemporary Accounting Research*, Vol. 29 No. 3, pp. 845–868.

Dickey, D. A. and Fuller, W. A. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, Vol. 74 No. 366, p. 427.

Feldman, R., Govindaraj, S., Livnat, J. and Segal, B. (2010), "Management's Tone Change, Post Earnings Announcement Drift and Accruals", *Review of Accounting Studies*, Vol. 15 No. 4, pp. 915–953.

Freybote, J. (2016), "Real Estate Sentiment as Information for REIT Bond Pricing", *Journal of Property Research*, Vol. 33 No. 1, pp. 18–36.

Freybote, J. and Seagraves, P. A. (2017), "Heterogeneous Investor Sentiment and Institutional Real Estate Investments", *Real Estate Economics*, Vol. 45 No. 1, pp. 154–176.

Granger, C. W. J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica*, Vol. 37 No. 3, p. 424.

Henry, E. and Leone, A. J. (2016), "Measuring Qualitative Information in Capital Markets Research. Comparison of Alternative Methodologies to Measure Disclosure Tone", *The Accounting Review*, Vol. 91 No. 1, pp. 153–178.

Heston, S. L. and Sinha, N. R. (2016), “News versus Sentiment - Predicting Stock Returns from News Stories”, *Finance and Economics Discussion Series*, Vol. 2016 No. 048, pp. 1–35.

Kahneman, D. and Tversky, A. (1979), “Prospect Theory: An Analysis of Decision under Risk”, *Econometrica*, Vol. 47 No. 2, p. 263.

Lang, M. (2018), “Essays on Sentiment Analysis through Textual Analysis in Real Estate Markets”, Dissertation, in IREIBS International Real Estate Business School (Ed.), *Schriften zu Immobilienökonomie und Immobilienrecht*, Vol. 88, Universitätsbibliothek Regensburg, Regensburg.

Lin, C. Y., Rahman, H. and Yung, K. (2009), “Investor Sentiment and REIT Returns”, *The Journal of Real Estate Finance and Economics*, Vol. 39 No. 4, pp. 450–471.

Ling, D. C., Marcato, G. and McAllister, P. (2009), “Dynamics of Asset Prices and Transaction Activity in Illiquid Markets: the Case of Private Commercial Real Estate”, *The Journal of Real Estate Finance and Economics*, Vol. 39 No. 3, p. 359.

Ling, D. C., Naranjo, A. and Scheick, B. (2014), “Investor Sentiment, Limits to Arbitrage and Private Market Returns”, *Real Estate Economics*, Vol. 42 No. 3, pp. 531–577.

Loughran, T. and McDonald, B. (2011), “When Is a Liability Not a Liability? - Textual Analysis, Dictionaries, and 10-Ks”, *The Journal of Finance*, Vol. 66 No. 1, pp. 35–65.

Marcato, G. and Nanda, A. (2016), “Information Content and Forecasting Ability of Sentiment Indicators: Case of Real Estate Market.”, *Journal of Real Estate Research*, Vol. 38 No. 2, 165-2013.

Newey, W. K. and West, K. D. (1987), “Hypothesis Testing with Efficient Method of Moments Estimation”, *International Economic Review*, Vol. 28 No. 3, p. 777.

Porter, M. F. (1980), “An Algorithm for Suffix Stripping”, *Program*, Vol. 14 No. 3, pp. 130–137.

Price, S. M., Seiler, M. J. and Shen, J. (2017), “Do Investors Infer Vocal Cues from CEOs During Quarterly REIT Conference Calls?”, *The Journal of Real Estate Finance and Economics*, Vol. 54 No. 4, pp. 515–557.

Said, S. E. and Dickey, D. A. (1984), “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order”, *Biometrika*, Vol. 71 No. 3, p. 599.

Seiler, M. J., Seiler, V. L. and Lane, M. A. (2012a), “Mental Accounting and False Reference Points in Real Estate Investment Decision Making”, *Journal of Behavioral Finance*, Vol. 13 No. 1, pp. 17–26.

Seiler, M. J., Seiler, V. L., Lane, M. A. and Harrison, D. M. (2012b), “Fear, Shame and Guilt: Economic and Behavioral Motivations for Strategic Default”, *Real Estate Economics*, Vol. 40 No. 4, S199-S233.

Shiller, R. (2007), “Understanding Recent Trends in House Prices and Home Ownership”, NBER Working Paper No. 13553, National Bureau of Economic Research, Cambridge, MA, 2007.

Soo, C. K. (2015), “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.

Tetlock, P. C. (2007), “Giving Content to Investor Sentiment - The Role of Media in the Stock Market”, *The Journal of Finance*, Vol. 62 No. 3, pp. 1139–1168.

Tsolacos, S., Brooks, C. and Nneji, O. (2014), “On the Predictive Content of Leading Indicators: The Case of U.S. Real Estate Markets”, *Journal of Real Estate Research*, Vol. 36 No. 4, pp. 541–574.

Tversky, A. and Kahneman, D. (1991), “Loss Aversion in Riskless Choice: A Reference-Dependent Model”, *The Quarterly Journal of Economics*, Vol. 106 No. 4, pp. 1039–1061.

Uysal, A. K. and Gunal, S. (**2014**), "The Impact of Preprocessing on Text Classification", *Information Processing & Management*, Vol. 50 No. 1, pp. 104–112.

Walker, C. B. (**2014**), "Housing Booms and Media Coverage", *Applied Economics*, Vol. 46 No. 32, pp. 3954–3967.

Walker, C. B. (**2016**), "The Direction of Media Influence - Real-Estate News and the Stock Market", *Journal of Behavioral and Experimental Finance*, Vol. 10, pp. 20–31.

3 News-Based Sentiment Analysis in Real Estate: A Machine-Learning Approach

3.1 Abstract

This paper examines the relationship between news-based sentiment, captured through a machine-learning approach, and the US securitized and direct commercial real estate markets. Thus, we contribute to the literature on text-based sentiment analysis in real estate by creating and testing various sentiment measures by utilizing trained support vector networks. Using a vector autoregressive framework, we find the constructed sentiment indicators to predict the total returns of both markets. The results show a leading relationship of our sentiment, even after controlling for macroeconomic factors and other established sentiment proxies. Furthermore, empirical evidence suggests a shorter response time of the indirect market in relation to the direct one. The findings make a valuable contribution to real estate research and industry participants, as we demonstrate the successful application of a sentiment-creation procedure that enables short and flexible aggregation periods. To the best of our knowledge, this is the first study to apply a machine-learning approach to capture textual sentiment relevant to US real estate markets.

Keywords: Textual Analysis, News-Based Sentiment, Machine-Learning, US Commercial Real Estate, Support Vector Machine

3.2 Introduction

Over the past decade, real estate researchers have intensified their efforts to investigate how sentiment affects individual decision-makers (Freybote and Seagraves, 2017), institutions (Das *et al.*, 2015) and hence, property markets themselves (Ling *et al.*, 2014; Marcato and Nanda, 2016). There is general consensus on the complexity of influencing factors, and that investors should not be considered as rational utility-maximizers only, thus indicating the overall importance of sentiment. Furthermore, real estate investors may be especially sensitive to sentiment, due to real estate market characteristics such as the relatively low market transparency and long transaction periods, leading to information asymmetries. Conducting a survey on decision-making among individuals actively involved in the property investing process, Gallimore and Gray (2002) found that individuals are in fact aware of the importance of sentiment for their own decisions.

Recent works further support the notion, that the augmentation of sentiment proxies in fundamental market models enhances their explanatory power. For example, Ling *et al.* (2014) confirm a relationship between investor sentiment and subsequent returns in the private commercial real estate market, which drives prices away from fundamentals. Walker (2014) showed similar findings for the UK housing market, suggesting media sentiment to have a significant impact on real house price changes.

This paper seeks to deepen the knowledge of a rather new field of sentiment analysis based on news items instead of traditional indicators such as investor surveys. Some initial research by Soo (2015), Walker (2014, 2016) and Nowak and Smith (2017) has assessed the relationship between textual sentiment measures and the residential real estate market, deploying sentiment-annotated word lists. However, no study evidently uses supervised machine-learning to extract news-based sentiment relevant to the US real estate market. Therefore, this paper examines the relationship between news-based sentiment, captured through a classification algorithm, and the US securitized and direct commercial real estate markets.

After training a support vector machine algorithm (SVM) for classification, we analyze approximately 54,500 real estate (RE) news headlines from the *S&P Global Market Intelligence* database (SNL) concerning their inherent sentiment. Thereby, the machine-learning algorithm assigns either a positive, negative or neutral score to each news headline, which is subsequently aggregated to different monthly measures of market sentiment. Based on psychological theory and existing research, we introduce an optimism indicator (*OI*), a pessimism indicator (*PI*) and a weighted sentiment quotient (*SQ*). A vector autoregressive framework (VAR) enables us to investigate the dynamic relationship between these three created sentiment measures and the securitized and direct real estate markets in the United States.

The findings indeed indicate strong and consistent evidence of a significant relationship between our sentiment indicators and real estate market movements. For both markets, especially the pessimism indicator provides additional information to macroeconomic fundamentals in explaining market returns. The predictive power of our indicator remains intact, even when controlling for the influence of other traditional sentiment measures, such as the *Survey of Consumers* of the University of Michigan or the *American Association of Individual Investors* (*AAII*) *Investor Sentiment Survey*. The *PI* drives total returns of the securitized and direct real estate market by one and – slightly delayed – by two and three months, respectively. As comparable results were not found for the optimism indicator, these findings indicate a potentially existing negativity bias of real estate market participants. As the analysis does not find a significant impact of preceding market performance on current sentiment measures, a statistically significant bi-directional relationship cannot be claimed.

These results provide an additional opportunity to better understand influences on real estate market returns that are not based on fundamental value changes. Furthermore, a new technique for extracting sentiment from one of the most widespread information sources – news – is applied, contrasted and discussed. The knowledge gained can be applied to every form of text corpus, such as earnings press releases, annual reports, IPO prospectus, corporate disclosures, analyst reports, tweets or blog posts. Hence, the study makes a valuable contribution to the extraction of sentiment itself and participates in the recently emerging strand of literature concerning textual analysis in real estate. Additionally, it sheds light on real estate news analytics, as an innovative source of sentiment and an opportunity to construct a leading market indicator.

This paper itself is organized as follows. In Section 3.3, we provide a synopsis of the relevant literature on textual analysis finding its way into the broad field of sentiment analysis. Furthermore, recent research on sentiment analysis in the context of real estate is discussed. The subsequent section introduces various datasets, while Section 3.5 presents the machine-learning approach, as well as the methods of aggregating the sentiment measures. Furthermore, the VAR framework is derived. Section 3.6 shows the empirical results and the conclusion draws upon the entire work and discusses implications of our findings for the industry, as well as future research.

3.3 Literature Review

3.3.1 Sentiment Analysis and the Subcategory of Textual Analysis

"The effects of noise on the world, and on our views of the world, are profound" (Black, 1986, p. 529). According to Black, noise has several meanings and impacts on economic activity in various ways; noise entails expectations, which do not follow any rational rules, is a form of uncertainty that changes investment flows, is information not yet arrived at every market participant, and subsumes the reasons for markets to be inefficient. Hence, noise enables trading in financial markets (Black, 1986). What Black laconically describes as "noise", can nowadays be considered at least partially as sentiment.

Following this rationale, there have been several attempts since the mid-1980s to explain asset prices deviating from intrinsic values, which are not based on underlying value changes (Brown and Cliff, 2004). After 2000, the debate on how to quantify sentiment intensified (Liu, 2012). In general, one can now distinguish between two different ways of measuring sentiment. On the one hand, there are indirect indicators, which are market-based, claiming to proxy sentiment such as closed-end fund discounts, buy-sell imbalance or mortgage fund flows (Brown and Cliff, 2004). On the other hand, one can rely on surveys as a direct measure of investor sentiment. Qiu and Welch (2006) discuss several survey-based sentiment indices, for example, the consumer confidence index or the AAII index, a survey of individual investors.

Recently, researchers have shown an increased interest in a new subcategory of sentiment analysis, so-called textual analysis. The digitalization of information and news, increasing computational power, and new techniques for analyzing text corpora fuel the rapid growth of this research area (Liu, 2012). A diverse variety of textual documents such as earnings press releases (Henry, 2008; Henry and Leone, 2016), news articles (Tetlock, 2007; Sinha, 2016; Hanna *et al.*, 2017), annual reports (Li, 2006) or IPO prospectus (Ferris *et al.*, 2013), corporate disclosures (Rogers *et al.*, 2011; Ozik and Sadka, 2012), and analyst reports (Twedt and Rees, 2012) were analyzed in order to extract sentiment and draw conclusions about market events.

Textual analysis techniques can neither be perfectly assigned to the group of indirect sentiment measures nor to direct ones and are therefore best described as in-between. Analyzing textual statements is not the same as surveys, where the participants are directly asked about their current state of sentiment. Nevertheless, the textual indicators are also not indirect proxies such as buy-sell imbalances, which theoretically proxy market sentiment but are originally a measure of other aspects of the market. They behave more like a mixture of both kinds of measures which justifies their in-between position.

When analyzing the relationship between sentiment and the market, textual analysis provides promising results for a wide range of domains such as market indices (Schumaker and Chen, 2009; Bollen *et al.*, 2011), exchange rates (Jin *et al.*, 2013; Chatrath *et al.*, 2014), company stock prices (Tetlock *et al.*, 2008), earnings (Li, 2010), trading volume or market volatility (Tetlock, 2007).

3.3.2 Dominant Methodologies in Textual Analysis

In recent years, two methodologies for conducting textual analysis have been predominant. Originally, the dictionary-based approach was introduced to the finance literature by Tetlock in 2007. It classifies phrases or sentences by comparing the textual documents word-by-word to pre-annotated word lists where each term is linked to a certain sentiment such as positive, negative or neutral. Examining news articles from *The Wall Street Journal*, Tetlock found that high media pessimism temporarily leads to downward pressure on market prices and higher market volatility. In a subsequent paper, Tetlock *et al.* (2008) again made use of the *Harvard University's General Inquirer (GI)* as sentiment dictionary in order to forecast firm earnings. Several papers followed his approach and applied both the methodology and the GI/Harvard dictionary in the most diverse contexts. Among others, Kothari *et al.* (2009)

investigated the relationship between company disclosures and the return volatility, as well as cost of capital and analyst forecast dispersion. Arguing that the meaning of words may depend on certain circumstances, Loughran and McDonald (2011) developed a financial-language-orientated word list especially for business communication. Based on their findings, researchers started to compare domain-specific dictionaries to general ones (Henry and Leone, 2016; Rogers *et al.*, 2011; Doran *et al.*, 2012) or added domain-specific words (Hanna *et al.*, 2017). Henry and Leone (2016) report that the investigation of financial disclosures with a domain-specific word list leads to superior results.

The second methodology focuses on sentiment classification algorithms such as support vector machines or the Naïve Bayes classifiers. Two of the earliest works of Pang *et al.* (2002) and Antweiler and Frank (2004) conducted an analysis with both techniques. Classifying movie reviews as positive or negative, Pang *et al.* (2002) showed that Naïve Bayes as well as SVM led to good results, whereby the SVM provided the most promising findings. Antweiler and Frank (2004) investigated more than 1.5 million message board postings on *Yahoo! Finance* and *Raging Bull* about a group of 45 companies and determined the predictive power of their sentiment measure on next day returns and volatility. Furthermore, they report that disagreement in sentiment during the period under consideration is linked to increased trading volume. At firm level, Li (2010) analyzed MD&As from 1994 to 2007 with the Naïve Bayes algorithm. The extracted tone is linked significantly to future earnings and liquidity and has predictive power with respect to future performance. Further techniques categorized by Khadjeh Nassirtoussi *et al.* (2014) are regression algorithms (Schumaker *et al.*, 2012), decision rules or decision trees (Rachlin *et al.*, 2007), combinatory algorithms and multi-algorithm experiments (Das and Chen, 2007).

Both methodologies have their respective advantages and disadvantages. In short, the dictionary-based approach is usually more transparent and easier to implement, once a dictionary is selected. Nevertheless, as literature has shown, choosing an appropriate pre-annotated word-list is crucial, as words may have different meanings in different contexts (see e.g. Loughran and McDonald, 2011). Hence, sentiment dictionaries need to be adapted first. Another disadvantage of the dictionary-based approach is that it is restricted to the words in the selected dictionary. Applying a machine-learning approach by contrast is more complicated, but at the same time a lot more flexible regarding future adjustments. A downside is the missing consensus in literature on the best way of deriving an appropriate training dataset necessary for the approach. Nevertheless, machine-learning approaches tend to yield a higher classification accuracy than the dictionary-based approaches (Li, 2010). In line with this, we apply an algorithm for classification, namely a support vector machine, to extract news-based sentiment relevant to the US real estate market.

3.3.3 Sentiment Analysis in the Context of Real Estate

As early papers only extend back to the beginning of 2000 (Barkham and Ward, 1999; Gallimore and Gray, 2002), the real estate sentiment literature lags behind related research in finance. However, there has lately been an increasing amount of literature on sentiment analysis in the context of real estate.

Conducting a survey among 983 UK property investors about their decision-making, Gallimore and Gray (2002) make the astounding discovery that personal feelings and the views of other market participants are almost equally important to fundamental market information. Subsequent research confirms these initial findings across real estate market sectors. Clayton *et al.* (2009) and Ling *et al.* (2014) examine the commercial real estate market, and find evidence that investor-sentiment measures among others in the form of the *Real Estate Research Corporation Investment Survey* have a significant linkage to pricing and market returns in subsequent periods. Changes in market sentiment during the downturn in the UK commercial property markets (i.e. the second half of 2007) motivated Crosby *et al.* (2010) to analyze the client influence on performance measurement appraisals. They found that pressure on fund managers might be an explanation for different appraisal outcomes. Lin *et al.* (2009) and Das *et al.* (2015) took a closer look at REIT performance, and Marcato and Nanda (2016) among others, at residential real estate returns.

Similar to the financial literature, real estate sentiment research was traditionally conducted facilitating direct and indirect sentiment measures, as so do all the above-mentioned research papers. Over time however, new ways of measuring sentiment have emerged. Online search engine volume provided by Google Trends have been successfully established as a new way of measuring real estate market sentiment (Hohenstatt *et al.*, 2011; Dietzel *et al.*, 2014; Rochdi and Dietzel, 2015). Equivalently, the stream of textual-analysis-based sentiment measures is slowly finding its way into real estate research. Some first attempts were made by Walker (2014), making use of the dictionary-based approach. He found that past newspaper articles about the housing market Granger-cause house price changes in in the UK, even when controlling for different control variables. His findings were confirmed on a city level in the US. With 37,500 local housing news articles, Soo (2015) successfully applied the dictionary-based approach and argues that her sentiment measure leads house-price movements by more than two years. In accordance with his findings in 2014, Walker (2016) found further evidence that the media is a reliable source of sentiment in the real estate housing market.

Together, these studies provide insights into sentiment analysis in the field of real estate, but little is known about the potential of other methods to investigate text corpora. Extracting relevant real estate sentiment is still limited mainly to dictionary-based approaches. No study has so far applied a machine-learning approach in a real estate context. Hence, the present paper is the first to use a sentiment classification algorithm to extract sentiment from qualified news items and quantify the performance in relation to the securitized and the direct commercial real estate markets.

Thus, we state our first research question as follows: *(1) Can sentiment measures created via machine-learning predict the securitized commercial real estate market?*

Furthermore, it is worth investigating, whether the results deviate, when switching to the direct real estate market. Hence, the second research question follows directly: *(2) Is the predictive power different for the direct real estate market?*

As there have been several attempts at measuring sentiment with direct and indirect indicators, the third research question considers measuring the relative quality: *(3) How do the created sentiment indicators perform in addition to established sentiment measures?*

Finally, research question 4 is based on the notion of an existing negativity bias (Rozin and Royzman, 2001), which refers to the idea that the human psychological state is affected more strongly by negative entities – in this case, news stories – than by positive ones. Given that Tetlock (2007) found corresponding evidence in terms of stock market sentiment, we construct various sentiment measures accordingly and formulate the fourth research question as: *(4) Is there evidence that market participants react differently to negative news in contrast to positive ones?*

3.4 Data

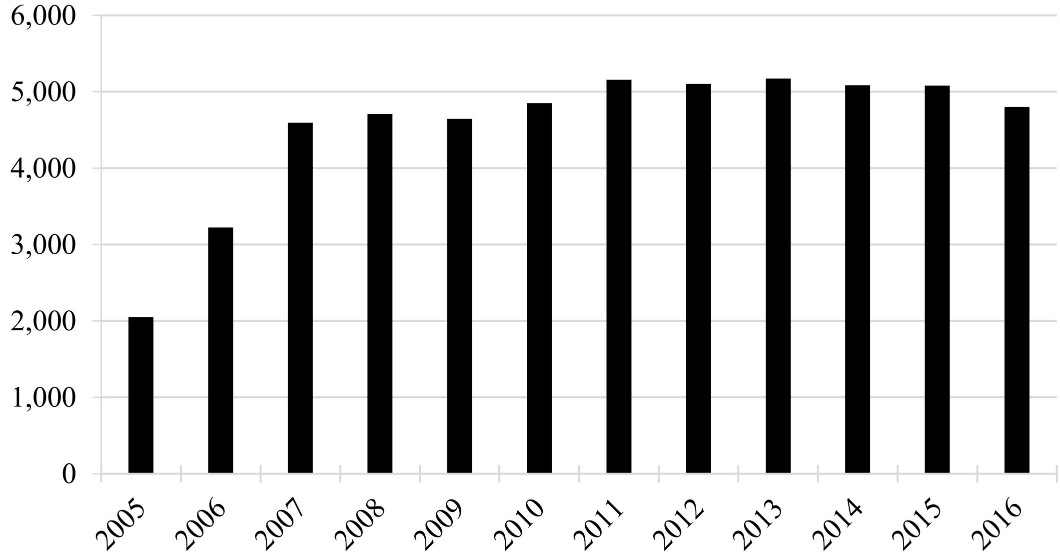
To examine the relationship between news-based sentiment and the real estate market in the United States, we use two types of dataset: (1) a news text corpus and (2) real estate return data, as well as further economic time series. The availability of historic news in the digital archive of our data source restricts the overall research period. Thus, we collect all data from January 1st, 2005 to December 31st, 2016. This 12-year period is worth investigating, since it contains a boom phase (until 2007), the housing bubble bust and the recession from 2007 to 2009, as well as the pronounced recovery market phase in the subsequent years.

