

Measuring sentiment in real estate - a comparison study

Article

Accepted Version

Nanda, A. and Heinig, S. (2018) Measuring sentiment in real estate - a comparison study. *Journal of Property Investment and Finance*, 36 (3). pp. 248-258. ISSN 1463-578X doi: <https://doi.org/10.1108/jpif-05-2017-0034> Available at <https://centaur.reading.ac.uk/72893/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1108/jpif-05-2017-0034>

Publisher: Emerald Group Publishing Limited

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1. Introduction

Understanding the role of investor sentiment is not new to economics and finance literature. While the role of sentiment is well understood but its measurement and impact under complex, informationally-asymmetric economic environments are not straightforward, as it can differ for different types of agents at various times across different asset classes. Any causal analysis will, therefore, suffer from substantial endogeneity issues leading to biased estimates and wrong statistical inferences. This is especially problematic in commercial real estate due to severe information asymmetry. The nature of the global real estate investment, especially due to the lessons learnt during the Global Financial Crisis, requires a deeper understanding of the risk-taking behaviour and the associated determinants. Much of the volatility can be captured through examining cap rate or yield, in which sentiment can play a significant role. Sentiment can lead to quick, major changes in risk premia and subsequent deflation in property values.

The study of sentiment is profoundly under-researched in real estate pricing models and the existing real estate investment literature. In this study, we demonstrate how the capturing sentiment in yield model can significantly improve explanatory power. We use the London West End office real estate market as our testing ground. We present three different ways of extracting sentiment in this paper and apply the sentiment indicators to a standard yield model. The first method uses indirect sentiment proxies, which mirror the market development through macroeconomic indicators. The second method utilises online search volume data, as a tool to approximate the beliefs of market participants. The last procedure presents a new and convenient way to mirror the market sentiment. We apply computational linguistic methods to extract the sentiment from various market reports. Our results suggest that the models benefit from the additional information which comes from the sentiment

indicators. All three methods outperform the base model. However, online search volume data is performing best among the three approaches and reaches the highest r-square value in comparison.

We have organised the paper as follows. In the next section, we review the relevant literature and situate our hypotheses within the literature. We then describe our methodology and our data. Finally, we present the empirical analysis and conclude with a summary of key findings in the last section.

2. Literature review

Originally established in behavioural finance and economics sentiment analysis is now a well-recognised subject in various fields. The main difference to classical finance theories is that irrational behaviour of market participants is seen as an appropriate mirror for sentiment. The academic literature is somewhat divided between two types of sentiment measures; indirect measures that are mostly based on solid economic indicators such as closed-end fund discounts, liquidity metrics, trading volumes, the interest rate, GDP etc. direct measures that draw on primary research using surveys and interviews. Lately, new forms of sentiment measures have also emerged. Much improved computing ability and widespread use of online search engines, online search volume data provided by Google Trends (GT)¹ can be a rich source of information. Several papers have by now made use of GT data in addressing the relevant research questions (see, for example, Choi and Varian (2009), Preis et al. (2010), Wu and Brynjolfsson (2013), Loughlin and Harnisch (2013)).

¹ Google Trends Data 2016. Google, <https://www.google.co.uk/trends/>, accessed on 05.05.2016

Traditionally, the University of Michigan Consumer Sentiment Index has been used by several studies to show improvement in explanatory power (Carroll et al. (1994), Bram and Ludvigson (1998), Howrey (2001), Easaw and Heravi (2004))

The majority of sentiment focused studies in real estate, such as Goodman (1994), Weber and Devaney (1996), Dua (2008), Nanda (2007), Croce and Haurin (2009) or Ling et al. (2014) tend to concentrate on the residential sector. Baker and Saltes (2005), Clayton et al. (2009), and Marcato and Nanda (2016) analyse the role of sentiment data in explaining market dynamics in commercial real estate. In a recent work, Das et al. (2015) use the flight to liquidity and category/style investing theory to explain the sentiment-induced trading behaviour of institutional investors. While the presented studies have shown, that sentiment has an impact on the real estate market, only a few studies have used sentiment indices in a yield modelling framework (see Ling et al. 2014; Clayton et al. 2009). Tsolacos (2012) analyses the application of sentiment indicators on the European real estate market, specifically pointing towards how sentiment indicators can improve forecasting ability. He uses the economic sentiment indicator (ESI) provided by the European Union for the three broad markets: Germany, France, and the UK.