3.4.1 News Data

The identification of a suitable text corpus that is relevant to the commercial real estate market is decisive to building an accurate real estate market sentiment measure. Hence, we base our sentiment analysis upon professional financial news from the *S&P Global Market Intelligence* platform. The platform offers real-time updates, trends, market activities and reporting which is specific to the real estate market. Due to the expertise of reporting on SNL, we assume the news to be more comprehensive and reliable than news usually directed to the public. Over the 12-year time span, 54,530 articles including the keyword "real estate" were collected. This corresponds to more than 370 real estate news items per month. Following Peramunetilleke and Wong (2002), who argue that headlines are normally short and straight-to-the-point, this paper analysis news headlines only.

Figure 3.1 presents the amount of real estate-related news published by SNL over the 12-year research period. During the boom market, from 2005 to 2007, the news coverage more than doubles from about 2,050 to 4,595 annual news. This might be the result of an increased interest in real estate, but probably also due to the rise of the internet and hence, more and more people reading news online.

Figure 3.1: SNL Real Estate News Coverage, 2005 – 2016



Notes: This figure plots the sample distribution of real-estate-related news published by *S&P Global Market Intelligence* (SNL) over the sample period, 2005:M1 to 2016:M12. All news was retrieved using the digital archive of SNL by selecting articles that contain the keyword "real estate".

During the bust of the subprime mortgage crisis, the annual news coverage stabilizes at around 4,700 news items, and reaches its peak in 2011 with 5,158 news annually. In comparison, the post-crisis level of annual news coverage is steadily higher than the prior bust-level in 2007/08. This may indicate an increased attention-level concerning real estate as an asset class.

3.4.2 Real Estate Data

The return data of the direct real estate market stems from a repeat-sales index provided by *CoStar*. More specifically, we select the *CoStar Commercial Repeat-Sale Index* (*CCRSI*) as an accurate and comprehensive measure of commercial real estate prices in the United States. As a measure of overall market performance, the value-weighted US Composite Price Index is chosen. The index is published monthly and is available at www.costargroup.com.

Furthermore, we derive the return data of the securitized market from the *National Association of Real Estate Investment Trusts* (*NAREIT*), selecting the *FTSE/NAREIT All Equity REIT Total Return Index* as a market-capitalization-weighted, free-float-adjusted index of equity REITs in the United States (www.reit.com). We use the monthly percentage changes of both indices to measure the total returns from the direct and securitized commercial real estate market, respectively.

3.4.3 Further (Economic) Data

To control for other potential influencing factors causing variations in real estate sentiment and returns, a selected set of control variables is included which covers macroeconomic, capital market and property market fundamentals bearing the potential to influence the business cycle. Two distinct sets of control variables are used for the direct and securitized market, respectively. To account for the impact of debt market conditions, the VAR frameworks firstly incorporate a measure of overall economic default risk (*SPREAD*), defined as the difference between Moody's Seasoned Baa- and Aaa-rated corporate bonds (e.g. Lin *et al.*, 2009; Ling *et al.*, 2014). Secondly, we include a term structure variable (*TERM*), as a mean for expectations of future economic developments, defined as the difference between the yields on the 10-year Treasury bond and the 3-month Treasury bill (e.g. Clayton *et al.*, 2009; Freybote and Seagraves, 2017). The analysis of the securitized real estate market controls for percentage changes of the Consumer Price Index (*CPI*) since real estate is often regarded as a hedge against inflation (e.g. Hoesli *et al.*, 2008).[20] To account for the performance of the general stock market, we also incorporate the return of the S&P500 composite index (*SP500*) in our analysis (e.g. Schätz and Sebastian, 2010; Das *et al.*, 2015). Additionally, incorporating initial claims of unemployment insurance (*UNEMPL*) controls for labor market developments and therefore proxies general macroeconomic conditions as well as space market demand when explaining direct real estate market returns (Brooks and Tsolacos, 1999). Total construction spending (*CONSTR*) is also implemented to account for the supply side or overall activity of the development industry (e.g. Dietzel *et al.*, 2014).

Table 3.1 presents the descriptive statistics of monthly returns and other variables. We state the mean, median, standard deviation, minimum and maximum. Total returns range from -6.87% to 3.18% and -31.67% to 31.02% for the direct and securitized market, respectively. The volatility, measured per standard deviation of the securitized

[20]Controlling for the CPI did not substantially alter the results in terms of sign, size and significance of the sentiment indicators and autoregressive components when running the direct commercial real estate market models. As we adapted the controls to better reflect markets without losing to many degrees of freedom we report regression equations without the CPI.

market is more than four times greater than of the direct one. The overall volatility in returns is the result of the boom and bust phases included in our sample period.

Table 3.1: Descriptive Statistics: Real Estate Returns and Economic Time Series

Statistic	**Mean**	**Median**	**SD**	**Min**	**Max**
CCRSI (%)	0.34	0.59	1.59	-6.87	3.18
NAREIT (%)	0.88	1.25	6.91	-31.67	31.02
SPREAD (%)	1.13	0.96	0.51	0.55	3.38
TERM (%)	1.87	2.01	1.08	-0.52	3.69
INFL (%)	0.17	0.19	0.43	-1.92	1.22
SP500 (%)	0.51	1.02	4.10	-16.94	10.77
UNEMPL	350,036	318,466	102,575	200,456	717,000
CONSTR	83,815	82,235	15,204	50,973	110,020

Notes: This table reports summary statistics of our monthly real estate return data and macroeconomic time series. *CCRSI* is the total return of the *CoStar Commercial Repeat-Sale Index*. *NAREIT* is the total return of the *FTSE/NAREIT All Equity REIT Total Return Index*. *SPREAD* is the difference between Baa- and Aaa-rated corporate bonds yields. *TERM* is the difference between the 10-year US Treasury bond and the 3-Month Treasury bill yields. *CPI* is the percentage change of the Consumer Price Index (CPI). *SP500* is the total return of the S&P 500 Composite Index. *UNEMPL* is the amount of unemployment initial claims in number of persons. *CONSTR* is the amount of construction spending in millions of dollars. Percentages are expressed in decimal form. The sample period is 2005:M1 to 2016:M12.

To test the robustness of our sentiment measures, we further control for a set of more "general" and well-established sentiment indicators such as the *Surveys of Consumers* of the University of Michigan (*CONSUSENTI*). We also incorporate the bullish and bearish measures of the *American Association of Individual Investors (AAII) Investor Sentiment Survey* (*AAIIBULL*, *AAIIBEAR*) as well as of the *Investors Intelligence US Advisors' Sentiment Report* (*ADVSENTBULL*, *ADVSENTBEAR*). From the *Economic Policy Uncertainty* platform, their *News-Based Policy-Related Uncertainty* measure (*ECOPOLUNCERTINEWS*), the *Overall Policy-Related Economic Uncertainty* indicator (*ECOPOLUNCERTIOVER*) or *Equity Market-Related Economic Uncertainty* (*ECOUNCERT*) is used. For a full description of all variables, see Table 3.2. All data was obtained from the *Federal Reserve Bank of St. Louis* (www.fred.stlousfed.org) and *Thomson Reuters Datastream* (www.financial.thomsonreuters.com) on a monthly basis.

Table 3.2: Data Description

Variable label	Description	Unit	Mnemonic	Source	RE Market
10Y	10-Year Treasury Constant Maturity Rate	Percent	GS10	FRED	Direct, Indirect
3M	3-Month Treasury Constant Maturity Rate	Percent	GS3M	FRED	Direct, Indirect
AAA	Moody's Seasoned Aaa Corporate Bond Yield	Percent	AAA	FRED	Direct, Indirect
AAIIBEAR	US Sentiment Survey: AAII % Bearish	Percent	USAAIIN	TR Datastream	Direct, Indirect
AAIIBULL	US Sentiment Survey: AAII % Bullish	Percent	USAAIIP	TR Datastream	Direct, Indirect
ADVSENTBEAR	Advisors' Sentiment Bearish	Percent	USIIBER	TR Datastream	Direct, Indirect
ADVSENTBULL	Advisors' Sentiment Bullish	Percent	USIIBUL	TR Datastream	Direct, Indirect
BAA	Moody's Seasoned Baa Corporate Bond Yield	Percent	BAA	FRED	Direct, Indirect
CONSTR	Total Construction Spending	Million USD	TTLCON	FRED	Direct
CONSUSENTI	University of Michigan: Consumer Sentiment	Index	UMCSENT	FRED	Direct, Indirect
CPI	Consumer Price Index for All Urban Consumers	Price Index	CPIAUCNS	FRED	Indirect
ECOPOLUNCERTINEWS	US Economic Policy Uncertainty Index – News-Based	Index	USEPUNEWR	TR Datastream	Direct, Indirect
ECOPOLUNCERTOVER	US Economic Policy Uncertainty Index – Overall	Index	USEPUPOLR	TR Datastream	Direct, Indirect
ECOUNCERT	US Equity Related Economic Uncertainty	Index	USEPUEQ	TR Datastream	Direct, Indirect
S_P500	S&P 500 Composite	Price Index	S&PCOMP	TR Datastream	Indirect
UNEMPL	Civilian Unemployment Rate	Percent	UNRATENSA	FRED	Direct

Notes: Series were taken from Federal Reserve Bank of St. Louis (FRED) and Thomson Reuters Datastream (TR). The data span for all series is 2005:M1 to 2016:M12.

3.5 Methodology

3.5.1 Sentiment Extraction via Machine-Learning

To extract sentiment from news headlines, this paper deploys a support vector machine as a supervised learning algorithm. Support vector machines or support vector networks are machine-learning techniques for two-group classification tasks proposed by Cortes and Vapnik (1995) during the nineties. In theory, each headline is depicted as an input vector in some high-dimensional feature space via a non-linear mapping technique chosen a priori, where a linear decision surface is constructed to distinguish between different classes. As supervised learning technique, this requires a pre-classified set of training data, which are used to construct the decision surface described above. Our training set comprises of a balanced sample of about 4,500 pre-classified headlines selected randomly within the full SNL text corpus.[21] Knowing the position of the hyperplane, subsequently allows identifying the category of additional headlines, depending on their position in the feature space, relative to the surface. More conveniently, one can imagine that training headlines – already assigned to one class or the other – are depicted as a set of data points in space and a simple hyperplane is constructed that separates the points from one class to the other. Given this so-called decision surface, one can afterwards determine the class of new dots or headlines solely by their position relative to this hyperplane.

Following Cortes and Vapnik (1995), a set of pre-classified training data $(y_1, \boldsymbol{x_1}), \dots, (y_l, \boldsymbol{x_l}), y_i \in \{-1,1\}$ is linearly separable, if the inequality $y_i(\boldsymbol{w x_i} + b) - 1 \geq 0, i = 1, \dots, l$ is fulfilled for all training elements.[22] Hence, the optimal hyperplane $\boldsymbol{w_0 x} + b_0 = 0$ is the decision surface that separates the training data with the maximal margin i.e. maximizes the distance $\rho(\boldsymbol{w_0}, b_0) = \frac{2}{||\boldsymbol{w}||} = \frac{2}{\sqrt{\boldsymbol{ww}}}$ between data points on the edge of each class.[23] These training vectors $y_i(\boldsymbol{w x_i} + b) - 1 = 0$ are called support vectors.

[21] Note that only the remaining headlines are used to calculate the sentiment indicators afterwards. This makes sure that algorithm 'tuning" does not influence classification results.
[22] For ease of reading, we stick to the common notation of matrices using bold characters.
[23] Because it is mathematically more convenient, the optimal hyperplane can be derived by minimizing 0,5 $\boldsymbol{w} \times \boldsymbol{w}$ subject to $y_i(\boldsymbol{w x_i} + b) - 1 \geq 0, i = 1, \dots, l$.

Cortes and Vapnik (1995) show that the vector $\boldsymbol{w}_o$, which determines the optimal decision surface, is a linear combination of these vectors:

$$\boldsymbol{w}_o = \sum_{i=1}^{l} \alpha_i^0 y_i \boldsymbol{x}_i, \tag{3.1}$$

where $\alpha_i^0 \geq 0$.

To find the parameters of α_i, the algorithm has to solve the following quadratic programing problem:

$$W(\boldsymbol{\Lambda}) = \boldsymbol{\Lambda}^T \boldsymbol{1} - \frac{1}{2} \boldsymbol{\Lambda}^T \boldsymbol{D} \boldsymbol{\Lambda}, \tag{3.2}$$

with respect to $\boldsymbol{\Lambda}^{\boldsymbol{T}} = (\alpha_1, \ldots, \alpha_l)$ subject to the constraints of $\boldsymbol{\Lambda}^{\boldsymbol{T}} \boldsymbol{Y} = 0$ and $\boldsymbol{\Lambda} \geq 0$, where ***1*** is a l-dimensional unit vector, $\boldsymbol{Y}^T = (y_1, \ldots, y_l)$ the *l*-dimensional vector of labels and $\boldsymbol{D}$ the symmetric *l* x *l* - matrix $D_{ij} = y_i y_j \boldsymbol{x}_i \boldsymbol{x}_j$ with $i, j = 1, \ldots, l$. Given $\boldsymbol{w_0}$, one can solve $\boldsymbol{w_0 x} + b_0 = 0$ for b_0, which provides us with all parameters required to state the optimal, maximal margin hyperplane. Hence, new data $\tilde{\boldsymbol{x}}$ can be classified applying a signum function:

$$f(\tilde{\boldsymbol{x}}) = sign(\boldsymbol{w_o} \tilde{\boldsymbol{x}} + b_0). \tag{3.3}$$

Positive results indicate a class of "+1" and vice versa. Referring back to the aforementioned "data points in space" example, equation (3.3) mathematically determines the position of new headlines relative to the decision surface and thereby assigns the class "+1" or vice versa.

Due to the possibility that that training data may not be separable by a hyperplane without classification errors, we follow Cortes and Vapnik (1995) and use a so-called *soft-margin classifier* by introducing some non-negative "slack" variable $\xi_i \geq 0, i = 1, \ldots, l$ and minimize $\frac{1}{2} \boldsymbol{ww} + C \sum_{i=1}^{l} \xi_i$ subject to $y_i(\boldsymbol{w x_i} + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$. The constant C is considered as a trade-off parameter between error and margin. Thus, one still has to solve (3.2) with respect to $\boldsymbol{\Lambda}^{\boldsymbol{T}} = (\alpha_1, \ldots, \alpha_l)$, but subject to slightly adjusted constraints of $\boldsymbol{\Lambda}^{\boldsymbol{T}} \boldsymbol{Y} = 0$ and $C \times \boldsymbol{1} \geq \boldsymbol{\Lambda} \geq 0$. This allows the technique to ignore some "misplaced dots" when constructing the decision surface.

To render the classification algorithm even more versatile, the data is not mapped into the input space, but some higher dimensional feature space using the so-called "kernel trick". This enables separating data by a decision surface, even when they are not linearly separable in the input space. Thereby, the hyperplane becomes flexile and now behaves more like a moldable blanket than a rigid plate when separating the data points. Mathematically, an N-dimensional vector function $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^N$ transfers the n-dimensional input vector $\boldsymbol{x}$ into the N-dimensional space. One then constructs the linear separator $\boldsymbol{w}$ and parameter *b,* using the transformed vectors $\phi(\boldsymbol{x}_i) = \phi_1(\mathbf{x}_i), \phi_2(\mathbf{x}_i), \ldots, \phi_N(\mathbf{x}_i), i = 1, \ldots, l$ in the same manner described above. "New" data can be classified by transforming the "data" vector into the feature space ($\widetilde{\boldsymbol{x}} \rightarrow \phi(\widetilde{\boldsymbol{x}})$) first, and then applying the sign function afterwards:

$$f(\widetilde{\boldsymbol{x}}) = sign(\boldsymbol{w_o}\phi(\widetilde{\boldsymbol{x}}) + b_0). \tag{3.4}$$

Additionally, in order to classify textual documents into three different sentiment categories, a few obstacles must be tackled. First, a support vector machine does not work without converting the textual documents into numeric vectors beforehand. Therefore, training headlines are split into single words or features. Combined with corresponding word frequencies, these features are then listed in a so-called document-term matrix, in which each training headline is represented by a numeric row vector. Hence, each feature of the training data set becomes one dimension of the input space. For new data, a vector is constructed by counting how often these training features are included in the headline, and using the respective frequencies as the coordinates of the corresponding dimension. Second, a support vector network just distinguishes between two classes. As we are using the categories "positive", "negative" and "neutral", this requires us to run three different support vector machines with two categories each. At the end, a voting system assigns headlines to the class with the highest number of votes.

3.5.2 Creating Real Estate Sentiment Measures

After classifying each headline as either positive, negative or neutral, the respective sentiments for monthly observation periods are aggregated. Because this study explores the relationship between news-based sentiment and the real estate market comprehensively, we do not restrict our analysis to a single sentiment measure, but propose three different ones.

As in Tetlock (2007), the first measure is based on the idea of negativity bias*,* according to which individuals are affected more strongly by negative rather than positive influences – even when of equal intensity (Rozin and Royzman, 2001). The so-called "Pessimism Indicator" (*PI*) is a measure of pessimism expressed in the news, which relates the number of negative headlines to the overall number of headlines for a given period. It is formally defined as follows:

$$PI_t = \frac{\sum_1^I negative\ headlines_{i,t}}{\sum total\ number\ of\ headlines_t}, \tag{3.5}$$

where i is a headline classified as negative and t is the period in which all headlines must be published to be taken into account.

Similar to Antweiler and Frank (2004), we propose a second sentiment measure capturing optimism (bullishness) in news: an "Optimism Indicator" (*OI*). As a contrary measure to the *PI,* it is defined as the number of positive headlines divided by the overall number of headlines for a given period. More formally:

$$OI_t = \frac{\sum_1^I positive\ headlines_{i,t}}{\sum total\ number\ of\ headlines_t}, \tag{3.6}$$

where i is a headline identified as positive and t the aggregation period.

Both *PI* and *OI* range from 0 to 1, whereby a higher value indicates a greater level of media-expressed pessimism or optimism, respectively. These measures can therefore be interpreted as percentages of pessimism and optimism in the news over the respective time period.

Thirdly, a relative measure is suggested, which accounts for both polarities, positivity as well as negativity expressed in news. The "Sentiment Quotient" (*SQ*) indicates the degree of optimism and pessimism in the news, excluding all neutral headlines. This measure is inspired by *yukkalab*, a company offering commercial sentiment analysis (www.yukkalab.com). The *SQ* is defined as the number of positive headlines in relation to the number of positive and negative headlines for a given period t. If the *SQ* is greater than 0.5, the positive headlines exceed the negative ones, indicating overall optimism in the news, and vice versa. In terms of computation, it can be stated as follows:

$$SQ_t = \frac{\sum_1^I positive\ headlines_{i,t}}{\sum_1^I positive\ headlines_{i,t} + \sum_1^J negative\ headlines_{i,t}}, \tag{3.7}$$

where i is a headline classified as positive, j is a headline identified as negative and t the time span used for aggregation.

Table 3.3 presents the descriptive statistics of all three sentiment measures. Mean, median, standard deviation, minimum and maximum are reported. During our sample period, the *PI* and *OI* range from 0.09 to 0.38 and 0.22 to 0.48, respectively. While the mean of the *PI* is 0.21, it is 0.35 for the OI. The average *SQ* is 0.63, consistently indicating an (on average) higher amount of news classified as positive than such classified as negative by the support vector network.

Table 3.3: Descriptive Statistics: News-Based Sentiment Measures

Statistic	**Mean**	**Median**	**SD**	**Min**	**Max**
PI	0.21	0.20	0.06	0.09	0.38
OI	0.35	0.35	0.06	0.22	0.48
SQ	0.63	0.65	0.09	0.39	0.77

Notes**:** This table reports summary statistics of our monthly sentiment measures. *PI is* the pessimism indicator, *OI* the optimism indicator and *SQ* the sentiment quotient. The sample period is 2005:M1 to 2016:M12.

3.5.3 Vector Autoregression

To formalize the analysis, a vector autoregression framework is employed. Given that vector autoregression does not require any a priori assumptions on existing causalities, this technique offers an effective way to investigate the dynamic relationship between sentiment indicators extracted from newspaper headlines and real estate markets. Furthermore, VARs are more flexible than univariate models and offer a rich structure which allows them to capture more features of the data (Brooks and Tsolacos, 2010).

The simplest form of the well-known standard-form or conventional VAR is a bivariate model comprising of a system of two regression equations, where two endogenous variables (y_{1t} and y_{2t}) are expressed as linear functions of their own and each other's lagged values and error terms:

$$\begin{aligned} y_{1t} = {} & \beta_{10} + \beta_{11}\, y_{1t-1} + \cdots + \beta_{1k}\, y_{1t-k} + \alpha_{11}\, y_{2t-1} + \cdots \\ & + \alpha_{1k}\, y_{2t-k} + u_{1t} \\ y_{2t} = {} & \beta_{20} + \beta_{21}\, y_{2t-1} + \cdots + \beta_{2k}\, y_{2t-k} + \alpha_{21}\, y_{1t-1} + \cdots \\ & + \alpha_{2k}\, y_{1t-k} + u_{2t}, \end{aligned} \tag{3.8}$$

where k is the number of lags and u_{it} a white noise disturbance term with $E(u_{it}) = 0$, ($i = 1,2$), $E(u_{1t}, u_{2t}) = 0$. In our case, y_{1t} are the return of the real estate market in period t, while y_{2t} is either the *PI*, the *OI* or the sentiment quotient for the respective month.