Clayton et al. (2009) see the real estate market influenced by sentiment, due to the lack of transparency, illiquidity, short sell constraints and the firm segmentation of the market. In a more recent study, Ling et al. (2014) examine the short-sale constraints in private real estate markets finding stronger effects

Marcato and Nanda (2016) analyse a range of sentiment measures and confirm several results from the earlier studies. To show a gain in forecasting accuracy, the authors used the University of Michigan index, the Architecture Billings Index (ABI) [Baker and Saltes

(2005)], and the Housing Market Index (HMI). Other fields for the application of sentiment analysis are REITs (see Chiang and Lee, 2009; Lin et al. 2009; Das et al. 2015).

Some US studies have applied the GT measures (see Choi et al., 2009 and Beracha et al., 2013) focusing on the housing market. Dietzel, Braun and Schäfers (2014) also used Google Search volume data to improve the ability to explain movements in the commercial real estate market. However, whether forecasting ability can be enhanced or not remains unclear from all these studies.

The application of natural language-based techniques to the real estate field is rather new. Besides the works of Soo (2014) and Walker C. (2014, 2016), it is our understanding, that no other studies have analysed this topic using a similar set of tools. However, the application of text analysis with word lists has been employed in various fields such as marketing (see He, 2012; Chen et al. 2016);, political science (Greene and Haber, 2016). In their studies, Soo and Walker focused on the housing market in the US and the UK. Driven by the underlying idea that news articles try to describe the real world and that readers believe that the description is truthful and adequate, the authors assumed that readers adopt the presented opinion and start to change their behaviour in the long run. Therefore, extracted sentiment from news articles can foreshadow market developments.

This short literature review indicates that sentiment is a useful indicator when it comes to market developments. Besides the focus on the residential market in the literature, the few studies which have analysed the commercial real estate market show, that sentiment should not be neglected in this relationship.

3. Methodology

The literature review revealed that in the last couple of years different approaches for sentiment extraction have emerged. In the following section, we will describe the three chosen methods as well as the selected model in more detail. We start with a well-established specification for the yield model. The yield or the cap rate represents a ratio of the net operating income over the price of the property. Several papers have examined and developed yield model (see, for example, Sivitanides et al. 2001; Hendershott and MacGregor, 2005a and b; Clayton et al., 2009; Chervachidze and Wheaton, 2013; McAllister and Nanda 2015; Devaney et al. 2016). Common determinants are a risk-free rate, the expected rent and a risk premium as represented by equation (1).

$\begin{aligned} Yield_{(office)r,t} = & \beta_0 + \beta_1 Risk\ Free\ Rate_{j,t} + \beta_2 Risk\ Premium_{j,t} \\ & + \beta_3 Rent\ measure_{j,t} + \varepsilon_{j,t} \end{aligned}$	<i>Equation 1</i>
---	-----------------------

Cushman & Wakefield (formerly DTZ) has provided the property yield. The yield is recorded on a quarterly level for the London West End office market. Similar to Clayton et al. (2009), McAllister and Nanda (2015), Sivitanides et al. (2001), as well as Devaney et al. (2016) we use the 10-year government bond rate for the UK as a risk-free rate. For the risk premium, we further follow Devaney et al. (2016), who calculated their risk premium based on the volatility of the equity market as an eight-quarter rolling standard deviation from the stock market return and subtract the national 10-year government bond rate (risk-free rate).

Next, we incorporate sentiment in the yield model. According to Sivitanides et al. (2001) who point out that yields are more volatile than rents and are likely to respond to new developments within the market sooner. While yields can incorporate recent and new developments, without much delay, rents are fixed over a longer period.

A sentiment variable which is able measures the market expectations is believed to improve a yield model significantly. We extend equation 1 with a sentiment indicator (equation 2).

$Yield_{(office)r,t} = \beta_0 + \beta_1 Risk\ Free\ Rate_{j,t} + \beta_2 Risk\ Premium_{j,t} + \beta_3 Rent\ measure_{j,t} + \beta_4 Sentiment_{j,t} + \varepsilon_{j,t}$	<i>Equation</i> 2
--	----------------------

The literature has differentiated between direct and indirect sentiment proxies. It is believed that the direct sentiment proxies can provide a better indication since they use the actual opinion of market participants. The sentiment measures that we present provide a mix of direct and indirect metrics.