Note that, based on economic theory, further control variables are included in our VAR framework as additional exogenous variables on the right-hand side of equation (3.8). This leads to the final model (3.9) which shows (3.8) in common matrix notation and uses **X** as a matrix of exogenous variables and **B** as a matrix of coefficients:

$$\boldsymbol{y_t} = \boldsymbol{A_0} + \boldsymbol{A_1 y_{t-1}} + \cdots + \boldsymbol{A_k y_{t-k}} + \boldsymbol{BX} + \boldsymbol{u_t}. \tag{3.9}$$

During the regression analysis, components of the VAR are tested using an Augmented Dickey-Fuller Test (ADF) to check for the existence of a unit root. Whenever the null hypothesis and therefore the required stationarity is rejected, variables are differenced once or used as growth rates to ensure statistical appropriateness. Additionally, the optimal lag length has to be determined for a well-specified VAR by making use of an array of selection criteria. Our decision was based mainly on the three most popular ones, the Akaike (AIC), the Bayesian (BIC) and the Hannan-Quinn information criterion (HQIC). All three rest on the notion that including an extra term might increase the goodness of the model, but that the model should be penalized at the same time for the increasing number of parameters one needs to estimate. Whenever the rise in goodness of fit outweighs the penalty term, the information criterion decreases. Accordingly, the lag length which minimizes the value of the information criteria is chosen (Brooks and Tsolacos, 2010). Whenever results are inconclusive, the likelihood ratio test and the final prediction error are utilized to guide the decision on the appropriate lag length.

We further apply the Breusch-Godfrey Lagrange Multiplier test to ensure that the residual series from an estimated model are not serially correlated. Looking for any patterns in the plotted residuals is in some cases difficult to interpret and is therefore only for verification. In addition, several diagnostic tests are performed, for example, residuals are tested for normality and homoscedasticity.

As the main interest of this paper is to investigate whether the created media sentiment measures do indeed have predictive power when explaining returns of the direct and indirect real estate market in the US, for each VAR, Granger causalities are tested and reported. Furthermore, we always state the variance decomposition of forecast errors using a Cholesky factorization.

3.6 Results

A quick recap: our analysis follows the theoretical premise that real estate market participants base their decisions on available information, as well as their own personal beliefs, which are not fully reflected in fundamental economic data. While researchers like Marcato and Nanda (2016) use readily available sentiment indices such as the *Architecture Billings Index* and the *National Association of Homebuilders/Wells Fargo Housing Market Index* to capture an aggregate of individual expectations in non-residential as well as residential real estate markets, respectively, we pursue another direction. Corresponding with Akerlof and Shiller (2010), we argue that "[a]ll of […] processes are driven by stories. The stories that people tell to themselves, about themselves, about how others behave, and even about how the economy as a whole behaves all influence what they do" (p. 173). Thus, our approach makes use of a trained support vector machine to measure market sentiments based on "published" news stories, which arguably bear the potential to influence the decision-making of informed commercial real estate market participants in the United States. As we do not know whether media simply reflects or causes market movements of the direct as well as indirect real estate markets, or whether there is a bi-directional relationship, all the following results aim to shed light on the dynamic as well as temporal dimension between these two possibly linked aspects. The analysis starts by looking at the securitized real estate market and proceeds by comparing the results to the findings from the direct real estate market.

3.6.1 Securitized Real Estate Market

Table 3.4 shows the endogenous dynamics between the *FTSE/NAREIT All Equity REIT Total Return Index* (*NAREIT*) and our three different sentiment indicators, using a VAR framework. All three models control for the same set of macroeconomic variables i.e. term, spread, inflation and the returns of the S&P 500, all models are robust in terms of diagnostic tests and show an optimal lag length of two. Although not shown explicitly in the tables, significant control variables carry the expected sign. The regressions are conducted on a monthly basis, as we are able to benefit from our manually constructed sentiment measures. As long as there are enough news stories provided, our indicators can be computed for any desired period. Thus, when analyzing the securitized real estate market, we are only limited by the frequency at which control variables are available. This differs from the work of other researchers such as Ling *et al.* (2014) and Das *et al.* (2015), in which the frequency of the sentiment measure e.g. the quarterly published *Real Estate Research Corporation* (RERC) survey is the limiting factor.

Table 3.4: VAR Estimation Results: News-Based Sentiment and Securitized Real Estate Market

	FTSE/NAREIT All Equity REIT Total Return Index (NAREIT)					
	Model 1		**Model 2**		**Model 3**	
NAREIT (-1)	-0.168	*	-0.193	**	-0.185	**
	[-1.88658]		[-2.17359]		[-2.09171]	
NAREIT (-2)	-0.200	**	-0.200	**	-0.193	**
	[-2.31786]		[-2.26888]		[-2.21658]	
Pessimism Indicator (-1)	-0.254	**				
	[-2.54932]					
Pessimism Indicator (-2)	-0.056					
	[-0.55530]					
Optimism Indicator (-1)			0.057			
			[0.69979]			
Optimism Indicator (-2)			0.053			
			[0.64730]			
Sentiment Quotient (-1)					0.128	**
					[2.00084]	

(Table continues on the following page.)

Table 3.4: VAR Estimation Results: News-Based Sentiment and Securitized Real Estate Market (continued)

Sentiment Quotient (-2)			0.049
			[0.75974]
Constant	0.005	0.004	0.004
	[1.11839]	[0.89662]	[0.98906]
Macroeconomic variables	**YES**	**YES**	**YES**
Adj. R-squared	0.64	0.62	0.63
F-statistic	16.37	15.21	15.83
Log likelihood	256.82	253.32	255.22
Akaike AIC	-3.40	-3.35	-3.38
Schwarz SC	-3.05	-3.00	-3.02
Granger causality			
Sentiment measure	**0.03**	0.74	0.13
NAREIT	0.54	0.69	0.91

Notes: This table reports results for the estimated VAR models with monthly *NAREIT* returns and news-based sentiment as endogenous variables. The set of macroeconomic control variables includes the difference between Baa- and Aaa-rated corporate bonds yields (*SPREAD*), the difference between the 10-year US Treasury bond and the 3-Month Treasury bill yields (*TERM*), the percentage change of the CPI (*INFL*) and the total return of the S&P 500 Composite Index (*SP500*). For brevity, we only report the results of the real estate return equations for each sentiment indicator. T-statistics are reported in brackets underneath the coefficient estimates. In terms of Granger causality, p-values are reported for both directions. P-values in bold show a significance up to 5%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2005:M4 to 2016:M12.

The regression equations of Models 1 to 3 show the expected statistical significance at the 1st and 2nd lag of the autoregressive *NAREIT* component and similar levels of goodness of fit around 62% to 64%. With regard to sentiment measures, all coefficients have the expected sign. While a rising pessimism indicator negatively affects market returns, the opposite is true for the optimism indicator and sentiment quotient. This corresponds to the way the indicators are created. *OI* and *SQ* facilitate the number of positive headlines, *PI* the number of negative headlines as the numerator. However, only the 1st lag of the *PI* and *SQ* are statistically significant at the 5% level. The optimism indicator has no significant impact at all on market returns. Granger causalities confirm these findings. In contrast to the *OI* of Model 2, the *PI* has predictive power at the 5% level. The sentiment quotient slightly misses the 10% level of significance. Note that for none of the three models *NAREIT* does Granger-cause the sentiment measures. Hence, the sentiment indicators are not significantly affected by past market performance, but provide additional information that is relevant to the securitized real estate market.[24] This indicates a non-existing endogenous dynamic between the securitized real estate market and the sentiment indicators in Model 2 and Model 3 and a one-sided relationship from the *PI* to market returns in Model 1. Variance decomposition figures up to 12 months, using the Cholesky decomposition, yield a contribution of 6.12% for the *PI*, 0.46% for the *OI* and 3.56% for the *SQ*, which is consistent with previous findings.

Overall, based on Table 3.4, the pessimism indicator shows the highest predictive power in explaining the growth of returns in the United States securitized real estate market. This is the case despite the fact that we used the same SNL dataset for all three indicators, as well as an identical trained support vector machine when classifying news items beforehand. A more pronounced market sensitivity to negative news was also found by Tetlock (2007), when analyzing the interactions between media and the general stock market. As his mathematically derived dictionary-based sentiment measure consisted primarily of negatively annotated word categories, he referred to it as pessimism factor. Furthermore, Loughran and McDonald (2011) also focus primarily on negative word lists in their seminal paper.

According to research question 3, the question remains as to whether our sentiment measures and especially the *PI,* retain their predictive power when including other sentiment measures. To check for robustness, and hence include a broad spectrum of other sentiment indicators at the same time, Table 3.5 contrasts the base Model 1 from Table 3.4 with two augmented regression models i.e. Models 4 and 5.[25] Facilitating other available sentiment measures, we run two principal component analyses – one for bearish and one for bullish market indicators – and include the extracted principal components as endogenous variables in our Model 1. This allows us to consider the opinion of individual investors (*AAIIBULL* and *AAIIBEAR*), as well as sentiment

[24] Note that our results do not automatically indicate that market participants are ignoring past market performance in terms of their sentiment about the future or that past market performance is not relevant for our constructed sentiment indicators at all. As our text corpus does not only contain news about past market movements, but also many other possible aspects concerning the real estate industry, past market performance is most likely only one factor driving sentiment indicator changes. Furthermore, different news might incorporate different levels of textual sentiment, are reported at different frequencies and can be forward- or backward-looking. Hence, this heterogeneity might be the reason why our models do not capture a statistically significant relationship between (pure) market performance in the past and future indicator changes.

[25] Note that when using the *OI* and the *SQ* instead of the *PI* in Model 4 and 5 of Table 3.5, the *OI* is still insignificant and the *SQ* drives returns one month ahead similar to the findings of Table 3.4.

expressed by stock market newsletter editors (*ADVSENTBULL*, *ADVSENTBEAR*). At the same time, we include further policy (*ECOPOLUNCERTINEWS*, *ECOPOLUNCERTIOVER*) as well as equity–market-related economic uncertainty (*ECOUNCERT*) – expressed by news coverage, disagreement among economic forecasters and federal tax code provisions – and consumer sentiment (*CONSUESENTI*). Again, all models yield an optimal lag length of 2 months.

Table 3.5: VAR Estimation Results: News-Based Sentiment and Securitized Real Estate Market – Controlling for Other Sentiment Indicators

	FTSE/NAREIT All Equity REIT Total Return Index (NAREIT)					
	Model 1		**Model 4**		**Model 5**	
NAREIT (-1)	-0.168	*	-0.142		-0.140	
	[-1.88658]		[-1.56605]		[-1.55211]	
NAREIT (-2)	-0.200	**	-0.110		-0.124	
	[-2.31786]		[-1.22274]		[-1.39038]	
Pessimism Indicator (-1)	-0.254	**	-0.249	**	-0.250	**
	[-2.54932]		[-2.52610]		[-2.52191]	
Pessimism Indicator (-2)	-0.056		-0.093		-0.081	
	[-0.55530]		[-0.93056]		[-0.80736]	
First component (bearish) (-1)			0.000			
			[0.03499]			
First component (bearish) (-2)			-0.011	**		
			[-2.05542]			
Second component (bearish) (-1)			-0.002			
			[-0.55066]			
Second component (bearish) (-2)			-0.007	*		
			[-1.74007]			
First component (bullish) (-1)					0.000	
					[0.05343]	
First component (bullish) (-2)					0.015	**
					[2.52453]	
Second component (bullish) (-1)					-0.001	
					[-0.33726]	

(Table continues on the following page.)

Table 3.5: VAR Estimation Results: News-Based Sentiment and Securitized Real Estate Market – Controlling for Other Sentiment Indicators (continued)

Second component (bullish) (-2)			-0.004
			[-1.18864]
Constant	0.005	0.005	0.005
	[1.11839]	[1.26585]	[1.20840]
Macroeconomic variables	**YES**	**YES**	**YES**
Adj. R-squared	0.64	0.66	0.65
F-statistic	16.37	14.32	14.12
Log likelihood	256.82	262.77	262.07
Akaike AIC	-3.40	-3.43	-3.42
Schwarz SC	-3.05	-3.00	-2.98
Granger causality (PI ~ NAREIT)			
Pessimism indicator	**0.03**	**0.04**	**0.04**
NAREIT	0.54	0.36	0.38
Granger causality (Sentiment PCA ~ NAREIT)			
First component		0.08	**0.02**
Second component		0.22	0.49
NAREIT on first component		0.26	0.54
NAREIT on second component		0.77	0.76

Notes: This table reports results for the estimated VAR models with monthly *NAREIT* returns, news-based sentiment and further sentiment proxies as endogenous variables. The set of macroeconomic control variables includes the difference between Baa- and Aaa-rated corporate bonds yields (*SPREAD*), the difference between the 10-year US Treasury bond and the 3-Month Treasury bill yields (*TERM*), the percentage change of the CPI (*INFL*) and the total return of the S&P 500 Composite Index (*SP500*). Principal components are constructed as described in the text. For brevity, we only report the results of the real estate return equations for each sentiment measure. T-statistics are reported in brackets underneath the coefficient estimates. In terms of Granger causality, p-values are reported for both directions. P-values in bold show a significance up to 5%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2005:M4 to 2016:M12.

Despite including additional sentiment components, the pessimism indicator retains sign, coefficient size and significance of the 1st lag at the 5% level. Changes in the *PI* still Granger-cause *NAREIT* market returns, while the reverse causation further on cannot be stated. Considering the coefficient estimations of the bearish and bullish sentiment components, one can observe a similar dynamic. Except for the 2nd lag of the bullish component, all 2nd-lag principal components (PCs) are statistically significant at the 10% or 5% levels and show the expected coefficient signs. However, while the first component of the bullish sentiment measure Granger-cause *NAREIT* returns at the 5% level, the results are slightly weaker for the first bearish component, which fails to reach the 5% level. In both cases, the second component does not Granger-cause *NAREIT* returns and *NAREIT* returns do not Granger-cause the sentiment PCs at all.

The variance decomposition figures show a contribution of 3.42% to 4.62% for the *PI,* while the first and second components of the bearish (bullish) indicator range up to 8.66% (7.54%) and 3.20% (2.65%), respectively. Overall, these results confirm that our pessimism indicator has some return-signaling effect in the securitized real estate market in the United States, besides the more general sentiment expressed by the principal components.

It is worth noting that the sentiment indicators constructed via support vector machine usually have a more timely impact on *NAREIT* returns than the general sentiment components. Usually, the 1st lag of the *PI* is the significant one, as opposed to the second of the sentiment PCs in Models 4 and 5. Provided one can adopt the presumption that investors require some time to gather information and subsequently form their own personal beliefs about the market, one could argue that this is induced by the temporal nature of the perception-building process. As survey-based indicators aggregate sentiment from market participants which should be at least partly influenced by news items, our news-based sentiment measures are positioned one step ahead, directly capturing sentiment from the information source. Thus, they should have a more timely impact on market returns. This theory would also explain why the news specific *PI*, as well as the general sentiment principal components, have predictive power on *NAREIT* returns in the same model. The PCs not only incorporate sentiment from news items, but also from other sources such as the abovementioned federal tax code provisions, which differentiates them from our purely news-based sentiment indicators.

3.6.2 Direct Real Estate Market

This and the following paragraphs repeat the entire process for the direct real estate market, further assessing the predictive power and robustness of our sentiment indicators according to research question 2. Thus, the VAR framework of Table 3.6 analyses the potentially endogenous relations between the three machine-learning sentiment indicators and the *CoStar Commercial Repeat Sales Index* (*CCRSI*), as a measure of direct market performance. Once again, all models control for economic default risk, expectations about future economic and labor market developments, as well as real estate supply, by including spread, term, initial unemployment claims and construction-spending variables.[26] Significant controls show the expected sign although left out in the tables for the sake of brevity. The analysis uses an optimal lag length of 8 months following the joint recommendations of several lag-length indicators such as Akaike, Schwarz and Hannan-Quinn information criteria, final prediction error as well as the sequential modified LR-test statistic. For ease of reading, sentiment measure means pessimism indicator in Model 6, optimism indicator in Model 7 and sentiment quotient in Model 8. Again, Table 3.6 states Granger causalities for both directions at the bottom of each column.

Table 3.6: VAR Estimation Results: News-Based Sentiment and Direct Real Estate Market

	CoStar Commercial Repeat-Sales Index (CCRSI)					
	Model 6		**Model 7**		**Model 8**	
	Pessimism Indicator		*Optimism Indicator*		*Sentiment Quotient*	
CCRSI (-1)	1.081	***	1.126	***	1.097	***
	[12.2066]		[12.2800]		[12.1386]	
CCRSI (-2)	-0.071		-0.097		-0.116	
	[-0.62895]		[-0.79707]		[-0.56192]	
CCRSI (-3)	-1.072	***	-1.069	***	-1.108	***
	[-10.0662]		[-9.22729]		[-9.98168]	
CCRSI (-4)	1.304	***	1.305	***	1.307	***
	[9.49656]		[8.90372]		[9.30795]	
CCRSI (-5)	-0.364	**	-0.313	**	-0.320	**
	[-2.63687]		[-2.10608]		[-2.25577]	
CCRSI (-6)	-0.494	***	-0.535	***	-0.549	***
	[-4.68369]		[-4.55049]		[-4.97542]	
CCRSI (-7)	0.831	***	0.818	***	0.840	***
	[7.42224]		[6.69008]		[7.36439]	
CCRSI (-8)	-0.395	***	-0.397	***	-0.386	***
	[-4.76615]		[-4.50364]		[-4.58881]	

(Table continues on the following page.)

[26] Note that when further controlling for lagged returns of the securitized real estate market, the findings of Table 3.6 do not change.

Table 3.6: VAR Estimation Results: News-Based Sentiment and Direct Real Estate Market (continued)

Sentiment measure (-1)	-0.026		-0.011		0.004	
	[-1.29342]		[-0.57983]		[0.29472]	
Sentiment measure (-2)	-0.060	**	0.014		0.028	
	[-2.32027]		[0.62850]		[1.59966]	
Sentiment measure (-3)	-0.087	***	0.019		0.045	**
	[-3.00079]		[0.83651]		[2.29096]	
Sentiment measure (-4)	-0.031		0.016		0.024	
	[-1.03654]		[0.72929]		[1.16132]	
Sentiment measure (-5)	0.010		-0.016		-0.005	
	[0.33006]		[-0.75305]		[-0.25187]	
Sentiment measure (-6)	0.038		-0.011		-0.008	
	[1.34760]		[-0.49419]		[-0.40430]	
Sentiment measure (-7)	-0.006		-0.004		0.008	
	[-0.23158]		[-0.17419]		[0.46902]	
Sentiment measure (-8)	-0.049	**	0.018		0.040	***
	[-2.39110]		[1.09460]		[2.82321]	
Constant	0.001		0.000		0.001	
	[0.85106]		[0.62156]		[0.78543]	
Macroeconomic variables	**YES**		**YES**		**YES**	
Adj. R-squared	0.81		0.78		0.81	
F-statistic	21.88		18.13		21.04	
Log likelihood	494.88		484.22		492.64	
Akaike AIC	-6.90		-6.74		-6.87	
Schwarz SC	-6.28		-6.12		-6.24	
Granger causality						
Sentiment measure	**0.00**		0.65		**0.01**	
CCRSI	0.99		0.74		0.92	

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 Baa- and Aaa-rated corporate bond yields (*SPREAD*), the difference between the 10-year US Treasury bond and 3-Month Treasury bill yields (*TERM)*, the amount of unemployment initial claims (*UNEMPL*) and the amount of construction spending (*CONSTR*). For the sake of brevity, we only report the results of the real estate return equations for each sentiment measure. T-statistics are reported in brackets underneath the coefficient estimates. In terms of Granger causality, values are reported for both directions. P-values in bold show a level of significance up to 5%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2005:M10 to 2016:M12.

Models 6 to 8 show a very pronounced autoregressive component; except for the 2nd lag, all other lagged values of the *CCRSI* are highly significant when explaining future market returns. Considering the pronounced cyclical behavior of the *CoStar Index* over the observation period with a boom phase until 2007, the bust of 2008/2009 and subsequent market recovery, this has to be expected. In terms of sentiment measures, Table 3.4 and Table 3.6 yield similar results for the indirect and direct commercial real estate markets. Once more, *PI* and *SQ* show the expected sign of significant lags, while the *OI* does not significantly predict direct market returns. However, the *CCRSI* reacts later to the sentiment indicators than the *NAREIT*. While the 1st lag appeared to be relevant in the REIT market, the 2nd, 3rd and – in terms of magnitude less pronounced – the 8th lag are now the three important ones. Although real estate is deemed to be slow, there is no intuitive explanation why the 8th lag is significant but not the 4th to 7th ones. Presumably, this is a sample effect.

Overall, the pessimism indicator predicts the direct real estate market best. Its changes Granger-cause market returns at the 1% level of significance. However, in contrast to previous results, the sentiment quotient now reaches similar levels of predictive power. This can also be seen when comparing the goodness of fit measures for Models 6 and 8 that are very similar in terms of magnitude. The variance decomposition up to 36 months corroborates these findings, as the *PI's*, *OI's* and *SQ's* contribution to forecast errors reach 20.94%, 3.50% and 15.47%, respectively.

Again, with a non-significant *OI*, one could argue that there is some evidence of an existing negativity bias of market participants. Nevertheless, the results in the direct real estate market are slightly less pronounced than in the securitized one. Note that *CoStar* returns do not Granger-cause any of the three sentiment indicators in Table 3.6. All existing endogenous relationships extend from changes in the indicators to market returns and not vice versa, or in a bi-directional manner. Hence, the indicators are again able to extract additional information from news that is relevant in explaining direct market movements.

Table 3.7 depicts the relative performance of our sentiment indicators created via machine-learning, in contrast to other more general sentiment measures. Models 9 and 10 augment Model 6 of Table 3.6 with the same first and second bullish and bearish components of the principal component analysis. Because the optimal lag length remains 8 months, we refrain from an extended VAR approach and incorporate the components only as additional exogenous controls. This is because the addition as endogenous variables would lead to a massive loss of degrees of freedom, due to two additional equations and two additional variables with 8 lags each, for which coefficients have to be estimated. Although still significantly explaining direct markets returns with the 2nd, 3rd and 8th lag, the results of Models 9 and 10 are slightly weaker in terms of significance, as well as coefficient magnitude in comparison to Model 6. Once again, a reverse causation cannot be stated.[27]

The variance decomposition shows a contribution of the *PI* up to 14.66% (19.63%) in the case of Model 9 (10). This leads us to the conclusion that there is indeed evidence of the pessimism indicator's return-signaling effect not only for the indirect but also for the direct real estate market.

Table 3.7: VAR Estimation Results: News-Based Sentiment and Direct Real Estate Market – Controlling for Other Sentiment Indicators

	CoStar Commercial Repeat-Sales Index (CCRSI)					
	Model 6		**Model 9**		**Model 10**	
	Pessimism Indicator		Sentiment Indices (bearish)		Sentiment Indices (bullish)	
CCRSI (-1)	1.081	***	1.103	***	1.120	***
	[12.2066]		[12.2347]		[12.5269]	
CCRSI (-2)	-0.071		-0.099		-0.091	
	[-0.62895]		[-0.87003]		[-0.81061]	
CCRSI (-3)	-1.072	***	-1.041	***	-1.062	***
	[-10.0662]		[-9.58587]		[-10.2399]	
CCRSI (-4)	1.304	***	1.302	***	1.298	***
	[9.49656]		[9.25507]		[9.47375]	
CCRSI (-5)	-0.364	**	-0.378	***	-0.369	***
	[-2.63687]		[-2.72855]		[-2.76909]	
CCRSI (-6)	-0.494	***	-0.468	***	-0.468	***
	[-4.68369]		[-4.38024]		[-4.56236]	
CCRSI (-7)	0.831	***	0.828	***	0.835	***
	[7.42224]		[7.24651]		[7.69898]	

(Table continues on the following page.)

[27] When substituting the *PI* by the *OI* or *SQ* in Table 3.7, the results of Table 3.6 with respect to the respective significance of the *OI* and *SQ* still hold.

Table 3.7: VAR Estimation Results: News-Based Sentiment and Direct Real Estate Market – Controlling for Other Sentiment Indicators (continued)

CCRSI (-8)	-0.395	***	-0.426	***	-0.428	***
	[-4.76615]		[-5.04040]		[-5.34518]	
Sentiment measure (-1)	-0.026		-0.025		-0.019	
	[-1.29342]		[-1.20390]		[-0.88914]	
Sentiment measure (-2)	-0.060	**	-0.058	**	-0.055	**
	[-2.32027]		[-2.20412]		[-2.14026]	
Sentiment measure (-3)	-0.087	***	-0.063	**	-0.057	*
	[-3.00079]		[-2.05561]		[-1.95132]	
Sentiment measure (-4)	-0.031		-0.006		-0.001	
	[-1.03654]		[-0.19909]		[-0.03824]	
Sentiment measure (-5)	0.010		0.021		0.028	
	[0.33006]		[0.69597]		[0.97256]	
Sentiment measure (-6)	0.038		0.032		0.038	
	[1.34760]		[1.10973]		[1.37261]	
Sentiment measure (-7)	-0.006		-0.007		-0.003	
	[-0.23158]		[-0.27399]		[-0.13043]	
Sentiment measure (-8)	-0.049	**	-0.039	*	-0.036	*
	[-2.39110]		[-1.81901]		[-1.79544]	
First component (bearish)			0.001			
			[1.46930]			
First component (bearish) (-1)			0.001			
			[1.15635]			
First component (bearish) (-2)			-0.001			
			[-1.63152]			
Second component (bearish)			-0.001			
			[-0.84100]			
Second component (bearish) (-1)			-0.001			
			[-0.84360]			
Second component (bearish) (-2)			-0.001			
			[-1.43150]			
First component (bullish)					-0.001	*
					[-1.76414]	
First component (bullish) (-1)					-0.001	
					[-1.43729]	
First component (bullish) (-2)					0.002	**
					[2.08032]	
Second component (bullish)					-0.001	
					[-1.67842]	
Second component (bullish) (-1)					-0.001	*
					[-0.77426]	

(Table continues on the following page.)

Table 3.7: VAR Estimation Results: News-Based Sentiment and Direct Real Estate Market – Controlling for Other Sentiment Indicators (continued)

Second component (bullish) (-2)			-0.001 *
			[-1.85649]
Constant	0.001	0.001	0.001
	[0.85106]	[1.06344]	[1.08778]
Macroeconomic variables	**YES**	**YES**	**YES**
Adj. R-squared	0.81	0.82	0.83
F-statistic	21.88	18.80	20.29
Log likelihood	494.88	500.72	505.20
Akaike AIC	-6.90	-6.90	-6.97
Schwarz SC	-6.28	-6.15	-6.21
Granger causality			
Pessimism indicator	0.00	0.07	0.03
CCRSI	0.99	0.99	1.00

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 Baa- and Aaa-rated corporate bonds yields (*SPREAD*), the difference between the 10-year US Treasury bond and the 3-Month Treasury bill yields (*TERM)*, the amount of unemployment initial claims (*UNEMPL*), the amount of construction spending (*CONSTR*) and further sentiment proxies per PCA. Principal components are constructed as described in the text. For brevity, we only report the results of the real estate return equations for each sentiment measure. T-statistics are reported in brackets underneath the coefficient estimates. In terms of Granger causality, values are reported for both directions. P-values in bold show a significance up to 5%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2005:M10 to 2016:M12.

3.6.3 Synopsis

Based on the notion of the general importance of news for the decision-making process of market participants, our research aimed to investigate the potential of sentiment indicators created via machine-learning and a dataset of news items. Research questions 1 and 2 deal with whether the readily constructed sentiment indicators are able to predict direct and indirect commercial real estate market returns and whether there are differences with respect to the markets. Our results indeed indicate predictive power for both markets, and the results are comparable with respect to the quality of individual sentiment measures. Furthermore, for neither of the two markets a reverse causation could be found.

However, the results deviate in market reaction times to changes in the sentiment indicators. During the 12-year sample period, returns in the securitized market respond to news-based sentiment one month earlier than *CoStar* returns. This might be the case because of the typical characteristics of the two markets; the direct real estate market is known to move slower than the indirect one. The main reason for the time difference is presumably the time-consuming transaction process in the direct real estate market with considerable time exposure to search execution, due diligence, financing, negotiation, clarification of tax and legal conditions and other administrative necessities. Until the change of ownership is actually executed, months can pass by, which allows all contract parties to create their own opinion about future market movements which are normally priced in their respective deal offers. The indirect market is able to execute transactions a lot quicker as investors can sell their shares immediately once their sentiment has changed.

In Table 3.4 and Table 3.6, not all sentiment indicators have the same prediction potential. While the optimism indicator – concentrating on positive news – showed no explanatory power, the *SQ* and *PI* measures – based on positive and negative news or negative news only – were both successful in explaining market movements. The *PI* worked very well for both markets, the *SQ* better in the direct than in the securitized one. Overall, this might be interpreted as evidence of an existing negativity bias of the market. At this point it is also worth mentioning, that although the *PI* is based on the idea of a negativity bias, this theory is just one possible explanation for the results found. Looking at the greater significance of negative news, as well as their timely delayed impact, theories such as loss-aversion and anchoring behavior of market participants can be applicable as well. Evidence from commercial real estate pricing suggests a variation of loss-aversion during the market cycle i.e. an increase in turn-around periods and a weakening during downturns (Bokhari and Geltner, 2011). Loss-averse sellers especially try to hold out when markets are turning downwards which are presumably also periods with an increased negative news coverage. This provides us with a complementary explanation to the idea of a negativity bias when considering the delayed impact and greater influence of the *PI*.

Additionally, the *PI* retained its impact and significance when controlling for other more general sentiment measures in both markets. Even more so, *NAREIT* returns reacted earlier to changes in our sentiment indicator, in contrast to changes in more general sentiment, further showing the capability of “new” sentiment measures created via textual analysis and a machine-learning approach. A possible explanation of this phenomenon could be that the created sentiment measures are more sector- or real-estate-specific than the alternative indicators used in the regression models of Table 3.4 and Table 3.6. By implementing both kinds of measures, the models presumably capture not only the sentiment of market participants with respect to the broader macroeconomic environment, but additionally sentiment with respect to real estate. Furthermore, one could argue that real-estate specific sentiment should affect the markets immediately with short notice, while macroeconomic sentiment requires more considerations by market participants with respect to the respective influence on real estate performance. This could arguably explain the varying timely impact of both kind of sentiment measures.

3.7 Conclusion

Due to the specific characteristics of real estate markets such as low transparency, information asymmetry, illiquidity as well as long transaction periods, one could argue that real estate is by nature more prone to sentiment than stock markets, for example. A number of articles have indeed dealt with the role of market sentiment measured by different proxies and found evidence of significance for real estate asset pricing. One area of research extracts sentiment by investigating text corpora. However, for real estate, related research focuses mainly on a dictionary-based approach. The ongoing digitalization of news and technical advances enables us to contribute to the literature on text-based sentiment analysis in the realm of real estate, by creating and testing sentiment measures constructed via a machine-learning approach. Hence, this paper examines the relationship between news-based sentiment, captured through support vector networks, and the US securitized and direct commercial real estate markets.

In order to extract sentiment from about 54,500 news items, provided by *S&P Global Market Intelligence* Platform (SNL), we train a support vector machine as a classification algorithm. Subsequently, the classification scores thus gained are aggregated into three different monthly sentiment measures, i.e. a pessimism and optimism indicator, as well as a “neutral” sentiment quotient. Applying a VAR framework and monthly real estate return data provided by *NAREIT* and *CoStar*, we analyze the dynamic relationship between our created sentiment measures and direct as well as securitized market returns.

The results indeed show evidence of a significant relationship between our sentiment indicators and real estate market movements. More precisely, the *PI* Granger-causes *NAREIT* returns and leads the market by one month, even when controlling for macroeconomic fundamentals. Furthermore, the text-based indicator provides information in explaining securitized market returns beyond more general market sentiment. Our results do not indicate a significant influence of past market performance on any of the three constructed sentiment measures. The direct real estate market yields similar findings. The pessimism indicator, as well as the sentiment quotient, drive total returns by 2, 3 and (8) months. For both measures, Granger causality remains significant when including macroeconomic and general sentiment controls. In equal measure to the REIT market, there is no bi-directional relationship. Overall, the findings are consistent with the notion of a slower-moving direct market, in contrast to the securitized one. These results highlight the importance of real estate news analytics as an innovative source of sentiment and indicate that news-based sentiment can be deployed as a leading market indicator.

Looking at the text-based sentiment indicators themselves, they are not only directly linked to real estate but also by construction more directly linked to market sentiment than indirect indicators such as mortgage fund flows. This means, they sit in-between these two types of sentiment indicators allowing them to combine the benefits of both. They are easier and faster to compute and directly related to the asset class. As we show the successful application of a sentiment-measuring method that also allows short and flexible aggregation periods, we contribute to real estate research and to industry participants as well. On the one hand, the methodologies explored and the results found might help to improve and explain real estate decision-making processes, for example by enhancing forecasting models to anticipate future market movements. On the other hand, a better understanding of past events is equally important. Looking at firm level, companies can use the applied methodologies in order to gain insights about market sentiment prevailing after reporting company news, publishing a new strategy or releasing a new product. Thus, understanding the extraction of sentiment from textual documents provides market participants and researchers with a flexible and adjustable tool that is both directly related to the asset class and quicker as well as easier to replicate as e.g. topic-specific surveys.

However, in order to create sentiment indicators for even smaller aggregation periods, a more extensive news dataset than the one we used would be required. Future research could therefore combine professional news with other sources such as news directed to the public, for example news from *The Wall Street Journal* or the *Financial Times*. Nevertheless, at higher frequency levels, efficiently controlling for macroeconomic fundamentals becomes progressively more complicated. In addition, a comparison to the established dictionary-based approach would be worthwhile. Due to different levels of transparency in other real estate markets, one could expect sentiment to be even more relevant in countries with a less advanced real estate industry, an issue that is also worth investigating.

3.8 References

Akerlof, G. A. and Shiller, R. J. (2010), *Animal Spirits - How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*, Princeton University Press, Princeton, NJ.

Antweiler, W. and Frank, M. Z. (2004), “Is All That Talk Just Noise? - The Information Content of Internet Stock Message Boards”, *The Journal of Finance*, Vol. 59 No. 3, pp. 1259–1294.

Barkham, R. J. and Ward, C. W.R. (1999), “Investor Sentiment and Noise Traders: Discount to Net Asset Value in Listed Property Companies in the U.K.”, *Journal of Real Estate Research*, Vol. 18 No. 2, 291–312.

Black, F. (1986), “Noise”, *The Journal of Finance*, Vol. 41 No. 3, pp. 528–543.

Bokhari, S. and Geltner, D. (2011), “Loss Aversion and Anchoring in Commercial Real Estate Pricing: Empirical Evidence and Price Index Implications”, *Real Estate Economics*, Vol. 39 No. 4, pp. 635–670.

Bollen, J., Mao, H. and Zeng, X. (2011), “Twitter Mood Predicts the Stock Market”, *Journal of Computational Science*, Vol. 2 No. 1, pp. 1–8.

Brooks, C. and Tsolacos, S. (1999), “The Impact of Economic and Financial Factors on UK Property Performance”, *Journal of Property Research*, Vol. 16 No. 2, pp. 139–152.

Brooks, C. and Tsolacos, S. (2010), *Real Estate Modelling and Forecasting*, Cambridge University Press, New York, NY.

Brown, G. W. and Cliff, M. T. (2004), “Investor Sentiment and the Near-Term Stock Market”, *Journal of Empirical Finance*, Vol. 11 No. 1, pp. 1–27.

Chatrath, A., Miao, H., Ramchander, S. and Villupuram, S. (2014), “Currency Jumps, Cojumps and the Role of Macro News”, *Journal of International Money and Finance*, Vol. 40, pp. 42–62.

Clayton, J., Ling, D. C. and Naranjo, A. (2009), "Commercial Real Estate Valuation - Fundamentals Versus Investor Sentiment", *The Journal of Real Estate Finance and Economics*, Vol. 38 No. 1, pp. 5–37.

Cortes, C. and Vapnik, V. (1995), "Support-Vector Networks", *Machine Learning*, Vol. 20 No. 3, pp. 273–297.

Crosby, N., Lizieri, C. and McAllister, P. (2010), "Means, Motive and Opportunity? Disentangling Client Influence on Performance Measurement Appraisals", *Journal of Property Research*, Vol. 27 No. 2, pp. 181–201.

Das, P. K., Freybote, J. and Marcato, G. (2015), "An Investigation into Sentiment-Induced Institutional Trading Behavior and Asset Pricing in the REIT Market", *The Journal of Real Estate Finance and Economics*, Vol. 51 No. 2, pp. 160–189.

Das, S. R. and Chen, M. Y. (2007), "Yahoo! - For Amazon: Sentiment Extraction from Small Talk on the Web", *Management Science*, Vol. 53 No. 9, pp. 1375–1388.

Dietzel, M. A., Braun, N. and Schäfers, W. (2014), "Sentiment-Based Commercial Real Estate Forecasting with Google Search Volume Data", *Journal of Property Investment & Finance*, Vol. 32 No. 6, pp. 540–569.

Doran, J. S., Peterson, D. R. and Price, S. M. (2012), "Earnings Conference Call Content and Stock Price - The Case of REITs", *The Journal of Real Estate Finance and Economics*, Vol. 45 No. 2, pp. 402–434.

Ferris, S. P., Hao, Q. and Liao, M.-Y. (2013), "The Effect of Issuer Conservatism on IPO Pricing and Performance", *Review of Finance*, Vol. 17 No. 3, pp. 993–1027.

Freybote, J. and Seagraves, P. A. (2017), "Heterogeneous Investor Sentiment and Institutional Real Estate Investments", *Real Estate Economics*, Vol. 45 No. 1, pp. 154–176.

Gallimore, P. and Gray, A. (2002), “The Role of Investor Sentiment in Property Investment Decisions”, *Journal of Property Research*, Vol. 19 No. 2, pp. 111–120.

Hanna, A. J., Tuner, J. D. and Walker, C. B. (2017), “News Media and Investor Sentiment over the Long Run”, QUCEH Working Paper Series No. 2017-06, Queen's University Centre for Economic History (QUCEH), Queen's University Belfast, Belfast, 2017.

Henry, E. (2008), “Are Investors Influenced By How Earnings Press Releases Are Written?”, *Journal of Business Communication*, Vol. 45 No. 4, pp. 363–407.

Henry, E. and Leone, A. J. (2016), “Measuring Qualitative Information in Capital Markets Research. Comparison of Alternative Methodologies to Measure Disclosure Tone”, *The Accounting Review*, Vol. 91 No. 1, pp. 153–178.

Hoesli, M., Lizieri, C. and MacGregor, B. (2008), “The Inflation Hedging Characteristics of US and UK Investments: A Multi-Factor Error Correction Approach”, *The Journal of Real Estate Finance and Economics*, Vol. 36 No. 2, pp. 183–206.

Hohenstatt, R., Käsbauer, M. and Schäfers, W. (2011), “'Geco' and its Potential for Real Estate Research - Evidence from the US Housing Market”, *Journal of Real Estate Research*, Vol. 33 No. 4, pp. 471–506.

Jin, F., Self, N., Saraf, P., Butler, P., Wang, W. and Ramakrishnan, N. (2013), “Forex-Foreteller”, in Ding, Y. (Ed.), *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM Press, New York, NY, pp. 1470–1473.

Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T. and Ngo, D. C. L. (2014), “Text Mining for Market Prediction - A Systematic Review”, *Expert Systems with Applications*, Vol. 41 No. 16, pp. 7653–7670.

Kothari, S. P., Li, X. and Short, J. E. (2009), “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*, Vol. 84 No. 5, pp. 1639–1670.

Li, F. (2006), “Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports?”, Working Paper, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, 2006.

Li, F. (2010), “The Information Content of Forward-Looking Statements in Corporate Filings - A Naïve Bayesian Machine Learning Approach”, Vol. 48 No. 5, pp. 1049–1102.

Lin, C. Y., Rahman, H. and Yung, K. (2009), “Investor Sentiment and REIT Returns”, *The Journal of Real Estate Finance and Economics*, Vol. 39 No. 4, pp. 450–471.

Ling, D. C., Naranjo, A. and Scheick, B. (2014), “Investor Sentiment, Limits to Arbitrage and Private Market Returns”, *Real Estate Economics*, Vol. 42 No. 3, pp. 531–577.

Liu, B. (2012), “Sentiment Analysis and Opinion Mining”, *Synthesis Lectures on Human Language Technologies*, Vol. 5 No. 1, pp. 1–167.

Loughran, T. and McDonald, B. (2011), “When Is a Liability Not a Liability? - Textual Analysis, Dictionaries, and 10-Ks”, *The Journal of Finance*, Vol. 66 No. 1, pp. 35–65.

Marcato, G. and Nanda, A. (2016), “Information Content and Forecasting Ability of Sentiment Indicators: Case of Real Estate Market”, *Journal of Real Estate Research*, Vol. 38 No. 2, 165-2013.

Nowak, A. and Smith, P. (2017), “Textual Analysis in Real Estate”, *Journal of Applied Econometrics*, Vol. 32 No. 4, pp. 896–918.

Ozik, G. and Sadka, R. (2012), “Media and Investment Management”, *SSRN Electronic Journal*.

Pang, B., Lee, L. and Vaithyanathan, S. (2002), "Thumbs up? Sentiment Classification using Machine Learning Techniques", in Unknown (Ed.), *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - EMNLP '02*, Association for Computational Linguistics, Morristown, NJ, pp. 79–86.

Peramunetilleke, D. and Wong, R. K. (2002), "Currency Exchange Rate Forecasting from News Headlines", in Zhou, X. (Ed.), *Proceedings of the 13th Australasian Database Conference*, Australian Computer Society, Sydney, pp. 131–139.

Qiu, L. X. and Welch, I. (2006), "Investor Sentiment Measures", NBER Working Paper No. 10794, National Bureau of Economic Research, Cambridge, MA, 2006.

Rachlin, G., Last, M., Alberg, D. and Kandel, A. (2007), "ADMIRAL - A Data Mining Based Financial Trading System", in Unknown (Ed.), *IEEE Symposium on Computational Intelligence and Data Mining, 2007*, IEEE Service Center, Piscataway, NJ, pp. 720–725.

Rochdi, K. and Dietzel, M. (2015), "Outperforming the Benchmark - Online Information Demand and REIT Market Performance", *Journal of Property Investment & Finance*, Vol. 33 No. 2, pp. 169–195.

Rogers, J. L., van Buskirk, A. and Zechman, S. L. C. (2011), "Disclosure Tone and Shareholder Litigation", *The Accounting Review*, Vol. 86 No. 6, pp. 2155–2183.

Rozin, P. and Royzman, E. B. (2001), "Negativity Bias, Negativity Dominance, and Contagion", *Personality and Social Psychology Review*, Vol. 5 No. 4, pp. 296–320.

Schätz, A. and Sebastian, S. P. (2010), "Real Estate Equities - Real Estate or Equities?", EPRA Research Paper, European Public Real Estate Association, Brussels, 2010.

Schumaker, R. P. and Chen, H. (2009), "Textual Analysis of Stock Market Prediction Using Breaking Financial News", *ACM Transactions on Information Systems*, Vol. 27 No. 2, pp. 1–19.