In general, we believe that the measured sentiment should have a negative impact on property yields. If sentiment rises, prices rise as well, which leads to a decrease in property yields and, vice versa. Different groups of investors might follow different goals within the market and hence should use different approaches to sentiment. In the following sections, we describe the three sentiment indicators and their construction in more detail.

We assume that macroeconomic sentiment proxies can give information about the market sentiment. Therefore, the first sentiment indicator we construct uses macroeconomic factors. The indicator combines two direct sentiment proxies and four indirect sentiment proxies. Different to other studies, such as Lee et al. (1991); Clayton et al. (2009) or Ling et al. (2014)

we chose a different set of sentiment proxies. All these studies focus on the US REIT market while our study is focusing on the direct real estate market in the UK. Similar to Tsolacos (2012), we use the Economic Sentiment Indicator (ESI) published by the European Commission. ESI summarises five sector-specific confidence indicators using a weighted aggregation of standardised input series. We also use the Business Climate Indicator (BCI), also published by the European Commission. BCI provides a composite indicator for the manufacturing sector in the Euro area.

We choose four indirect sentiment proxies to focus on macroeconomic factors that can capture the effect of sentiment changes in the broader economy. The national stock market index provides information about the national economy. Among others, Baker and Wurgler (2006, 2007), Tetlock (2007), and Kurov (2010) have shown that investor sentiment influences the stock markets. We use the quarterly return from the stock exchange. Further, we use the government bond rate, capturing several country-specific risks, such as inflation, interest rate risk and state of public finances. As has repeatedly been applied in the literature over last few decades, we include a consumer confidence index, which can inform consumer behaviour trends. Also, the credit rating can act as a sentiment indicator, especially from the foreign investors' perspectives.

The above indicators will have a high level of correlations. Moreover, to extract the irrational component, we apply a two-stage process. In a first step, we apply an orthogonalization process to the individual sentiment proxies. Here, we try to remove known macroeconomic influences. We focus on the main factors, which drive the economy, such as the change of GDP, the forecasted change of GDP, the interest rate, the logarithm of the consumer price indicator, the logarithm of consumer spending, the level of unemployment, as well as the

percentage change in the industry production of the country (*c_gdp, fc_gdp, intr, logcpi, logcsp, unemp, indpropc*).

It is necessary to regress each of the proxies against those factors, while the intercept is excluded from the regression. The residuals of these six orthogonalized regressions are then combined to provide a measure of the irrational and unexplainable part.

In the second part, we combine the six individual residual time series into one sentiment index by applying a Principal Component Analysis (PCA). We first standardise the residuals for a better comparison. Because some variables may react to changes in the sentiment more quickly than others, we use both the standardised variable and a lagged version of it in the PCA specification. We have lagged the variables by one period, given that we deal with quarterly data. In total twelve variables are included in the PCA structure. We obtain the first principal component with the highest eigenvalue from this analysis.²

To select whether the lagged or unlagged variable of the six residuals should be included in the final index, we compare the correlation of each factor loading with the first stage index from the PCA. We remove these factor loadings, which have a smaller correlation, from the final sentiment calculation. This leaves some lagged and unlagged variables for the final compilation.

Based on this selection, we construct the final sentiment index. To verify that we have not lost any significant explanatory power by removing the weak factors. The correlation between the first stage index and the constructed sentiment indicator should still be in a reasonable and comparable range.

² In unreported tests, we construct the sentiment indicators following the Kaiser criterion. Here all principal components with an eigenvalue greater than 1 have been used. We then construct a weighted average of the number of recommended components based on their explanation power. However, the results do not suggest that this method is superior.

Next, we analyse online search volume data. Studies [Dietzel et al. (2014), Preis et al. (2010)] have shown that online search volume data can provide a useful indication of the consumers' intentions. It is likely that a significant number of searches are purely for information gathering or general curiosity of learning about current topics within the economy. Nonetheless, some searches can lead to economic activity or transaction, which is what we are hoping to capture.

The company stores the searches and use them to aggregate a search term index over time. Specific category related searches, are only counted by the Google algorithm in that specific category if a series of related searches are performed. We assume that the volume of online searches within the specific category may be able to reflect the sentiment of the market segment and represent a suitable way of measuring the underlying mood. To construct our Google Trends indicator we used a set of 90 specific search words, focusing on the office, retail fundamentals as well as news related to the property service sector firms. The GT tool allows the user to download the time series for each search term. Since London is a major city within England, Google provides a national index as well as an indication about the city-specific share of the national index. To construct the London specific Google Trends sentiment index, we have broken the national series down to the city level. This provides a better indication of the city-specific development. From 90 selected search terms, a total number of 57 words generated a result for London. Finally, we have aggregated all results into a single index.