Schumaker, R. P., Zhang, Y., Huang, C.-N. and Chen, H. (2012), "Evaluating Sentiment in Financial News Articles", *Decision Support Systems*, Vol. 53 No. 3, pp. 458–464.

Sinha, N. R. (2016), "Underreaction to News in the US Stock Market", *Quarterly Journal of Finance*, Vol. 06, pp. 1–46.

Soo, C. K. (2015), "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.

Tetlock, P. C. (2007), "Giving Content to Investor Sentiment - The Role of Media in the Stock Market", *The Journal of Finance*, Vol. 62 No. 3, pp. 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M. and Macskassy, S. (2008), "More Than Words - Quantifying Language to Measure Firms' Fundamentals", *The Journal of Finance*, Vol. 63 No. 3, pp. 1437–1467.

Twedt, B. and Rees, L. (2012), "Reading Between the Lines - An Empirical Examination of Qualitative Attributes of Financial Analysts' Reports", *Journal of Accounting and Public Policy*, Vol. 31 No. 1, pp. 1–21.

Walker, C. B. (2014), "Housing Booms and Media Coverage", *Applied Economics*, Vol. 46 No. 32, pp. 3954–3967.

Walker, C. B. (2016), "The Direction of Media Influence - Real-Estate News and the Stock Market", *Journal of Behavioral and Experimental Finance*, Vol. 10, pp. 20–31.

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

4.1 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-labelled 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

4.2 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 *et al.*, 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-labelled text documents.[28] 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 labelled data shortage by utilizing a new source of distant-labelled 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.[29]

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.[30] 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. Bianchi and Guidolin, 2014; Lizieri *et al.*, 1998). 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

[28] To gradually improve a deep-learning algorithm's capabilities, permanent human intervention is required.
[29] A publication assessing a potential link of the constructed sentiment indicator to direct real estate market liquidity is intended by the authors.
[30] 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.

are performed for this purpose. Besides being required in 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 4.3, research with respect to text-based sentiment in finance and real estate is re-considered as an introduction to the more theory-driven Sections 4.4 and 4.5. 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 4.6 presents the results of the VAR, Markov-switching and logit procedure. Section 4.7 discusses implications and provides concluding remarks.

4.3 Literature Review

4.3.1 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 *et al.* (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, p. 5). 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. Boudoukh *et al.*, 2013; Engelberg *et al.*, 2012; Ferguson *et al.*, 2015; García, 2013; Groß-Klußmann and Hautsch, 2011; Hanna *et al.*, 2017; Heston and Sinha, 2016; Ozik and Sadka, 2012; Sinha, 2016; Sun *et al.*, 2016, as well as the seminal articles by Tetlock, 2007 and Tetlock *et al.*, 2008). Another stream of literature addresses the influence of earnings press releases on a broad variety of performance measures (see e.g. Davis *et al.*, 2015; Davis and Tama-Sweet, 2012; Henry, 2008; Henry and Leone, 2016; Huang *et al.*, 2014) and annual reports (see e.g. Feldman *et al.*, 2010; Jegadeesh and Wu, 2013; Kothari *et al.*, 2009; Li, 2010; Loughran and McDonald, 2011, 2015).

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. Henry and Leone, 2016; Davis *et al.*, 2012; Davis and Tama-Sweet, 2012 for an application of the program *DICTON*). 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*.

4.3.2 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 *et al.*, 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 *et al.*, 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 *et al.*, 2009), private market returns (Ling *et al.*, 2014), trading behavior (Das *et al.*, 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 *et al.* (2009), several scholars drew on *Google* search query data to analyze various aspects of the real estate market in the United States. Hohenstatt *et al.* (2011) provide evidence that *Google Trends*[31] 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 *et al.* (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 aforementioned *DICTON* software to investigate the relationship between the UK

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

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 *et al.* (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 *et al.* (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-labelled 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-labelling in real estate research.

4.4 Data

The outlined study utilizes four types of data: *Seeking Alpha*[32] (SA) long and short idea sections (as explained below) serve as the training dataset 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*

[32] *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 favorable 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.

4.4.1 Seeking Alpha

For the construction of the sentiment index, a two-part process is proposed. As this paper refrains from manually labelling training data for the ANN, a dataset of distant supervision-labelled text documents[33] is required. Summary sections of *Seeking Alpha* long and short ideas are collected for this objective. The following example from the dataset 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 dataset 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 long 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 dataset. The final training dataset consists of a balanced sample of 17,822 SA texts provided by 3,107 different authors and containing an average of 381 characters.

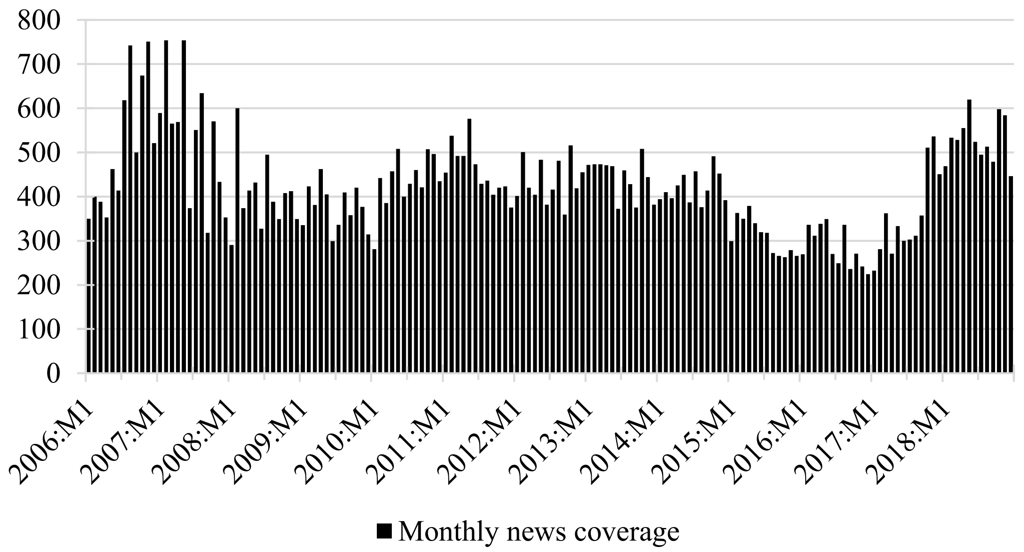
4.4.2 S&P News Database

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*

[33] Distant supervision-labelling is defined as the absence of an annotator providing the classification of the training data manually.

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 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 4.1).

Figure 4.1: S&P News Distribution over Study Period



Notes: Figure 4.1 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.

4.4.3 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 *et al.* (2009), Ling *et al.* (2014) and Hausler *et al.* (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 *et al.*, 2009). Following Das *et al.* (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 *et al.*, 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 4.1.

Table 4.1: Descriptive Statistics

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

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 to 2018:M12.

4.5 Methodology

4.5.1 Artificial Neural Network

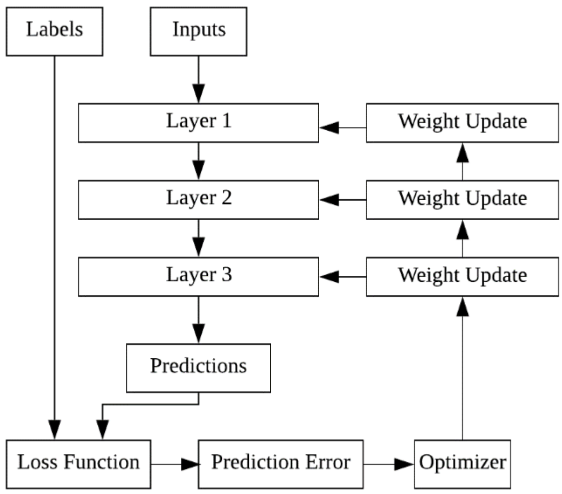
Artificial neural network research, often falsely perceived as a young field, actually emerged as early as the 1950s, with Rosenblatt (1958) often being considered the inventor of the first "real" ANN. Due to the extensive computational requirements and lack of mathematical algorithms to back the concepts, 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 datasets. 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 *et al.*, 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.[34] 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

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

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 4.2.

Figure 4.2: Basic Structure of an Artificial Neural Network



Notes: Figure 4.2 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.

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.[35] 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

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

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 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.[36] 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.[37] 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.

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.[38]

[36] 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.
[37] *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.
[38] A document feature matrix, also referred to as a sparse matrix, contains a column for each word in the respective dataset and a row for each text document in the dataset. 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.

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:[39]

$$max(0, dot(Input, W) + b). \tag{4.1}$$

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:

$$\frac{1}{1+e^{-t}} \; with\; t = dot(Input, W) + b. \tag{4.2}$$

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.[40] The optimization process thus contains a total of 174 gradient updates.[41]

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

$$\frac{1}{n}\sum_{k=1}^{n} -1\big(y_k log(p_k) + (1 - y_k) log(1 - p_k)\big). \tag{4.3}$$

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

[40] 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.

[41] Updates per epoch: 29 (≈14,258/500); Updates over all epochs: 174 (=29*6).

y_k is a binary variable taking the value 1 if *Seeking Alpha* text k is labelled as a long idea, and 0 if *Seeking Alpha* text k is labelled 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).[42] The updates for all parameters W and b are calculated with the following equations:

$$
\begin{aligned}
v_{dW_t} &= \beta v_{dW_{t-1}} + (1-\beta)(dW_t)^2 \\
v_{db_t} &= \beta v_{db_{t-1}} + (1-\beta)(db_t)^2 \\
W_{t+1} &= W_t - \frac{\eta}{\sqrt{v_{dW_t} + \varepsilon}}(dW_t) \\
b_{t+1} &= b_t - \frac{\eta}{\sqrt{v_{db_t} + \varepsilon}}(db_t).
\end{aligned}
\tag{4.4}
$$

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.

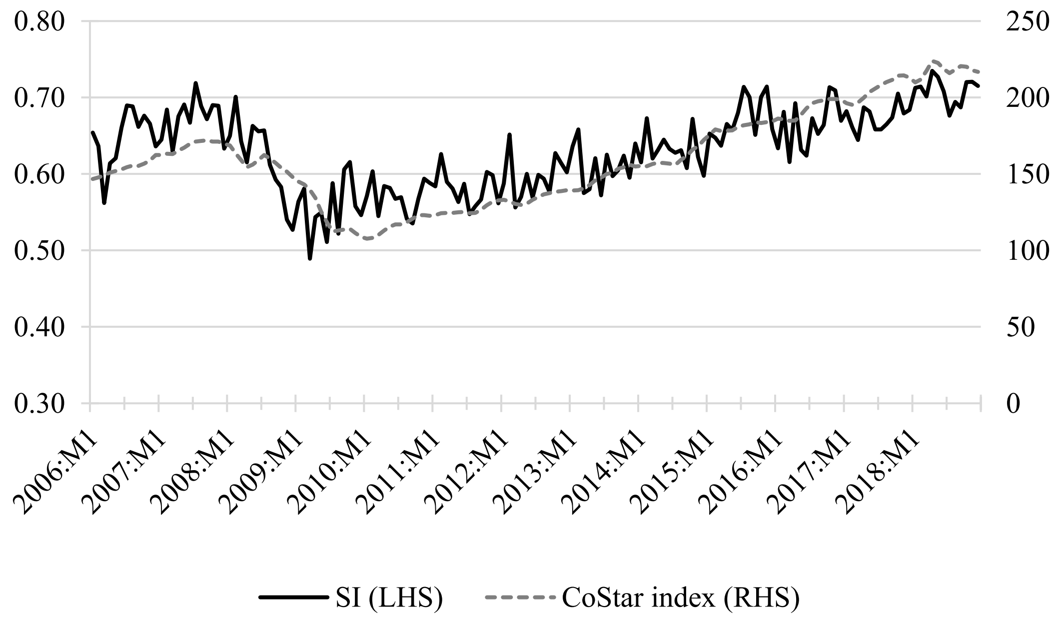
The training process described above is used to train 10 ANN models, in order to increase the robustness of the predictions. The average prediction value for each S&P news article is used to calculate its sentiment score. The monthly sentiment index value is then computed as the average sentiment score 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

[42] 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.

interpreted as a probability value. In the regression analyses, first differences of the monthly sentiment index score are used.

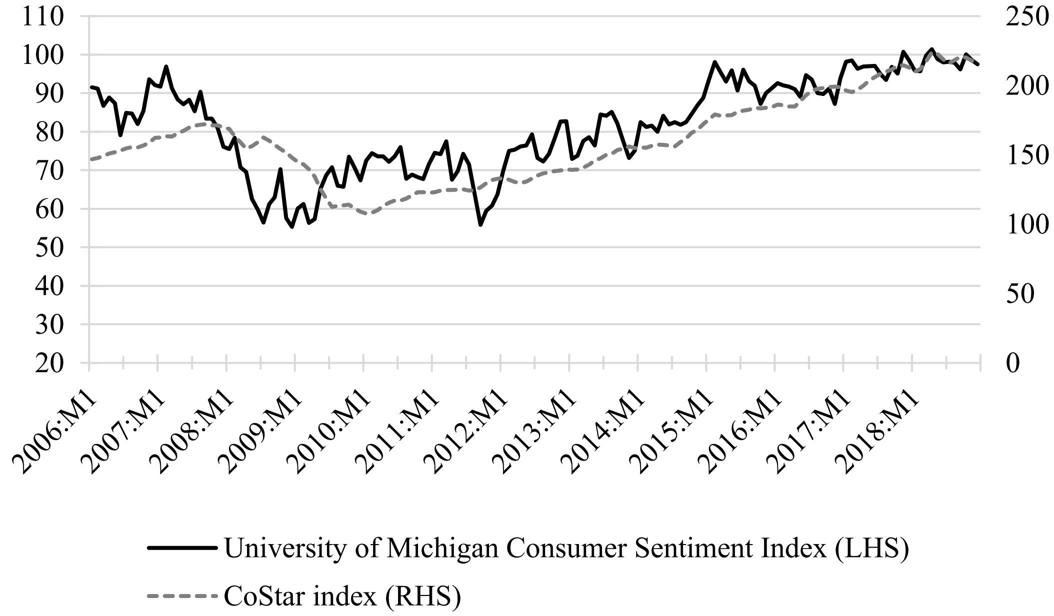
SI yields a mean value of 0.63 and a standard deviation of 0.05. This matches the average positive market performance of the CCRSI of 0.26% during the sample period. To provide some initial visual results, Figure 4.3 contrasts the SI with the CCRSI, as well as the *University of Michigan Consumer Sentiment 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 4.3: Temporal Progression of the SI



(Figure continues on the following page.)

Figure 4.3: Temporal Progression of the SI (continued)



Notes: The top chart in Figure 4.3 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 4.3 repeats the same lineup for the *University of Michigan Consumer Sentiment Index* (MCSI). The sample period is always 2006:M1 to 2018:M12.

4.5.2 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.

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:

$$\begin{aligned} y_{1,t} = {} & \alpha_{1,0} + \alpha_{1,1}\, y_{1,t-1} + \cdots + \alpha_{1,k}\, y_{1,t-k} + \alpha_{1,1}\, y_{2,t-1} + \cdots \\ & + \alpha_{1,k}\, y_{2,t-k} + u_{1,t} \\ y_{2,t} = {} & \alpha_{2,0} + \alpha_{2,1}\, y_{2,t-1} + \cdots + \alpha_{2,k}\, y_{2,t-k} + \alpha_{2,1}\, y_{1,t-1} + \cdots \\ & + \alpha_{2,k}\, y_{1,t-k} + u_{2,t}. \end{aligned} \tag{4.5}$$

$u_{i,t}$ 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 (4.5) are eventually adjusted by including a combined set of additional exogenous controls $\boldsymbol{z_t}$ with coefficient matrix $\boldsymbol{B}$.[43] This leads to the widely used standard-form VAR which can be estimated using ordinary least squares (OLS):

$$\boldsymbol{y_t} = \boldsymbol{A_0} + \boldsymbol{A_1 y_{t-1}} + \cdots + \boldsymbol{A_k y_{t-k}} + \boldsymbol{B z_t} + \boldsymbol{u_t}. \tag{4.6}$$

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

[43] Bold characters denote matrices.

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.

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. $\mu_t(m) = \boldsymbol{x_t\beta_m}$ 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:

$$y_t = \mu_t(m) + \sigma(m)\epsilon_t = \boldsymbol{x_t\beta_m} + \sigma(m)\epsilon_t.^{44} \qquad (4.7)$$

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*:

$$p = \begin{bmatrix} P(s_t = 1|s_{t-1} = 1) & P(s_t = 2|s_{t-1} = 1) \\ P(s_t = 1|s_{t-1} = 2) & P(s_t = 2|s_{t-1} = 2) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}. \qquad (4.8)$$

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:

$$L_t(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}) = \sum_{m=1}^{2} \frac{1}{\sigma_m} \phi\left(\frac{y_t - \mu_t(m)}{\sigma_m}\right) P(s_t = m|\Im_{t-1}, \boldsymbol{\delta}), \qquad (4.9)$$

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 $\Im_{t-1}$

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

the information set available at period *t-1*. Thus, the full log-likelihood for all time periods *T* is given by equation (4.10):

$$l(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}) = \sum_{t=1}^{T} log \left[\sum_{m=1}^{2} \frac{1}{\sigma_m} \phi \left(\frac{y_t - \mu_t(m)}{\sigma_m} \right) P(s_t = m | \Im_{t-1}, \delta) \right]. \quad (4.10)$$

Equation (4.10) can then be maximized with respect to $\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\delta}$. Due to the nature of transition probabilities, equation (4.10) 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 (4.10) 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.

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, pp. 525–527), the class of binary response models can be written as:

$$P(y = 1 | \boldsymbol{x}) = P(y = 1 | x_1, x_2 \dots x_k), \quad (4.11)$$

where $\boldsymbol{x}$ is a (*1* x *k*) - matrix of explanatory variables and y a binary response variable taking either value 1 or 0. Assuming that the response probability is linear in a set of parameters β_k, equation (4.11) can be written as:

$$P(y = 1|\boldsymbol{x}) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) = G(\beta_0 + \boldsymbol{x\beta}), \qquad (4.12)$$

with *G* being a nonlinear function taking values between 1 and 0, $0 < G(\,) < 1$, and $\boldsymbol{\beta}$ a (*k* x *1*)-matrix of coefficients. From the set of possible functions *G*, this paper employs the "logit"-link[45] $G(\beta_0 + \boldsymbol{x\beta}) = exp(\beta_0 + \boldsymbol{x\beta})/[exp(\beta_0 + \boldsymbol{x\beta}) + 1]$. Using maximum-likelihood estimation, coefficients can be calculated from the following equation:

$$\text{logit}[P(y = 1|\boldsymbol{x})] = \ln\left(\frac{P(y = 1|\boldsymbol{x})}{1 - P(y = 1|\boldsymbol{x})}\right) = \beta_0 + \boldsymbol{x\beta} + u. \qquad (4.13)$$

In order to analyze the relationship between market turns and the ANN-based sentiment indicator, y_t for month *t* is set to 1 for periods in which the *CCRSI* return is greater than or equal to 0 and 0 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 (4.13) becomes:

$$\text{logit}[P(y_t = 1|\boldsymbol{x}, SI)] = \beta_0 + \boldsymbol{x_t\beta} + \sum_i \gamma_{t-i}\, SI_{t-i} + u_t. \qquad (4.14)$$

The optimal lag length of the sentiment indicator i is chosen analogously to the Markov-switching model. However, 5 (7) 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.

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,

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

used, for example, by Diebold and Rudebusch (1989) to score leading indicators. These first metric is Brier's (1950) Quadratic Probability Score (QPS):

$$\text{QPS} = \text{T}^{-1} \sum_{t=1}^{T} (\hat{y}_t - y_t)^2, \tag{4.15}$$

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:

$$\text{LPS} = -\text{T}^{-1} \sum_{t=1}^{T} [(1 - y_t)ln(1 - \hat{y}_t) + y_t ln(\hat{y}_t)]. \tag{4.16}$$

4.6 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 *et al.* (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.

4.6.1 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 4.2 shows the results of estimating the endogenous relationship between the constructed monthly sentiment indicator and CCRSI returns, following equation (4.6). 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.

Table 4.2: 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]	

(Table continues on the following page.)

Table 4.2: VAR Estimation Results (continued)

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]	
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]	

(Table continues on the following page.)

Table 4.2: VAR Estimation Results (continued)

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

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%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2006:M10 to 2018:M12.

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 4.2, 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 2nd (and 5th) 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 1st, 2nd and 3rd 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.[46]

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 3rd and the 5th 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 2nd 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 4.3), 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 2nd lag of Model 3 becomes significant again and t-statistics for the 6th and 7th lag increase. Furthermore, the

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

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.

4.6.2 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 4.3 shows the estimation results of equation (4.7). Minimizing the average of AIC, HQ and BIC suggests a need to include the 7th, 8th and 9th 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 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 4.3. However, the results do not change substantially when similar controls with identical lags as in the VAR are included.

Table 4.3: 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:

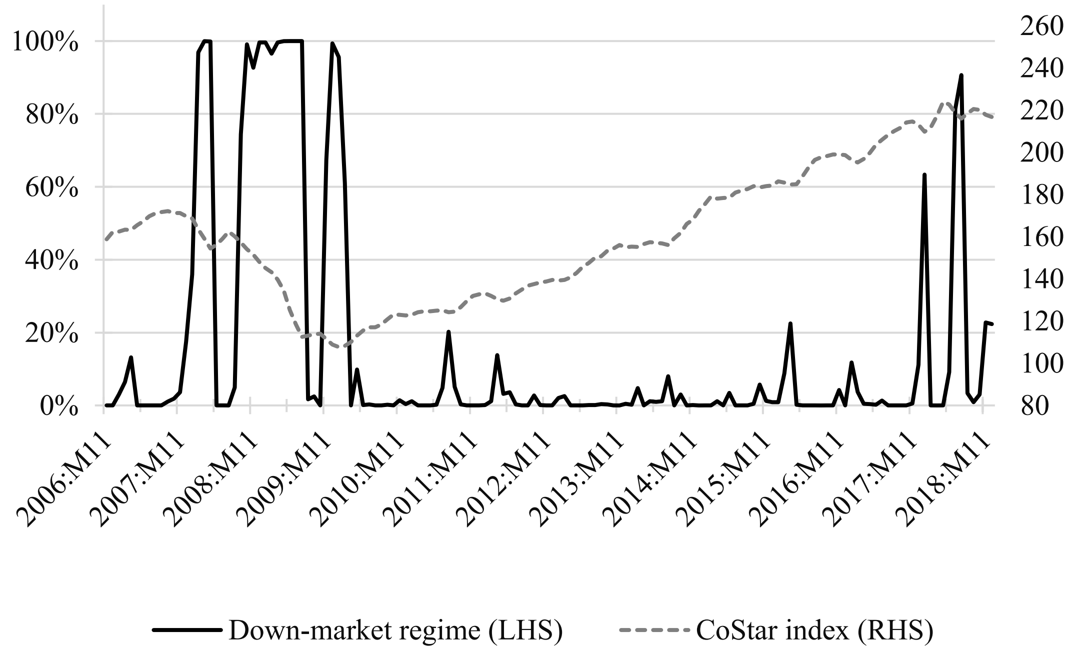
	Regime 1	*Regime 2*
(months)	25.99	4.90

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.

Figure 4.4 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% threshold.

Figure 4.4: Markov-Switching – Filtered Probabilities



Notes: This figure depicts filtered probabilities computed by the Markov-switching model estimated in Table 4.3. 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 (4.6) 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 4.4 facilitates Model 2 of Table 4.2 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.

Table 4.4: 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]	

(Table continues on the following page.)

Table 4.4: VAR Estimation Results in Up- and Down-Market Periods (continued)

TERM (-3)	0.285		0.030		0.690	
	[0.83080]		[0.08376]		[0.82694]	
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	

(Table continues on the following page.)

Table 4.4: VAR Estimation Results in Up- and Down-Market Periods (continued)

Granger Causality			
Sentiment indicator	**0.071**	0.755	**0.004**
CCRSI	0.117	0.366	0.572

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%. * denotes significance at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2006:M10 to 2018:M12.

In Model 2, the 6th and 7th 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 7th 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.

4.6.3 Binary Return Forecasts

Finally, following the reasoning of Tsolacos *et al.* (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 (4.14). 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 4.5. 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.

Table 4.5: 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	

(Table continues on the following page.)

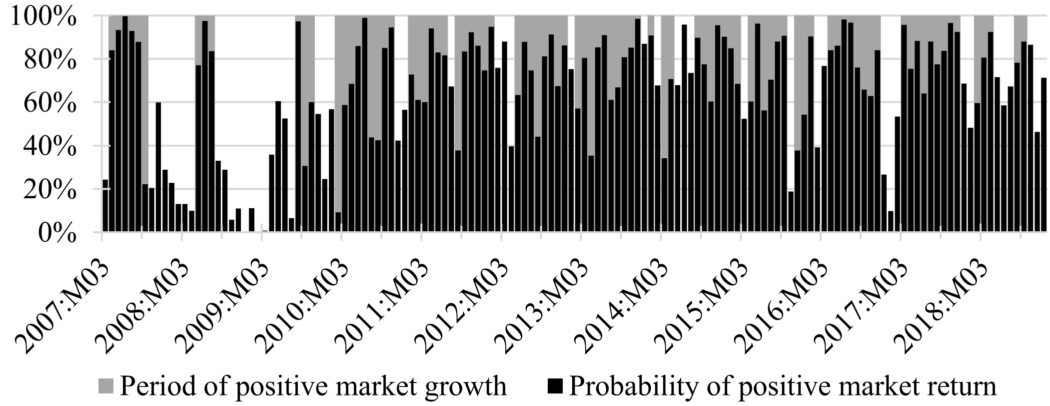
Table 4.5: Logit Estimation Results (continued)

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	

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). * denotes significance of z-statistics at 90%, ** significance at 95%, *** significance at 99%. The sample period is 2007:M03 to 2018:M12.

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 4.5 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 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 4.6, 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.[47]

Figure 4.5: 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 4.5. CCRSI returns are included as a second series to indicate periods of positive market growth. The sample period is 2007:M03 to 2018:M12.

Table 4.6: 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

(Table continues on the following page.)

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

Table 4.6: Forecast Performance (continued)

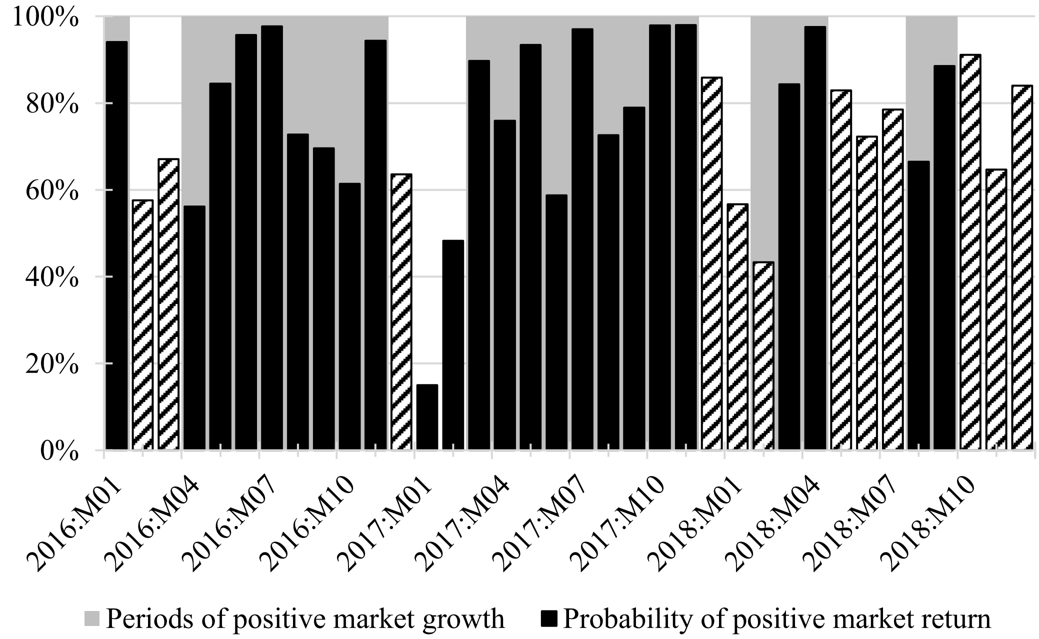
	Out-of-sample forecast performance	
	Logit model	**Naïve model**
Hit rate / correct-sign prediction	66.67 %	63.89 %
Brier's Quadratic Probability Score (QPS)	0.213	0.233
Log Probability Score (LPS)	0.604	0.660

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 4.5, 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.

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 4.6), 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 periods in the past, as well as during the forecasting period. Figure 4.6 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.

Figure 4.6: 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.

4.6.4 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 4.4 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.

4.7 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-labelled 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-labelled 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.

4.8 References

Antweiler, W. and Frank, M. Z. (2004), “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards”, *The Journal of Finance*, Vol. 59 No. 3, pp. 1259–1294.

Barkham, R. J. and Ward, C. W. R. (1999), “Investor Sentiment and Noise Traders: Discount to Net Asset Value in Listed Property Companies in the U.K”, *Journal of Real Estate Research*, Vol. 18 No. 2, pp. 291–312.

Beracha, E. and Wintoki, M. B. (2013), “Forecasting Residential Real Estate Price Changes from Online Search Activity”, *Journal of Real Estate Research*, Vol. 35 No. 3, pp. 283–312.

Bianchi, D. and Guidolin, M. (2014), “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*, Vol. 49 No. 1, pp. 116–164.

Borovkova, S. and Dijkstra, M. (2018), “Deep Learning Prediction of the EUROSTOXX 50 with News Sentiment”, *SSRN Electronic Journal*.

Boudoukh, J., Feldman, R., Kogan, S. and Richardson, M. (2013), “Which News Moves Stock Prices? A Textual Analysis”, NBER Working Paper No. 18725, National Bureau of Economic Research, Cambridge, 2013.

Braun, N. (2016), “Google Search Volume Sentiment and its Impact on REIT Market Movements”, *Journal of Property Investment & Finance*, Vol. 34 No. 3, pp. 249–262.

Brier, G. W. (1950), “Verification of Forecasts Expressed in Terms of Probability”, *Monthly Weather Review*, Vol. 78 No. 1, pp. 1–3.

Broyden, C. G. (1970), “The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations”, *IMA Journal of Applied Mathematics*, Vol. 6 No. 1, pp. 76–90.

Clayton, J., Ling, D. C. and Naranjo, A. (2009), "Commercial Real Estate Valuation: Fundamentals Versus Investor Sentiment", *The Journal of Real Estate Finance and Economics*, Vol. 38 No. 1, pp. 5–37.

Cortes, C. and Vapnik, V. (1995), "Support-Vector Networks", *Machine Learning*, Vol. 20 No. 3, pp. 273–297.

Cutler, D. M., Poterba, J. M. and Summers, L. H. (1989), "What Moves Stock Prices?", *The Journal of Portfolio Management*, Vol. 15 No. 3, pp. 4–12.

Das, P. K., Freybote, J. and Marcato, G. (2015), "An Investigation into Sentiment-Induced Institutional Trading Behavior and Asset Pricing in the REIT Market", *The Journal of Real Estate Finance and Economics*, Vol. 51 No. 2, pp. 160–189.

Das, S. R. and Chen, M. Y. (2007), "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web", *Management Science*, Vol. 53 No. 9, pp. 1375–1388.

Davis, A. K., Ge, W., Matsumoto, D. and Zhang, J. L. (2015), "The Effect of Manager-specific Optimism on the Tone of Earnings Conference Calls", *Review of Accounting Studies*, Vol. 20 No. 2, pp. 639–673.

Davis, A. K., Piger, J. M. and Sedor, L. M. (2012), "Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language", *Contemporary Accounting Research*, Vol. 29 No. 3, pp. 845–868.

Davis, A. K. and Tama-Sweet, I. (2012), "Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A", *Contemporary Accounting Research*, Vol. 29 No. 3, pp. 804–837.

Diebold, F. X. and Rudebusch, G. D. (1989), "Scoring the Leading Indicators", *The Journal of Business*, Vol. 62 No. 3, pp. 369–391.

Dietzel, M. A. (2016), "Sentiment-based Predictions of Housing Market Turning Points with Google Trends", *International Journal of Housing Markets and Analysis*, Vol. 9 No. 1, pp. 108–136.

Dietzel, M. A., Braun, N. and Schäfers, W. (2014), "Sentiment-based Commercial Real Estate Forecasting with Google Search Volume Data", *Journal of Property Investment & Finance*, Vol. 32 No. 6, pp. 540–569.

Din, A., Hoesli, M. and Bender, A. (2001), "Environmental Variables and Real Estate Prices", *Urban Studies*, Vol. 38 No. 11, pp. 1989–2000.

Engelberg, J. E., Reed, A. V. and Ringgenberg, M. C. (2012), "How are Shorts Informed?", *Journal of Financial Economics*, Vol. 105 No. 2, pp. 260–278.

Feldman, R., Govindaraj, S., Livnat, J. and Segal, B. (2010), "Management's Tone Change, Post Earnings Announcement Drift and Accruals", *Review of Accounting Studies*, Vol. 15 No. 4, pp. 915–953.

Ferguson, N., Philip, D., Lam, H. Y. T. and Guo, J. M. (2015), "Media Content and Stock Returns: The Predictive Power of Press", *Multinational Finance Journal*, Vol. 19 No. 1/1, pp. 1–31.

French, K. R. and Roll, R. (1986), "Stock Return Variances", *Journal of Financial Economics*, Vol. 17 No. 1, pp. 5–26.

Freybote, J. (2016), "Real Estate Sentiment as Information for REIT Bond Pricing", *Journal of Property Research*, Vol. 33 No. 1, pp. 18–36.

Freybote, J. and Seagraves, P. A. (2017), "Heterogeneous Investor Sentiment and Institutional Real Estate Investments", *Real Estate Economics*, Vol. 45 No. 1, pp. 154–176.

Freybote, J. and Seagraves, P. A. (2018), "The Impact of Investor Sentiment on Commercial Real Estate Market Liquidity", *Journal of Real Estate Research*, Vol. 40 No. 4, pp. 597–627.

García, D. (2013), "Sentiment during Recessions", *The Journal of Finance*, Vol. 68 No. 3, pp. 1267–1300.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant, L. (2009), “Detecting Influenza Epidemics Using Search Engine Query Data”, *Nature*, Vol. 457 No. 7232, pp. 1012–1014.

Groß-Klußmann, A. and Hautsch, N. (2011), “When Machines Read the News: Using Automated Text Analytics to Quantify High Frequency News-implied Market Reactions”, *Journal of Empirical Finance*, Vol. 18 No. 2, pp. 321–340.

Hamilton, J. D. (1989), “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle”, *Econometrica*, Vol. 57 No. 2, p. 357.

Hanna, A. J., Turner, J. D. and Walker, C. B. (2017), “News Media and Investor Sentiment over the Long Run”, QUCEH Working Paper Series 2017-06, Queen's University Centre for Economic History (QUCEH), Queen's University, Belfast, 2017.

Hausler, J., Ruscheinsky, J. and Lang, M. (2018), “News-based Sentiment Analysis in Real Estate: A Machine Learning Approach”, *Journal of Property Research*, Vol. 35 No. 4, pp. 344–371.

Henry, E. (2008), “Are Investors Influenced By How Earnings Press Releases Are Written?”, *Journal of Business Communication*, Vol. 45 No. 4, pp. 363–407.

Henry, E. and Leone, A. J. (2016), “Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone”, *The Accounting Review*, Vol. 91 No. 1, pp. 153–178.

Heston, S. L. and Sinha, N. R. (2016), “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., Lizieri, C. and MacGregor, B. (2008), “The Inflation Hedging Characteristics of US and UK Investments: A Multi-Factor Error Correction Approach”, *The Journal of Real Estate Finance and Economics*, Vol. 36 No. 2, pp. 183–206.

Hohenstatt, R., Käsbauer, M. and Schäfers, W. (2011), “"Geco" and its Potential for Real Estate Research: Evidence from the U.S. Housing Market”, *Journal of Real Estate Research*, Vol. 33 No. 4, pp. 471–506.

Huang, X., Teoh, S. H. and Zhang, Y. (2014), “Tone Management”, *The Accounting Review*, Vol. 89 No. 3, pp. 1083–1113.

Jegadeesh, N. and Wu, D. (2013), “Word Power: A New Approach for Content Analysis”, *Journal of Financial Economics*, Vol. 110 No. 3, pp. 712–729.

Kearney, C. and Liu, S. (2014), “Textual Sentiment in Finance: A Survey of Methods and Models”, *International Review of Financial Analysis*, Vol. 33, pp. 171–185.

Kim, C.-J. (1994), “Dynamic Linear Models with Markov-switching”, *Journal of Econometrics*, Vol. 60 No. 1-2, pp. 1–22.

Kothari, S. P., Li, X. and Short, J. E. (2009), “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*, Vol. 84 No. 5, pp. 1639–1670.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012), “ImageNet Classification with Deep Convolutional Neural Networks”, in Pereira, F., Burges, C.J.C., Bottou, L. and Weinberger, K.Q. (Eds.), *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc, pp. 1097–1105.

Li, F. (2010), “The Information Content of Forward-Looking Statements in Corporate Filings-A Naïve Bayesian Machine Learning Approach”, *Journal of Accounting Research*, Vol. 48 No. 5, pp. 1049–1102.

Lin, C. Y., Rahman, H. and Yung, K. (2009), “Investor Sentiment and REIT Returns”, *The Journal of Real Estate Finance and Economics*, Vol. 39 No. 4, pp. 450–471.

Ling, D. C., Naranjo, A. and Scheick, B. (2014), “Investor Sentiment, Limits to Arbitrage and Private Market Returns”, *Real Estate Economics*, Vol. 42 No. 3, pp. 531–577.

Liu, B. (2012), “Sentiment Analysis and Opinion Mining”, *Synthesis Lectures on Human Language Technologies*, Vol. 5 No. 1, pp. 1–167.

Lizieri, C., Satchell, S., Worzala, E. and Dacco, R. (1998), “Real Interest Regimes and Real Estate Performance: A Comparison of UK and US Markets”, *Journal of Real Estate Research*, Vol. 16 No. 3, pp. 339–356.

Loughran, T. and McDonald, B. (2011), “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks”, *The Journal of Finance*, Vol. 66 No. 1, pp. 35–65.

Loughran, T. and McDonald, B. (2015), “The Use of Word Lists in Textual Analysis”, *Journal of Behavioral Finance*, Vol. 16 No. 1, pp. 1–11.

Loughran, T. and McDonald, B. (2016), “Textual Analysis in Accounting and Finance: A Survey”, *Journal of Accounting Research*, Vol. 54 No. 4, pp. 1187–1230.

Marcato, G. and Nanda, A. (2016), “Information Content and Forecasting Ability of Sentiment Indicators: Case of Real Estate Market”, *Journal of Real Estate Research*, Vol. 38 No. 2, pp. 165–203.

Ozik, G. and Sadka, R. (2012), “Media and Investment Management”, *SSRN Electronic Journal*.

Peterson, S. and Flanagan, A. (2009), “Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal”, *Journal of Real Estate Research*, Vol. 31 No. 2, pp. 147–164.

Rochdi, K. and Dietzel, M. (2015), “Outperforming the Benchmark: Online Information Demand and REIT Market Performance”, *Journal of Property Investment & Finance*, Vol. 33 No. 2, pp. 169–195.

Roll, R. (1988), "R2", *The Journal of Finance*, Vol. 43 No. 3, pp. 541–566.

Rosenblatt, F. (1958), "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review*, Vol. 65 No. 6, pp. 386–408.

Ruscheinsky, J. R., Lang, M. and Schäfers, W. (2018), "Real Estate Media Sentiment through Textual Analysis", *Journal of Property Investment & Finance*, Vol. 36 No. 5, pp. 410–428.

Sinha, N. R. (2016), "Underreaction to News in the US Stock Market", *Quarterly Journal of Finance*, Vol. 06 No. 02 1650005, 1-46.

Smales, L. A. (2014), "News Sentiment in the Gold Futures Market", *Journal of Banking & Finance*, Vol. 49, pp. 275–286.

Soo, C. K. (2015), "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., Najand, M. and Shen, J. (2016), "Stock Return Predictability and Investor Sentiment: A High-Frequency Perspective", Working Paper, Department of Finance, Strome College of Business, Old Dominion University, Norfolk, 2016.

Tetlock, P. C. (2007), "Giving Content to Investor Sentiment: The Role of Media in the Stock Market", *The Journal of Finance*, Vol. 62 No. 3, pp. 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M. and Macskassy, S. (2008), "More than Words: Quantifying Language to Measure Firms' Fundamentals", *The Journal of Finance*, Vol. 63 No. 3, pp. 1437–1467.

Tieleman, T. and Hinton, G. E. (2012), "Lecture 6.5-Rmsprop, Coursera: Neural Networks for Machine Learning", Technical Report.

Tsolacos, S., Brooks, C. and Nneji, O. (2014), "On the Predictive Content of Leading Indicators: The Case of U.S. Real Estate Markets", *Journal of Real Estate Research*, Vol. 36 No. 4, pp. 541–573.

Walker, C. B. (2014), “Housing Booms and Media Coverage”, *Applied Economics*, Vol. 46 No. 32, pp. 3954–3967.

Walker, C. B. (2016), “The Direction of Media Influence: Real-estate News and the Stock Market”, *Journal of Behavioral and Experimental Finance*, Vol. 10, pp. 20–31.

Werbos, P. J. (1974), *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Thesis (Ph. D.).

Wooldridge, J. M. (2016), *Introductory Econometrics: A Modern Approach,* Sixth Edition, Cengage Learning, Australia.

5 Artificial Intelligence, News Sentiment and Property Market Liquidity

5.1 Abstract

This paper examines a text-based sentiment indicator to explain variations in direct property market liquidity in the United States. In a deep-learning framework, market sentiment is extracted from 66,070 US real estate market news articles provided by the S&P Global Intelligence database, using the medium of an artificial neural network. For the training process, 17,822 distant-labelled investment ideas from the crowd-sourced investment advisory platform Seeking Alpha are used. According to the results of the estimated autoregressive distributed lag models, 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, which should be accounted for by market participants in terms of their investment decisions but also when benchmarking their portfolios to market indices. This paper not only expands the literature on text-based sentiment indicators in real estate, but is also the first to demonstrate the application of AI for sentiment extraction from news articles in a market liquidity setting.

Keywords: Artificial intelligence, Market liquidity, Sentiment, News analytics, Commercial real estate, Deep learning

5.2 Introduction

With respect to direct real estate, scholars such as Fisher *et al.* (2003) and Clayton *et al.* (2009) highlight the time-varying nature of market liquidity in contrast to other asset classes. Impressively demonstrated during the last market cycle, "ease of selling" increases during up-market periods, and decreases accordingly in down-market phases. It can be argued that this peculiarity of the property market may be driven by the characteristics of real assets which are usually large-volume, heterogeneous and traded infrequently in segmented, local markets. However, in accordance with Liu (2015), who demonstrates a relationship between sentiment and liquidity for the stock market, Freybote and Seagraves (2018) have more recently pointed out the influence of market participants' sentiment on liquidity in direct property markets.