The last sentiment indicator is based on the quantification of texts. We use a simple text analysis approach called the NRC method. The method has been developed by Mohammad

and Turney (2010) for the National Research Council of Canada (NRC). The developed application uses text documents and identifies the sentiment within each document. NRC belongs to the class of lexicon classifications. Different to other comparable applications, such as BING (Hu and Liu, 2004) or AFINN (Nielsen, 2011), NRC focus on emotionally-charged words. The authors used Amazons Mechanical Turk to categorise a total of 14,182 words into either a positive (2,312 words), a neutral (8,546 words) or a negative category (3,324 words). Each of these categories carries a numerical value -1 for negative, 0 for neutral and 1 for all positive words. A computer algorithm calculates a final score on a document level to sort the whole document into one of the categories.

In this paper, we have used a document corpus of 233 market reports from different service providers (e.g. Jones Lang LaSalle, Savills and Knight Frank) for the London commercial real estate market. Since each market report is published at a certain time, we were able to develop a time series by the generated sentiment values. For some quarters if we were able to collect multiple market reports, we have then calculated an average value.

The NRC application is run in R with the help of the ‘*syuzhet*’ package by Jockers (2016). Before the algorithm is employed to compare the market reports with the word lists, each text document is *cleaned* according to natural language processing standards. For example, numbers and additional white spaces have been removed. Further, all words have been transformed into lower case, and the words have been stemmed.

Exhibit 1 illustrates the three constructed sentiment indicators and their behaviour over time.

[INSERT Exhibit 1]

4. Data description

For this study, we used different datasets. In the following section, we describe the data sources in more detail.

This paper analyses the London West End office market from 2004q1 until 2014q4 (44 quarters). The dataset comprises specific real estate data and a range of macroeconomic variables. Cushman & Wakefield (formerly DTZ) provided the property data. The macroeconomic data was collected via Thomson Reuters Datastream. We further used data from the OECD, the IMF, and the European Commission. The search volume data is available from 2004 onwards. To improve data comparability, the provided data is adjusted across search terms. These results are scaled to range from 0 to 100. Unfortunately, Google does not give the actual numbers for each search query. Besides the possibility of analysing different search terms in the various regions and at various points in time, the application offers a chance to search within different categories. One of these categories is ‘property’ (0–29)³. The categorical filter function eliminates various meanings of words, for better results. However, an explanation of how the algorithm knows that certain words have been searched within this category is missing. Dietzel et al. (2014) explain that the categorisation is based on the individual search behaviour. Each query is placed into a framework of searches, which considers queries before and after the specific search.

For the textual sentiment indicator, we have used a total of 233 market reports related to the London commercial real estate (CRE) market. The CRE topic includes multiple sub-topics such as *capital markets, care homes, commercial real estate, development, E-tailing, hotels, investment, office, retail, rural and student housing*. These reports have been collected from

³ The source code of the Google Trends webpage uses those codes for each of the categories.

BNPRRE, Cushman & Wakefield, CBRE, Colliers, DTZ, Jones Lang LaSalle, Knight Frank and Savills. The data was collected from 2005 to 2014. The reason for this mismatch is the available data on the different websites from the service agencies.

Exhibit 2 illustrates the used data sources, their origin as well as some descriptive statistics.

[INSERT Exhibit 2]

5. Results

In this section, we present our results for the different methods regarding the yield model with and without the various sentiment indicators.

Exhibit 3 reports four models. Model 1 is a standard yield model which largely conforms frequently reported results from the literature. In models 2, 3 and 4, we add three types of sentiment indicators to the standard yield model from column 1. All independent variables are significant, and the r-square value reaches a value of 0.297.

First, we add the macroeconomic sentiment indicator. Our theoretical assumption that the sentiment has an adverse impact on the property yield has been confirmed within all three models (the macroeconomic, the Google Trends and the textual sentiment models). While all the other model components are highly significant, the rent variable has become insignificant. The number of observations has dropped by one observation due to the construction of the indicator. The macroeconomic sentiment reaches an r-squared value of 0.485 and outperforms the base model.