By introducing a novel approach to extracting prevailing market sentiment from news articles by means of a deep-learning approach, this study not only extends research on sentiment in commercial real estate markets, but also the very limited literature on investor sentiment as an explanatory factor for the variation in commercial real estate market liquidity. At first, an artificial neural network (ANN) is trained on a distant-labelled dataset from the investment advisory platform *Seeking Alpha,* in order to classify news articles from the *S&P Global Market Intelligence* database regarding their inherent sentiment in a second step. By calculating an aggregate sentiment score for the news articles in a respective month, this procedure enables creating a monthly market sentiment indicator which can be analyzed for its influence on private real estate market liquidity.

With respect to text-based sentiment analysis, this approach has the potential to extract a rich information structure from news articles as ANNs do not rely on a predefined set of rules to indicate on the sentiment polarity expressed by the respective article's author. By using a distant-labelled dataset, the ANN itself decides which features should be accounted for to provide the most accurate sentiment classification. Thus, the resulting sentiment indicator may not only be superior to other text-based classifiers, but also exceed the capabilities of surveys or market-based proxies, such as mortgage fund flows or closed-ended fund discounts. Furthermore, the approach benefits from a direct link to market sentiment, as it can be calculated in real-time and is less cost- and time-consuming than surveys or manually classified machine-learning approaches.

During 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 two different liquidity proxies, in the case of the market depth dimension, 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 resulting in an influence on market liquidity.

The remainder of this paper is structured as follows. Section 5.3 provides a short overview of relevant and related literature. Section 5.4 and 5.5 describe the dataset, the sentiment extraction procedure, and the econometric approach used to estimate the results following in Section 5.6. Section 5.7 concludes.

5.3 Literature Review

The properties of market ***liquidity for the general stock market*** have undergone extensive empirical research during the last few decades. Chordia *et al.* (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 for 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 scarce. 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 *et al.* (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 maker's price adjustment to order flow. However, applying the framework of Kyle (1985) and DeLong *et al.* (1990) to direct property markets poses difficulties: No short-sale constraints exist in the models, thus noise traders increase trading both when sentiment is high and low. Additionally, the framework rests 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.[48] 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 result 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 an extension of their model, the authors additionally incorporate a higher propensity of the 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 for 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 extra space for contrary private signals. On the other hand, the high transaction fees in the property market might stifle this behavior. The answer on the question of which effect prevails is insofar an empirical one. 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 in particular for a direct property market application.

The first paper to analyze the potential ***relationship between sentiment and liquidity*** for the commercial real estate market is provided by Clayton *et al.* (2008). The authors examine potential explanations of time variation in commercial real estate market liquidity. In a subsequent empirical analysis facilitating 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) carry out a detailed analysis on the sentiment-liquidity relationship for the office market, using Markov-switching models. The authors use quarterly data for their analyses, facilitate activity (turnover) and market depth (Amihud) liquidity measures, and the *Real Estate Research Corporation (RERC)/Situs* survey as well as *Real Capital Analytics 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 times of high liquidity, and a larger

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

impact on the market-depth dimension (Amihud) of liquidity in times 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 present paper posits that additional insights can be gained from an analysis which refines several dimensions of previous work on the topic. At first, despite the high quality of NCREIF data, quarterly analysis prevents a fine-grained analysis of a potential mix of contemporary and lagged effects of sentiment on liquidity, due to its high degree of aggregation. It might be revealing to decompose the effect into its time-dependent components by incorporating a distributed lag structure into quantitative analyses. The rationale behind this approach lies in the specifics of the direct property market; Ametefe *et al.* (2016) analyze the inefficiencies in direct property markets and among others, emphasize the decentralized structure of the market and the resulting, often time-consuming need to find a counterparty. Together with long time frames to complete transactions (see IPF, 2004; Scofield, 2013; Devaney and Scofield, 2015), sentiment-driven buy or sell decisions may only 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) 165 days.[49] With many transactions in Devaney and Scofield's sample finishing substantially faster or slower, a sufficiently long time period for the market-wide sentiment-liquidity relationship has to be considered.

Secondly, the use of an alternative measure of real estate investor sentiment might have the potential to strengthen the empirical power of the analyses. This paper therefore facilitates a novel text-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 be a superior sentiment indicator, but can also be applied to short aggregation periods. The obtained sentiment scores are used to create an index proxying overall investor

[49] Although Devaney and Scofield (2015) analyze the UK real estate transaction market, conclusions for the US market should be valid, as both markets are highly developed.

sentiment in the US property market on a monthly basis. The application of news articles might allow for a more unmediated investigation, compared for example, to buy-sell indices, which constitute the aggregated results of potentially month-long transaction processes, initially possibly triggered by sentiment. With the utilization of the described deep learning model, this paper additionally extends the so far only AI based sentiment extraction approach in real estate research of Hausler *et al.* (2018).

5.4 Data and Methodology

The paper facilitates several data sources. For the ANN training procedure, text data from the crowd-sourced financial content platform *Seeking Alpha* (SA) is utilized. The sentiment measure itself is based on the vast *S&P Global Market Intelligence* (S&P) news database. In order to construct the liquidity measures required for the regression analyses, both *CoStar* and *Real Capital Analytics* (RCA) data are used. Finally, data required for several control variables is gathered from the webpage of the *Federal Reserve Bank of St. Louis* (FRED).

5.4.1 Sentiment Index

The chosen distant labelling approach for training the artificial neural network requires a large amount of financial text data with distinct, unambiguous sentiment polarity. *Seeking Alpha*, as a crowd-sourced platform providing investment information in its large long idea/short idea sections is well suited for the intended approach and has already found its way into academic research through an application as a news provision database for Chen *et al.* (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, *Seeking Alpha's* long and short ideas contain a short summary section which delineates the quintessence of the text.[50] As those summary sections succinctly cover the authors' positive or negative opinion on the equity or market in question, they serve as a reliable data source to isolate textual sentiment in a

[50] 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."

financial context. For the ANN's training process, a balanced sample of long and short summary sections containing 17,822 SA texts is thus collected.[51]

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 gauge of the sentiment prevailing in the real estate 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.

Text classification procedures normally consist of four stages: pre-processing, feature extraction, feature selection and classification (Uysal and Gunal, 2014). To provide the ANN with comparable data for the later steps, identical pre-processing steps have to be carried out both on the S&P and the SA text datasets. Additionally, unicode categories 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 unintended association of sentiment with certain date or time expressions. As a further extension, the training and classification datasets 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.[52]

For feature extraction, feature selection and classification, SA investment ideas are annotated with a distant supervision label of 0 if they are from the short idea category,

[51] The sample consists of texts from 3,107 different freelance authors, the average length of each text is 381 characters.
[52] This paper facilitates 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.

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 datasets. By this means, textual documents are expressed as binary vectors, which are interpretable by the neural network. Note that embedding layers and a larger word corpus were tested, but neither increased performance.

This study uses a random sample of 80% of the 17,822 one-hot encoded SA texts for the training of the sentiment classification ANN. The remaining 20% 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 facilitate ReLU (Rectified Linear Unit) activation functions. The reduction of nodes per layer is used in order to gradually reduce the complexity of the feature space. In formal terms, each of the ReLU layers processes data according to the following equation:

$$max(0, dot(Input, W) + b), \tag{5.1}$$

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.[53]

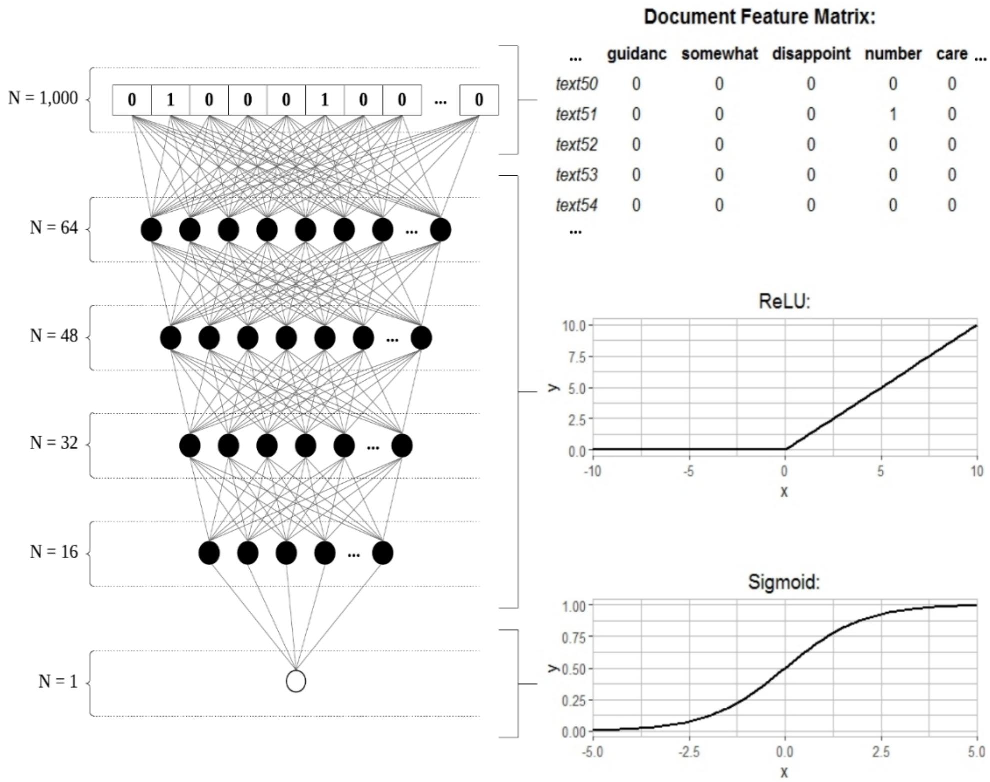
While initially set ANN weights are random, the training process carries out a step-wise adjustment process based on a feedback signal. This is provided by the combination of a sigmoid layer and a loss function. The sigmoid function, as the last layer of the ANN, squashes output values into the spectrum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

$$\hat{y}_k = \frac{1}{1 + e^{-t}} \; with \; t = dot(Input, W) + b. \tag{5.2}$$

Figure 5.1 provides a summary overview of the conceptual layout of the multilayer perceptron facilitated in this paper.

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

Figure 5.1: ANN Layout



Notes: Figure 5.1 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.

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 label value y_k for each textual document k:

$$L = \frac{1}{n}\sum_{k=1}^{n}\left[-1\left(y_k log(\hat{y}_k) + (1 - y_k) log(1 - \hat{y}_k)\right)\right]. \qquad (5.3)$$

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, facilitating Root Mean Square Propagation (RMSprop) as the optimizer algorithm (Tieleman and Hinton, 2012) is executed. 6 epochs, each containing all batches, are performed. Hence, weights W and b undergo a total amount of 174 updates specified by the equations:

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

$$\begin{aligned} v_{dW}(t) &= \beta v_{dW}(t-1) + (1-\beta)\left(\frac{\partial L}{\partial W}(t)\right)^2 \\ \Delta W(t) &= -\frac{\eta}{\sqrt{v_{dW}(t)+\varepsilon}}\left(\frac{\partial L}{\partial W}(t)\right) \\ \Delta b(t) &= -\frac{\eta}{\sqrt{v_{db}(t)+\varepsilon}}\left(\frac{\partial L}{\partial b}(t)\right), \end{aligned} \tag{5.4}$$

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, respectively. 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 constitutes 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 resulting trained model, a sentiment score for each document in the S&P dataset is estimated. Aggregating scores on a monthly basis, the mean score of each document published in the respective month is utilized as its sentiment value. For the study period between January 2006 and December 2018, the average monthly sentiment score (*SM*) is 0.63, and the standard deviation 0.05.

5.4.2 Liquidity Proxies

In their analysis of the literature on liquidity in financial markets, Ametefe *et al.* (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 aggravated by limited data availability and conceptual differences between both markets. For the *tightness* dimension of liquidity, Ametefe *et al.* suggest several bid-ask spread proxies, although for the direct property markets, these proxies are unavailable.[54] For the fifth dimension, namely *immediacy*, Ametefe *et al.* (2016) merely suggest real estate time on market as a proxy. To depict this dimension, a representative dataset of time-on-market information would be required. Due to the unavailability of such specific datasets, this study focuses on the 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 *et al.*, 2009; Glascock and Lu-Andrews, 2014; Freybote and Seagraves, 2018) is used to cover the dimensions *depth* and *resilience*. The measure is calculated as:[55]

$$AMI_t = log\left(\frac{|R_t|}{Vol_t}\right). \tag{5.5}$$

AMI$_t$ captures the absolute value of the price impact (*R*) of the 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.[56] The numerator is represented by the absolute of the return on the *CoStar Commercial Repeat-Sale Index* for month *t*.[57] The application of the Amihud measure 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 *et al.* (2016) for the fourth liquidity dimension, *breadth*. The measure *VOL*$_t$ is the transaction volume of the direct US property market for month *t* in billion USD.[58] By incorporating trading volume into the analysis, Baker and Stein's (2004) supposed positive relationship to

[54] The conversion of Ametefe *et al.'s* (2016) tightness proxy *relative quoted spread* to a direct real estate market use case is theoretically possible, but only feasible with the facilitation of a private dataset containing the required bid and ask prices of property transactions.
[55] 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 amount over time, by scaling it with the consumer price index for the US.
[56] RCA collects data on transactions of the volume USD 2.5 million or greater.
[57] RCA also provides a transaction-based monthly direct real estate index of the US market; however, the construction methodology of the index leads to an unacceptable level of autocorrelation which inevitably causes severe problems in the upcoming quantitative analyses.
[58] 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 is available). This study seeks to analyze monthly time series and facilitates RCA data, for which no turnover count was available.

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 Demsetz, 1968 or Glosten and Milgrom, 1985). Monthly transaction volume data for this study is again obtained from RCA.

5.4.3 Control Variables

In order to control for the effect of other potentially influential factors explaining variation in direct property market liquidity, a set of control variables is incorporated into the regression analyses. Liu (2015) considers the possibility that sentiment might merely capture macroeconomic conditions. For this reason, the paper controls for the state of the general economy as an explanatory factor for 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 condition 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 United States (*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 Commercial Repeat-Sale Index* (CCRSI).[59] Descriptive statistics for the liquidity, sentiment and control variables for the study period between January 2006 and December 2018 can be found in Table 5.1.

[59] Variables proxying the US general stock 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 next 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.

Table 5.1: 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	1876.28	2061.00	794.04	391.00	3651.00
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)	1038.44	1064.51	168.45	754.71	1324.35
CPI (%)	0.16	0.17	0.39	-1.92	1.01
UNRATE (%)	6.37	5.65	1.99	3.70	10.00

Notes: Table 5.1 reports summary statistics of the constructed sentiment measure *SM* as well as four different proxies of direct real estate market liquidity. *AMI* aims at covering 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 section 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 to 2018:M12.

5.5 Regression Analysis

Given that this paper seeks to decompose the potential effect of sentiment on liquidity into its contemporary and lag components, the slow nature of the direct property market must be reflected in the empirical models by the addition of distributed lag terms. An analysis of the liquidity measures facilitated in this paper furthermore reveals a strong negative auto-correlation.[60] For this reason, regression analysis requires the utilization of autoregressive terms (i.e. lagged liquidity variables). The integration of both distributed lags as well as autoregressive components requires the use of autoregressive distributed lag (ARDL) models. By including the dependent

[60] 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.

variable besides other explanatory variables as regressors, ARDL models allow a simultaneous assessment of a potential long- and short-run relationship between market liquidity, sentiment and macroeconomic controls. ARDL models have gained particular attention through the work of Pesaran and Shin (1998) and Pesaran *et al.* (2001) on cointegrating relationships. In formal terms, equation (5.6) depicts the applied model:

$$
\begin{aligned}
LIQ_t = \alpha_0 + \sum_{i=1}^{I} \alpha_i LIQ_{t-i} + \sum_{j=0}^{J} \beta_j SM_{t-j} + \sum_{k=1}^{K} \sum_{l_k=0}^{L} \gamma_{k,l_k} x_{k,t-l_k} \\
+ \sum_{m=2}^{12} \delta_m Month_m + \varepsilon_{t,}
\end{aligned}
\tag{5.6}
$$

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 set of macroeconomic controls, $Month_m$ monthly dummy variables and ε_t a random disturbance term.

Running augmented Dickey-Fuller tests indicates that some variables are stationary in levels (i.e. I(0)), while others are integrated of order 1. Thus, to ascertain unbiased and consistent estimates, the research framework must ultimately account for a potential existing cointegrating relationship. By estimating equation (5.6) in first differences and including the 1st lag of all regressors in levels, an unconstrained error correction model (ECM) is derived. Subsequently, the bound-testing procedure of Pesaran *et al.* (2001) is conducted. In case of the presence 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 ascertain an unbiased and consistent estimation. Bound-testing however, finds no evidence of a long-run relationship, so that each series of equation (5.6) 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 on a fixed lag of up to 9 months in the OLS models, so as to provide a complete picture of the relationship up to 3 quarters in the past ($I = J = 1, \ldots, 9$). 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).[61]

5.6 Results

Ordinary least squares estimation of equation (5.6) in first differences leads to the results depicted in columns 1 and 2 of Table 5.2. The results in column 1 exhibit an OLS regression facilitating *AMI* as the dependent variable, column 2 the results for *VOL*.

Table 5.2: Liquidity and Sentiment: Autoregressive Distributed Lag Models

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

(Table continues on the following page.)

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

Table 5.2: Liquidity and Sentiment: Autoregressive Distributed Lag Models (continued)

SM (t-9)	-6.095		28.955	
	(4.057)		(33.929)	
AMI (t-1)	-0.489	***		
	(0.103)			
AMI (t-2)	-0.387	***		
	(0.113)			
AMI (t-3)	-0.565	***		
	(0.111)			
AMI (t-4)	-0.206	*		
	(0.116)			
AMI (t-5)	-0.157			
	(0.113)			
AMI (t-6)	-0.223	*		
	(0.113)			
AMI (t-7)	-0.163			
	(0.108)			
AMI (t-8)	-0.165			
	(0.106)			
AMI (t-9)	0.019			
	(0.093)			
VOL (t-1)			-0.837	***
			(0.095)	
VOL (t-2)			-0.555	***
			(0.117)	
VOL (t-3)			-0.298	**
			(0.118)	
VOL (t-4)			-0.250	**
			(0.118)	
VOL (t-5)			-0.202	
			(0.124)	
VOL (t-6)			-0.115	
			(0.126)	
VOL (t-7)			-0.162	
			(0.133)	
VOL (t-8)			0.056	
			(0.124)	
VOL (t-9)			0.035	
			(0.103)	

(Table continues on the following page.)

Table 5.2: Liquidity and Sentiment: Autoregressive Distributed Lag Models (continued)

Macroeconomic 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 5.2 reports findings 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 facilitating Amihud's (2002) measure for illiquidity (*AMI*), representing the price impact of transaction volume. Column 2 shows the coefficients for transaction volume (*VOL*). Standard errors are reported in brackets underneath the coefficient estimates. 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 facilitates 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. The sample period is 2006:M11 to 2018:M12.

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% level for the first 3 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 market after periods of very high increases in transaction volumes, which effects *VOL* directly and *AMI* indirectly through the lower denominator value during periods consecutive to such "high volume" periods.

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. Specifically, parameters of the contemporary sentiment value and lags 1, 2, 4, 6 and 7 are at least significant at the 5% level, and lag 8 is furthermore still significant at the 10% level. All sentiment coefficient values exhibit 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* seems to track the pattern of times to completion, with around 87% of the property purchases and 86% of the property sales transactions requiring a time period of no more than 239 days. The effect of sentiment on liquidity thus seems to dribble into the market over an extended period, initially conceivably caused 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, facilitating differences in transaction volume (*VOL*) as a proxy for the breadth dimension of liquidity, exhibits similar, but weaker results. All sentiment parameters show a positive relationship with differences in volume, although only the contemporary volume parameter and the 2nd lag are significant at the 10% level. These results yield the conclusion that positive (negative) sentiment stimulates (stifles) the overall amount of trading, but that the effect of sentiment on price impact (i.e. *AMI*) seems to exceed the effect of *VOL*.[62] However, the significance of the 2nd 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 the direct property market, which moderate the effect of sentiment the extended model of Baker and Stein (2004) posits on trading volume.

In order to secure the robustness of the regression results, several diagnostics tests have been performed. To identify potential problems with auto-correlation 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% level. For this reason, the regressions are re-estimated, facilitating Newey-West standard errors (Newey and West, 1987). The results remain basically unchanged. Furthermore, a Breusch-Pagan Test is performed, resulting in no heteroscedasticity problems in the regression residuals for both

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

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 be a main driver of a strong relationship between sentiment and Amihud. To eliminate that possibility, a model including contemporary *VOL* as well as 9 lags is estimated. The untabulated results show an increased strength of the effect of sentiment on Amihud.

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 units (*UNITS*) traded 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 5.3 in the appendix. Contemporary sentiment in the *PROPS* regression exhibits a positive parameter value which is furthermore significant at the 1% percent level. The 2nd lag of *PROPS* is also positive and significant at the 10% level. *UNITS* is significant and positive at the 5% 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 dimension of liquidity, as proxied by *AMI*. Market participants in the direct commercial real estate market seem to exhibit sentiment-induced behavior as a trigger for a transaction. Due to the long transaction periods, the effect of sentiment on liquidity however, only gradually manifests itself over the following months.

5.7 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 artificial neural network. Highly sentiment-loaded text documents from crowd-sourced investment content provider *Seeking Alpha* serve as a distant-labelled dataset and were facilitated to train the discrimination between positive and negative sentiment to an artificial neural network. The trained network is then used 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 so doing, and through aggregating monthly polarity scores of the single articles, a monthly real estate sentiment index is designed. In a next step, the potential of the sentiment index to explain 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), suggests that an effect of sentiment on liquidity might manifest in a lagged fashion. Furthermore, an increase of 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 depict market breadth. With the intention to examine the hypothesis of a lagged relationship between sentiment and liquidity empirically, autoregressive distributed lag (ARDL) models are estimated. OLS estimation exhibits strong evidence supporting an intertemporal relationship between the facilitated measure of sentiment and the depth and resilience dimension of liquidity. Regressions yield several significant lags for Amihud, which range up to order 8. The relationship between sentiment and the breadth dimension of liquidity is somewhat weaker, with a significant 2nd lag however still prevailing.