The second indicator relies on the standardised online search volume from Google searches. All model components are highly significant, and the r-square value is more than twice the size of the base model (0.699).

The textual sentiment indicator shows a similar picture, with significant variables and an r-squared value of 0.586. Different to the other models, the number of observation has dropped to 27, due to data availability.

All sentiment coefficients are significant up to a 5% level. The size of the effects differ significantly – the macroeconomic sentiment effect is about -0.052. The other two indicators

have an effect size of -0.205 and -0.311 respectively. This seems to suggest the greater ability of the GT and text-based sentiment indicators to capture the effects.

[INSERT Exhibit 3]

An advantage of the latter two indices lies in the time component. The macroeconomic indicator relies on different sentiment proxies, which need to be published and collected first. Same accounts for the textual sentiment indicator presented here. Therefore, both use past data and can only be reproduced after their source data has been published. Especially the textual sentiment indicator relies on the publication of market reports, which across different asset classes and service firms, are released on an irregular basis. However, this issue could be overcome with a different underlying source, since the displayed result suggests, that the wording in the market reports carries enough information to conclude the market movement. This is not surprising given the very nature of the market report.

The online search volume indicator, on the other hand, allows for an easy access and a simple extraction of the sentiment. Given the positive results of this study, online search volume should be considered as a valid source. The main advantage of the indicator can be found in the time aspect. The quarterly frequency in this study could be adjusted down to a monthly, weekly and nowadays even daily level. Market movements can be anticipated nearly simultaneously, and investors and market participants could react accordingly.

6. Conclusion

Price determination in the property market is often fraught with irrational behaviour and trends that can not be explained by fundamental drivers of the sectors. Such disconnect can lead to under or over-pricing of assets which then culminate into cycles and frequent fluctuations. This makes projections around investment very difficult. This applies in particular to the property sector with thin markets and infrequent transactions.

In this paper, we show that traditional forecasting models can benefit from the inclusion of sentiment indicators. The model performance is improved, and forecasting ability is stronger by such integration. We have showcased this through well-established yield modelling.

Moreover, we also contribute to the literature by demonstrating innovative ways of using new forms of data to construct sentiment indicators. Computational linguistic methods have been used to construct sentiment indices.

We believe that such tools can be an integral part of the repertoire of sentiment analysis methods. Future studies should focus on refining the underlying data for sentiment construction.

References

- Baker, K., Saltes, D., 2005. Architecture billings as a leading indicator of construction. *Business Economics* 40 (4), 67–73.
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61 (4), 1645–1680.
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129–151.
- Barkham, R.J., Ward, C.W.R., 1999. Investor sentiment and noise traders: Discount to net asset value in listed property companies in the UK. *Journal of Real Estate Research* 18 (2).
- Beracha, E., Wintoki, M.B., 2013. Forecasting residential real estate price changes from online search activity. *Journal of Real Estate Research* 35 (3).
- Bram, J., Ludvigson, S., 1998. Does consumer confidence forecast household expenditure? A sentiment index horse race. *Federal Reserve Bank of New York Economic Policy Review* 4, (2).
- Carroll, C.D., Fuhrer, J.C., Wilcox, D.W., 1994. Does consumer sentiment forecast household spending? If so, why? *The American Economic Review* 84 (5), 1397–1408.
- Case, K.E., Shiller, R.J., 1989. The efficiency of the market for single-family homes. *The American Economic Review* 79 (1), 125–137.
- Chervachidze, S., Wheaton, W., 2013. What determined the great cap rate compression of 2000–2007, and the dramatic reversal during the 2008–2009 financial crisis? *The Journal of Real Estate Finance and Economics* 46 (2), 208–231.
- Chiang, K.C.H., Lee, M.-L., 2009. The role of correlated trading in setting REIT prices. *The Journal of Real Estate Finance and Economics* 39 (4), 450–471.

- Chichernea, D., Miller, N., Fisher, J., Sklarz, M., White, B., 2008. A cross-sectional analysis of cap rates by MSA, *Journal of Real Estate Research* 30 (3), 249–292.
- Choi, H., Varian, H., 2009. Predicting the present with Google Trends. Google Inc.
- Clayton, J., Ling, D.C., Naranjo, A., 2009. Commercial real estate valuation: Fundamentals versus investor sentiment. *The Journal of Real Estate Finance and Economics* 38 (1), 5–37.
- Croce, R.M. and Haurin, D.R., 2009, Predicting turning points in the housing market. *Journal of Housing Economics*, 18 (4), 281-293.
- Das, P.K., Freybote, J., 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* 51 (2), 160–189.
- Devaney, S., Livingstone, N., McAllister, P. and Nanda, A., 2016, IPF report - "Unravelling Liquidity In International Commercial Real Estate Markets", published by Investment Property Forum (IPF).
- Dietzel, M.A., Braun, N., Schäfers, W., 2014. Sentiment-based commercial real estate forecasting with Google Search volume data. *Journal of Property Investment & Finance* 36 (6) 540–569.
- DiPasquale, D., Wheaton, W.C., 1992. The cost of capital, tax reform, and the future of the rental housing market. *Journal of Urban Economics* 31 (3), 337–359.
- Dua, P., 2008, Analysis of consumers' perceptions of buying conditions for houses. *The Journal of Real Estate Finance and Economics*, 37 (4), 335-350.
- Duca, J.V., Ling, D.C., 2015. The other (commercial) real estate boom and bust: The effects of risk premia and regulatory capital arbitrage. Federal Reserve Bank of Dallas Research Department Working Paper 1504.

- Easaw, J.Z., Heravi, S.M., 2004. Evaluating consumer sentiments as predictors of UK household consumption behaviour – Are they accurate and useful? *International Journal of Forecasting* 20 (4).
- Fan, C.S., Wong, P., 1998. Does consumer sentiment forecast household spending? The Hong Kong case. *Economics Letters by Elsevier* 58 (1), 77–84.
- Finn, Årup Nielsen (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs, DTU Informatics, Technical University of Denmark, Lyngby, Denmark, arXiv preprint arXiv:1103.2903.
- Freybote, J., 2016. Real estate sentiment as information for REIT bond pricing. *Journal of Property Research*, 1–19.
- Freybote, J., Seagraves, P.A., 2016. Heterogeneous investor sentiment and institutional real estate investments. *Real Estate Economics*.
- Goodman, John L. Jr., 1994, Using Attitude Data to Forecast Housing Activity, *The Journal of Real Estate Research* 9 (4), 445-453
- Hendershott, P.H., MacGregor, B.D., 2005a. Investor rationality: Evidence from UK property capitalization rates. *Real Estate Economics* 33 (2), 299–322.
- Hendershott, P.H., MacGregor, B.D., 2005b. Investor rationality: An analysis of NCREIF commercial property data. *Journal of Real Estate Research* 27 (4), 445–475.
- Hengelbrock, J., Theissen, E., Westheide, C., 2013. Market response to investor sentiment. *Journal of Business Finance & Accounting* 40 (7&8), 901–917.
- Howrey, E.P., 2001. The predictive power of the index of consumer sentiment. *Brookings Papers on Economic Activity*, 2001 (1), 175–207.
- Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). ACM.

- Hutchison, N., Fraser, P., Adair, A. and Srivatsa, R., 2012, Regime shifts in ex post UK commercial property risk premiums, *Journal of Property Research*, 29 (3), 247-269
- Jin, C., Soydemir, G., Tidwell, A., 2014. The US housing market and the pricing of risk: Fundamental analysis and market sentiment. *Journal of Real Estate Research*, 36 (2).
- Jockers, M. (2016). Package 'syuzhet'.
- Katona, G., 1968, Consumer behavior: Theory and findings on expectations and aspirations. *The American Economic Review*, 58(2), 19-30.
- Kurov, A., 2010. Investor sentiment and the stock market's reaction to monetary policy. *Journal of Banking & Finance* 34 (1), 139–149.
- Lee, C.; Shleifer, A.; Thaler, R.H. 1991. Investor sentiment and the closed-end fund puzzle. *Journal of Finance* 46 (March), 75–110.
- Lin, C.Y.; Rahman, H., Yung, K., 2009. Investor sentiment and REIT returns. *The Journal of Real Estate Finance and Economics*, 39 (4), 450–471.
- Ling, D.C., Naranjo, A., Scheick, B., 2014. Investor sentiment, limits to arbitrage and private market returns. *Real Estate Economics* 42 (2), 531–577.
- Loughlin, C., Harnisch, E., 2013. The viability of StockTwits and Google Trends to predict the stock market. www.stocktwits.com/research/Viability-of-StockTwits-and-Google-Trends-Loughlin_Harnisch.pdf.
- Malgarini, Marco; Margani, Patrizia (2007) Psychology, consumer sentiment and household expenditures, *Applied Economics*, Volume 39, Issue 13, pp. 1719-1729
- Marcato, G. and Nanda, A., 2016, Information Content and Forecasting Ability of Sentiment Indicators: Case of Real Estate Market, *Journal of Real Estate Research* 38 (2)
- McAllister, P. and Nanda, A., 2015, Does foreign investment affect US office real estate prices?, *The Journal of Portfolio Management*, 41 (6), pp. 38-47.