A shortcoming of this study lies in the unavailability of liquidity proxies for two of the dimensions, as outlined by Ametefe *et al.* (2016). Future research could facilitate datasets which allow for the construction of alternative liquidity proxies which represent the dimensions of tightness and immediacy, so as to provide a more complete picture of the sentiment-liquidity relationship. Furthermore, in the context of AI-based sentiment analysis, the authors of this study belief that there is a vast potential for future research and practical application. With the collection of a broader spectrum of distant-supervision labels and the extension of the amount of news constituting the sentiment index, an even more complete depiction of the facets of real estate market sentiment might be feasible.

5.8 Appendix

Table 5.3: Liquidity and Sentiment: Autoregressive Distributed Lag Models with Alternative Liquidity Measures

	Dependent variable			
	Props		**Units (bn)**	
	(1)		(2)	
C	-273.408		-0.007	
	(166.305)		(0.020)	
SM	5668.750	***	0.441	**
	(1416.210)		(0.172)	
SM (-1)	2280.774		0.086	
	(1737.298)		(0.209)	
SM (-2)	3753.825	*	0.046	
	(1909.511)		(0.223)	
SM (-3)	-2398.345		-0.130	
	(1871.445)		(0.222)	
SM (-4)	1388.745		0.272	
	(1822.705)		(0.209)	
SM (-5)	-1320.822		-0.030	
	(1702.090)		(0.195)	
SM (-6)	320.037		0.040	
	(1671.634)		(0.191)	
SM (-7)	24.047		0.077	
	(1617.013)		(0.190)	
SM (-8)	-1400.794		-0.089	
	(1451.007)		(0.182)	
SM (-9)	-382.467		0.028	
	(1301.999)		(0.166)	
PROPS (-1)	-0.768	***		
	(0.094)			
PROPS (-2)	-0.546	***		
	(0.115)			
PROPS (-3)	-0.279	**		
	(0.119)			

(Table continues on the following page.)

Table 5.3: Liquidity and Sentiment: Autoregressive Distributed Lag Models with Alternative Liquidity Measures (continued)

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)	

(Table continues on the following page.)

Table 5.3: Liquidity and Sentiment: Autoregressive Distributed Lag Models with Alternative Liquidity Measures (continued)

Macroeconomic 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 5.3 reports findings 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 facilitating the total number of traded properties (*PROPS*). Column 2 shows the coefficients for the total number of traded units in bn (*UNITS*). Standard errors are reported in brackets underneath the coefficient estimates. The contemporary value and 9 lags of *SM* were used together with 9 autoregressive terms of either *PROPS* or *UNITS* in both regressions. The *PROPS (UNITS)* regression furthermore facilitates 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. The sample period is 2006:M11 to 2018:M12.

5.9 References

Acharya, V. V. and Pedersen, L. H. (2005), "Asset Pricing with Liquidity Risk", *Journal of Financial Economics*, Vol. 77 No. 2, pp. 375–410.

Ametefe, F., Devaney, S. and Marcato, G. (2016), "Liquidity: A Review of Dimensions, Causes, Measures, and Empirical Applications in Real Estate Markets", *Journal of Real Estate Literature*, Vol. 24 No. 1, pp. 1–29.

Amihud, Y. (2002), "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects", *Journal of Financial Markets*, Vol. 5 No. 1, pp. 31–56.

Baker, M. and Stein, J. C. (2004), "Market Liquidity as a Sentiment Indicator", *Journal of Financial Markets*, Vol. 7 No. 3, pp. 271–299.

Brounen, D., Eichholtz, P. and Ling, D. (2009), "The Liquidity of Property Shares: An International Comparison", *Real Estate Economics*, Vol. 37 No. 3, pp. 413–445.

Chen, H., De, P., Hu, Y. and Hwang, B.-H. (2014), "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media", *Review of Financial Studies*, Vol. 27 No. 5, pp. 1367–1403.

Chordia, T., Roll, R. and Subrahmanyam, A. (2000), "Commonality in Liquidity", *Journal of Financial Economics*, Vol. 56 No. 1, pp. 3–28.

Clayton, J., Ling, D. C. and Naranjo, A. (2009), "Commercial Real Estate Valuation: Fundamentals Versus Investor Sentiment", *The Journal of Real Estate Finance and Economics*, Vol. 38 No. 1, pp. 5–37.

Clayton, J., MacKinnon, G. and Peng, L. (2008), "Time Variation of Liquidity in the Private Real Estate Market: An Empirical Investigation", *Journal of Real Estate Research*, Vol. 30 No. 2, pp. 125–160.

DeLong, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990), "Noise Trader Risk in Financial Markets", *Journal of Political Economy*, Vol. 98 No. 4, pp. 703–738.

Demsetz, H. (1968), "The Cost of Transacting", *The Quarterly Journal of Economics*, Vol. 82 No. 1, pp. 33–53.

Devaney, S. and Scofield, D. (2015), "Liquidity and the Drivers of Search, Due Diligence and Transaction Times for UK Commercial Real Estate Investments", *Journal of Property Research*, Vol. 32 No. 4, pp. 362–383.

Fisher, J., Gatzlaff, D., Geltner, D. and Haurin, D. (2003), "Controlling for the Impact of Variable Liquidity in Commercial Real Estate Price Indices", *Real Estate Economics*, Vol. 31 No. 2, pp. 269–303.

Freybote, J. and Seagraves, P. A. (2018), "The Impact of Investor Sentiment on Commercial Real Estate Market Liquidity", *Journal of Real Estate Research*, Vol. 40 No. 4, pp. 597–627.

Glascock, J. and Lu-Andrews, R. (2014), "An Examination of Macroeconomic Effects on the Liquidity of REITs", *The Journal of Real Estate Finance and Economics*, Vol. 49 No. 1, pp. 23–46.

Glosten, L. R. and Milgrom, P. R. (1985), "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics*, Vol. 14 No. 1, pp. 71–100.

Hausler, J., Ruscheinsky, J. and Lang, M. (2018), "News-Based Sentiment Analysis in Real Estate: A Machine Learning Approach", *Journal of Property Research*, Vol. 35 No. 4, pp. 344–371.

Investment Property Forum (2004), "Liquidity in Commercial Property Markets", Working Paper Five: Liquidity - Findings and Recommendations, Investment Property Forum, London, 2004.

Kyle, A. S. (1985), "Continuous Auctions and Insider Trading", *Econometrica*, Vol. 53 No. 6, pp. 1315–1335.

Liu, S. (2015), "Investor Sentiment and Stock Market Liquidity", *Journal of Behavioral Finance*, Vol. 16 No. 1, pp. 51–67.

Newey, W. K. and West, K. D. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, Vol. 55 No. 3, pp. 703–708.

Pastor, L. and Stambaugh, R. F. (2003), "Liquidity Risk and Expected Stock Returns", *Journal of Political Economy*, Vol. 111 No. 3, pp. 642–685.

Pesaran, M. H. and Shin, Y. (1998), "An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis", in Strøm, S. (Ed.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, *Econometric Society monographs*, Cambridge University Press, Cambridge, pp. 371–413.

Pesaran, M. H., Shin, Y. and Smith, R. J. (2001), "Bounds Testing Approaches to the Analysis of Level Relationships", *Journal of Applied Econometrics*, Vol. 16 No. 3, pp. 289–326.

Porter, M. F. (1980), "An Algorithm for Suffix Stripping", *Program*, Vol. 14 No. 3, pp. 130–137.

Scofield, D. (2013), "Time to Completion Liquidity in UK Commercial Real Estate Investment: 2000-2008", *Journal of European Real Estate Research*, Vol. 6 No. 1, pp. 34–47.

Tieleman, T. and Hinton, G. E. (2012), "Lecture 6.5-Rmsprop, Coursera: Neural Networks for Machine Learning", Technical Report.

Uysal, A. K. and Gunal, S. (2014), "The Impact of Preprocessing on Text Classification", *Information Processing & Management*, Vol. 50 No. 1, pp. 104–112.

6 Conclusion

6.1 Executive Summary

Back in 2009, in the immediate aftermath of the Great Financial Crisis, the winners of the *Nobel Memorial Prize in Economic Sciences* George Akerlof and Robert J. Shiller posed a crucial question with respect to the influence of stories on human behavior: "But what if stories themselves move markets? […] Then economists have gone overboard. The stories no longer merely *explain* the facts; they *are* the facts." (Akerlof and Shiller, 2009, p. 54). Proceeding from this question, the main focus of this dissertation is on assessing the influence of textual sentiment extracted by means of novel machine- and deep-learning procedures with respect to different dimensions of real estate markets. Therefore, in a series of four research papers, different news corpora were analyzed facilitating several sentiment classifiers, varying both in complexity and precision. In order to provide an overall picture, the following sections will carry out a comprehensive summary of the main findings of each individual research paper. Subsequently, results are aggregated to highlight congruencies, limitations as well as further research opportunities.

Paper 1 | On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination

Constituting the opening paper of the series, "On the Relationship between Market Sentiment and Commercial Real Estate Performance – A Textual Analysis Examination" represents the first attempt to quantify a potential relationship between media-expressed sentiment and the performance of private commercial real estate in the United States. Beforehand, a comparable study was conducted for the securitized real estate market only (see Ruscheinsky *et al.*, 2018). To measure the level of market sentiment, the study relies on abstracts of news articles published in the *Wall Street*

Journal from January 2001 to December 2016 containing the keyword “real estate”, “REIT” or both. In order to extract text-inherent sentiment, the study applies a sentiment dictionary, originally developed by Loughran and McDonald (2011) for the field of general finance and later on augmented with real estate terms by Ruscheinsky *et al.* (2018). Abstract sentiment scores are aggregated for each quarter in the form of an absolute and weighted positive-negative-ratio – the latter one implicitly accounting for the relative strength of sentiment expressed – and regressed on total returns of the *NCREIF* index representing quarterly performance of direct commercial real estate in the US. Hereby, the paper accounts for a possible bi-directional relationship and different timings of the relation by estimating a multiple linear regression (MLR) as well as a vector autoregressive (VAR) framework. Additionally, the behavior in solely positive and negative return periods is examined.

With respect to the aforementioned link between media sentiment and market returns, MLR results overall suggest a leading relationship of the textual sentiment indicator by two quarters. When facilitating weighted positive-negative sentiment ratios, the impact on returns is even more pronounced. These findings are confirmed within the estimated VAR framework when considering a circular link. The sentiment indicator Granger-causes market returns at a 1% level, while the opposite cannot be stated. In the course of robustness tests, a leading relationship to capital appreciations is determined by ignoring the income component of *NCREIF* returns. Interestingly, the analysis reveals a particularly strong predictive power in down-market quarters. This finding will be reconsidered in study 3 of the dissertation. Overall, paper 1 provides convincing evidence of the importance of media-expressed sentiment in the US direct real estate market. The promising results thus led to the decision to pursue the idea and probe more complex sentiment classifiers as well as to delve into increasingly sophisticated relationships between textual sentiment and real estate markets in the following three studies.

Paper 2 | News-Based Sentiment Analysis in Real Estate: A Machine-Learning Approach

In this regard, paper 2 “News-Based Sentiment Analysis in Real Estate: A Machine Learning Approach” refines and reexamines the findings of paper 1 by applying a more advanced sentiment classifier, a new and thematically broader text corpus and by studying the influence within both the securitized and direct commercial real estate market. This research setup aims at assessing a more powerful sentiment extraction procedure by facilitating a support vector machine but also on simultaneously examining the sentiment-return relationship in both markets instead of relying on two separate studies. The novel text corpus furthermore is used to confirm the robustness of the antecedent study’s findings. Hence, textual sentiment is extracted from expertise news articles provided by the *S&P Global Intelligence Database* which contain the keyword “real estate” and were published between 2005 and 2016. In contrast to paper 1, headlines were studied to investigate the inherent sentiment of very short and concise text documents. Due to restricted news data availability, the observation period of paper 1 and paper 2 do not provide a perfect match. Nevertheless, both studies cover the boom market of 2005 to 2007, the Great Financial Crisis and the prolonged recovery period afterwards. The relative timely congruence should thus allow for a comparison of their results. In this regard, estimations are made once more within a VAR framework. However, as monthly data is used, *NCREIF* returns have to be replaced by the *CoStar Commercial Repeat-Sales Index* for direct and the *FTSE/NAREIT All Equity REIT Total Return Index* for listed real estate in the United States. By controlling for alternative, well-recognized sentiment indicators in order to quantify the extent to which different measures overlap and the additional application of a pure optimism (*OI*) and pessimism indicator (*PI*) besides the neutral sentiment quotient (*SQ*), paper 2 provides several novelties.

With respect to research questions 1 and 2, challenging a leading relationship with respect to market returns of direct and listed real estate, results indicate predictive power for both markets in accordance with previous literature on the general stock market and also the first paper of this series. The *PI* as well as the *SQ* lead *NAREIT* returns by one month, even when macroeconomic controls are included in the regression analyses. More precisely, *PI* Granger-causes securitized market returns at a 5% level of significance. The impact on direct market returns is of comparable quality, but slightly delayed. The (2nd) and 3rd lag of the *PI* as well as of the *SQ* significantly explain future market returns. However, in none of both markets, a reverse influence – past market performance on sentiment – was found. The one-sided relationship is in line with the main findings of paper 1. Notably, *PI* retains impact and significance when controlling for an aggregate measure of other sentiment indicators and shows a more timely impact on *NAREIT* returns as opposed to those alternatives. This encouraging finding indicates the capability of a machine-learning-based classifier to provide a real-estate specific sentiment indicator for market participants. As an interesting side aspect, the measure focusing on negative news, i.e. the *PI*, delivers most consistent results in the regressions, with that affirming research question 4 of a different reaction of market participants to negative news.

Paper 3 | On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach
Paper 4 | Artificial Intelligence, News Sentiment and Property Market Liquidity

Based on these results, paper 3 and paper 4 attempt a couple of major steps ahead. Both draw on distant supervision-labelled investment ideas from the financial content service *Seeking Alpha* to train an artificial neural network (ANN). Hence, for the first time in real estate, a deep-learning approach is used to classify an extended and up-to-date text corpus provided once more by the *S&P Global Intelligence Database*. Using external training data further makes the approach additionally independent from human intervention and thus avoids man-made errors. Furthermore, ANNs can incorporate a much richer information structure when classifying textual data in comparison to a SVM approach and especially sentiment dictionaries. The approach is therefore the most versatile and powerful of all classifiers tested. In order to capitalize on the semantic capabilities of the deep learning approach, the papers refrain from using abstracts or headlines only and make use of full news articles. Beyond that, news up to December 2018 are covered and a novel sentiment aggregation method is applied.

While paper 3 "On the Predictive Potential of Investor Sentiment: A Deep-Learning Approach" assesses predictive potential with respect to the direct real estate market in the United States, paper 4 "Artificial Intelligence, News Sentiment and Property Market Liquidity" sheds light on the link between sentiment and the trading liquidity of the direct US real estate market. In the spirit of paper 2, a monthly analysis is conducted for both studies. With respect to the chosen econometric approach, the third paper relies on VAR models, Markov-switching (MS) and logit regressions to investigate the indicator's behavior in different market regimes and up-and down-market periods. Furthermore, actual in- and out-of-sample forecasts are conducted. In contrast, paper 4 uses an autoregressive distributed lag model to evaluate the impact of sentiment on market liquidity.

Similar to the first study, paper 3 suggests a different influence of the deep-learning-based indicator during boom and bust periods of the market. Although textual sentiment significantly explains market returns over the whole sample period, the relationship is more pronounced in negative return months, while statistically non-existing in an up-market sample. This behavior can also be shown within the MS regressions. However, when forecasting market returns, the indicator struggles with capturing sudden market swings. This issue could possibly be resolved by improved training or filtering of the sentiment time series by facilitating a larger text corpus. Notably, the indicator's influence – most distinct in lags 6 and 7 of the VAR model – is more in line with the finding of a leading relationship of 2 quarters in paper 1 instead of 2 to 3 months in paper 2.

With respect to market liquidity examined in paper 4, the depth, resilience and breath dimensions of the US direct commercial real estate market are investigated by regressing the sentiment indicator on Amihud's (2002) price impact measure as well as monthly property transaction volume. In order to evaluate the robustness of the results, the number of transacted properties and transacted units are used. The findings support the theoretically derived relationship of a negative relationship between sentiment and price impact, as well as a positive relationship with trading volume as posited by Baker and Stein (2004). In this regard, paper 4 evidently confirms the predictive potential of a textual sentiment indicator with respect to the liquidity dimension of real estate markets. The paper therefore provides a valuable supplement to the first, second and third study and further highlights the dire need to assess the indicators predictive quality with respect to other dimensions of the market beyond the presented ones.

6.2 Final Remarks

In an attempt to boil down the past five chapters in one question, one could ask with Antweiler and Frank (2004): "Is all the talk just noise […]?" Answering this bold and simple question with respect to US real estate news articles was this research project's interest. In the spirit of pioneering studies on textual analysis in accounting and finance, such as those of the previously cited authors Antweiler and Frank (2004), Das and Chen (2007), Tetlock (2007) and Loughran and McDonald (2011), a combined effort over four studies was undertaken to provide a concise but also nuanced answer. Drawing on three different sentiment classifiers, namely a sentiment dictionary, a machine-learning and a deep-learning approach as well as a range of econometric approaches and specialized research questions, the answer to the question has to be "no". This talk is definitely not just noise and influences returns and liquidity in direct and also partially in securitized real estate markets. All three studies on aggregated market performance in the United States confirm this notion with respect to returns despite relying on different sets of text data, three different text sections – abstracts, headlines and full articles – and three different sentiment classifiers. Although the timely impact on market returns deviates, the studies form a concordant and conclusive portrayal of the sentiment-return relationship. Considering the additional findings of the fourth paper, the quadruplet of studies highlight the vast potential of textual sentiment indicators with regard to explaining the behavior of real estate markets. Despite their initial complexity, the results especially point to the power of machine- and deep-learning sentiment classifiers in providing researchers and market practioneers with leading market indicators.

However, this research topic is far from being sufficiently covered and a number of future research opportunities is obvious. Although related research in mainstream finance is highly developed (see Kearney and Liu, 2014 and Loughran and McDonald, 2016 for two comprehensive surveys), the topic is barely covered in real estate. To the best of the author's knowledge, neither volatility models nor the impact on rents and cap rates nor sentiment-based trading strategies are published to date. This is despite textual sentiment might ease decision-making of real estate executives or foster the evaluation of past business strategies. The lack of real estate specific sentiment indices also leaves room for an additional research effort. The techniques proposed within the four studies would allow the construction of easily adaptable, quick-reacting indicators for regional markets or certain subsectors of real estate. Especially studies within the housing market – so far only covered by some scholars (see e.g. Walker, 2014; Soo, 2015 and Nowak and Smith, 2017) – with more non-professional players and a presumably higher sensitivity towards sentiment, might benefit.

In more detail, the aforementioned four research studies could be refined with respect to the influence of market transparency and the time-dependency of results. While fully elaborating the ideas is not intended, some words might suffice to guide future research. Despite the US market being regarded as one of the most transparent ones in the world, its susceptibility to textual sentiment was demonstrated. Presumably the influence might be even more pronounced in less information-efficient markets. Capturing respective dynamics in other countries is therefore certainly worth investigating. Additionally, due to an ongoing professionalization of the real estate industry, the impact of textual sentiment might deviate over time. In the future, extending sample periods and examining changes in the impact of textual sentiment might be thus considered worthwhile. The influence might even be dependent on different phases of the market cycle.

In this regard, by providing a first impression of the power and success of machine- and deep-learning approaches for textual sentiment analysis in real estate, this dissertation wishes to encourage other researchers to pursue the topic, come up with additional research ideas and even more motivating results.

6.3 References

Akerlof, G. A. and Shiller, R. J. (2009), *Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*, Princeton Univ. Press, Princeton, N.J.

Amihud, Y. (2002), “Illiquidity and Stock Returns: Cross-Section and Time-Series Effects”, *Journal of Financial Markets*, Vol. 5 No. 1, pp. 31–56.

Antweiler, W. and Frank, M. Z. (2004), “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards”, *The Journal of Finance*, Vol. 59 No. 3, pp. 1259–1294.

Baker, M. and Stein, J. C. (2004), “Market Liquidity as a Sentiment Indicator”, *Journal of Financial Markets*, Vol. 7 No. 3, pp. 271–299.

Das, S. R. and Chen, M. Y. (2007), “Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web”, *Management Science*, Vol. 53 No. 9, pp. 1375–1388.

Kearney, C. and Liu, S. (2014), “Textual Sentiment in Finance: A Survey of Methods and Models”, *International Review of Financial Analysis*, Vol. 33, pp. 171–185.

Loughran, T. and McDonald, B. (2011), “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks”, *The Journal of Finance*, Vol. 66 No. 1, pp. 35–65.

Loughran, T. and McDonald, B. (2016), “Textual Analysis in Accounting and Finance: A Survey”, *Journal of Accounting Research*, Vol. 54 No. 4, pp. 1187–1230.

Nowak, A. and Smith, P. (2017), “Textual Analysis in Real Estate”, *Journal of Applied Econometrics*, Vol. 32 No. 4, pp. 896–918.

Ruscheinsky, J. R., Lang, M. and Schäfers, W. (2018), “Real Estate Media Sentiment through Textual Analysis”, *Journal of Property Investment & Finance*, Vol. 36 No. 5, pp. 410–428.

Soo, C. K. (2015), “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, 2015.

Tetlock, P. C. (2007), “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”, *The Journal of Finance*, Vol. 62 No. 3, pp. 1139–1168.

Walker, C. B. (2014), “Housing Booms and Media Coverage”, *Applied Economics*, Vol. 46 No. 32, pp. 3954–3967.