- Mohammad, Saif M.; Turney, Peter D. (2010). Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon, In Proceedings of the NAACLHLT 2010 workshop on computational approaches to analysis and generation of emotion in text, pp. 26-34. Association for Computational Linguistics.
- Nanda, A. 2007, Examining the NAHB/Wells Fargo Housing Market Index (HMI). Housing Economics.
- Preis, T., Reith, D., Stanley, E., 2010. Complex dynamics of our economic life on different scales: Insight from search engine query data. Philosophical Transactions of the Royal Society A 2010 (368), 5707–5719. doi:10.1098/rsta.2010.0284
- Sivitanides, P., Southard, J., Torto, R.G., Wheaton, W.C., 2001. The determinants of appraisal-based capitalization rates, Real Estate Finance, Torto Wheaton Research.
- Sivitanidou, R., Sivitanides, P., 1999. Office capitalization rates: Real estate and capital market influences. Journal of Real Estate Finance and Economics 18 (3), 297–322.
- Soo, C. K. (2014). QUANTIFYING ANIMAL SPIRITS.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance 62 (3), 1139–1168.
- Tsolacos, S., 2012. The role of sentiment indicators for real estate market forecasting. Journal of European Real Estate Research 5 (2), 109–120.
- Walker, C. B. (2014). Housing booms and media coverage. Applied Economics, 46(32), 3954-3967.
- Walker, C. B. (2016). The direction of media influence: Real-estate news and the stock market. Journal of Behavioural and Experimental Finance, 10, 20-31.
- Weber, W. and Devaney, M., 1996, Can consumer sentiment surveys forecast housing starts?. Appraisal Journal, 64, 343-350.

Wu, L., Brynjolfsson, E., 2013. The future of prediction: How Google searches foreshadow housing prices and sales. Working Paper – Draft 2013

Exhibit 1: Sentiment graphs: 2004q1 - 2014q4 (44 quarters)

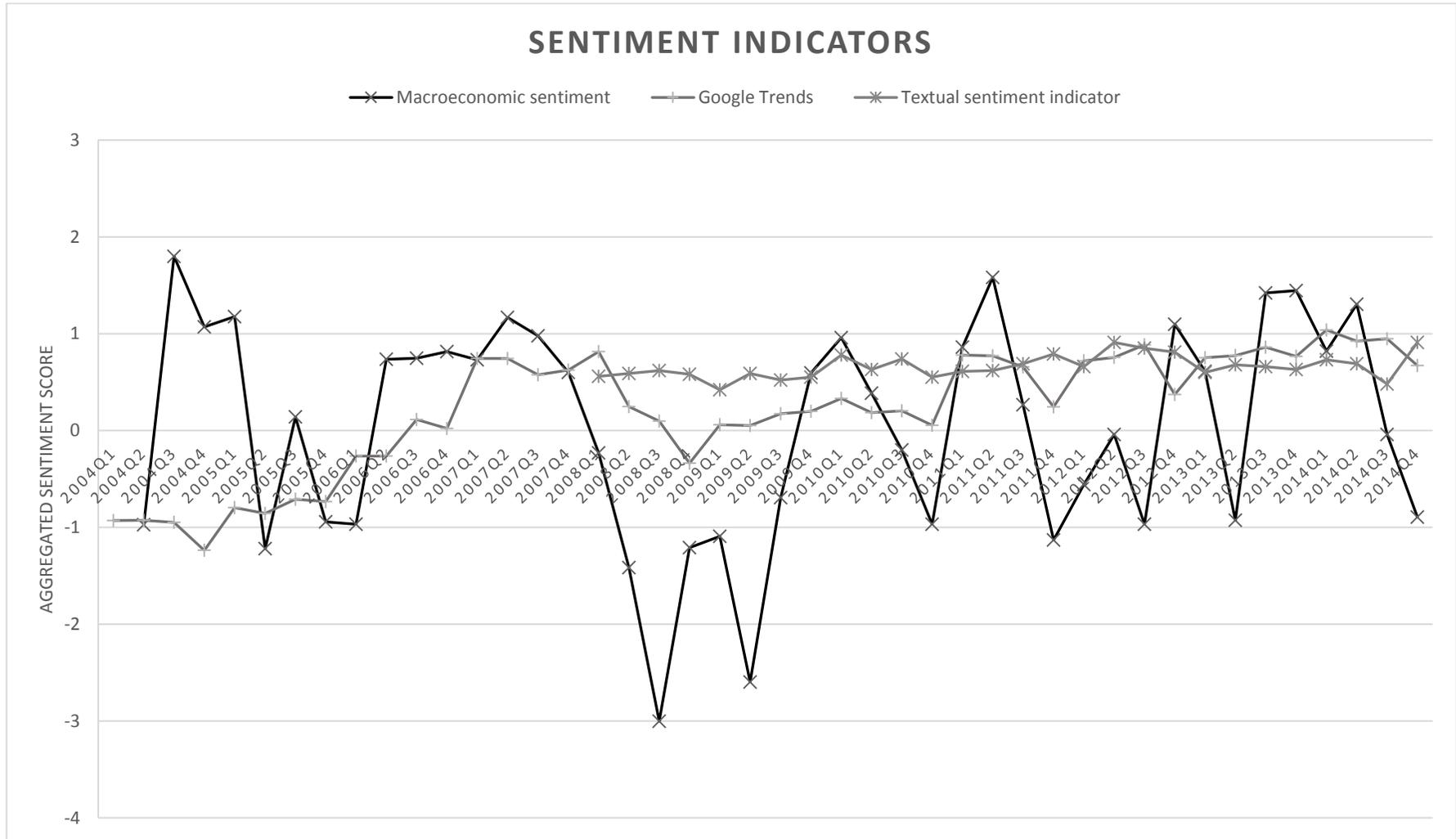


Exhibit 2: Variable Description: 2004q1 - 2014q4 (44 quarters)

LABELS	METHOD	SOURCE	Obs	Mean	Std. Dev.
Log of office yield	Logarithm of the London West End Prime office yield	Cushman-Wakefield data	44	1.5231	0.1254
Office rent 4 quarter moving average	four-quarter moving average of the long-run deviation of the log of real rent	computed based on DTZ/Cushman-Wakefield data	44	0.0013	0.0626
Government bond rate	10 year government bond	Thomson Reuters Datastream	44	3.6980	1.0760
Risk premium	eight-quarter rolling standard deviation from the stock market return and subtract the national 10-year government bond rate (risk free rate)	Computed by authors	44	2.9543	3.9355
Macroeconomic Sentiment	based on an orthogonalization process of 6 different direct and indirect sentiment proxies (ESI, BCI, stock market return, consumer confidence, credit rating, government bond)	Computed by authors	43	0.0285	1.1329
Google Trends	indicator is based on 57 search terms and the related search volume from Google Trends	Google Trends	44	0.2080	0.6330
London CRE Market Reports	one period lagged indicator for all London market reports analysed with the NRC method (233 documents)	all London CRE market reports from BNPPRE, C&W, CBRE, Colliers, DTZ, JLL, KF, Savills (all companies)	27	0.6615	0.1232

Exhibit 3: Regression Models

Dependent variable: log of office (yield)

VARIABLES	Standard yield model (Hendershott)	Standard yield model with Macroeconomic _Sentiment	Standard yield model with Google Trends	Standard yield model with text- based sentiment
Office rent 4 quarter moving average	0.549* [0.276]	0.279 [0.249]	1.744*** [0.249]	1.802** [0.647]
10 year - Government bond rate	0.083*** [0.014]	0.078*** [0.015]	0.026* [0.014]	0.080*** [0.022]
Risk premium	0.018*** [0.004]	0.015*** [0.004]	0.020*** [0.003]	0.025*** [0.009]
Macroeconomic Sentiment		-0.052*** [0.013]		
Google Trends			-0.205*** [0.025]	
London CRE Market Reports				-0.311* [0.175]
Constant	1.166*** [0.057]	1.187*** [0.061]	1.410*** [0.061]	1.326*** [0.180]
Observations	44	43	44	27
R-squared	0.297	0.485	0.699	0.586

Robust standard errors in brackets; *** p<0.01, ** p<0.05, * p<0.1