The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/1463-578X.htm

Outperforming the benchmark: online information demand and REIT market performance

Karim Rochdi and Marian Dietzel
International Real Estate Business School, University of Regensburg, Regensburg, Germany

Received 26 November 2014
Revised 26 November 2014
Accepted 11 December 2014

Abstract

Purpose – The purpose of this paper is to investigate whether there is a relationship between asset-specific online search interest and movements in the US REIT market.

Design/methodology/approach – The authors collect search volume (SV) data from "Google Trends" for a set of keywords representing the information demand of real estate (equity) investors. On this basis, the authors test hypothetical investment strategies based on changes in internet SV, to anticipate REIT market movements.

Findings – The results reveal that people's information demand can indeed serve as a successful predictor for the US REIT market. Among other findings, evidence is provided that there is a significant relationship between asset-specific keywords and the US REIT market. Specifically, investment strategies based on weekly changes in Google SV would have outperformed a buy-and-hold strategy (0.1 percent p.a.) for the Morgan Stanley Capital International US REIT Index by a remarkable 15.4 percent p.a. between 2006 and 2013. Furthermore, the authors find that real-estate-related terms are more suitable than rather general, finance-related terms for predicting REIT market movements.

Practical implications – The findings should be of particular interest for REIT market investors, as the established relationships can potentially be utilized to anticipate short-term REIT market movements.

Originality/value – This is the first paper which applies Google search query data to the REIT market.

Keywords Real estate, REIT, Google Trends, Information demand, Investment strategy, Search query data

Paper type Research paper

1. Introduction

It is common knowledge that an investor's decision about whether to invest in or divest from the stock market is determined by a variety of factors. Besides business-related news, it might be factors like natural disasters, the resignation of business leaders, terrorist attacks, let alone all kinds of economic fundamentals and political reports that make markets fluctuate. No one would seriously claim to be able to foresee such events with any accuracy or reliability. However, broken down to the very basics, the price of a stock is still determined by demand and supply, by one party who is willing to buy and another who is willing to sell. Apart from trading computers, the largest share of all financial transactions is still conducted by human beings who make a buy or sell decision. There can be no doubt that this decision is influenced by the abovementioned events, but in between an event or the release of certain news and a (human) financial transaction, people gather further information. In a world of smart phones, tablets and laptops, the internet has become the main source of this information. Therefore, big data and search query data in particular are becoming increasingly interesting for

Journal of Property Investment & Finance
Vol. 33 No. 2, 2015
pp. 169-195

1463-578X
DOI 10.1108/JPIF-11-2014-0069

those researching equity markets, as it represents an appropriate instrument for quantifying internet users' interests and motives.

In a seminal article, Preis *et al.* (2013a) use Google search volume (SV) data in search-interest-based trading strategies to examine the relationship between online search behavior and stock market movements (Dow Jones Industrial Average, DJIA). They find that online search behavior does indeed serve as an indicator of stock market movements. This raises the question of whether searchers leave more (less) traces online when researching more (less) complex asset classes, since estimating the fair value of a non-transparent good or market is much more time-consuming and elaborate. Thus, we hypothesize that the more research-intensive an asset, that is, the more information is needed before making a buy or sell decision, the better the chances of predicting the searchers' behavior with regard to their subsequent (trans)actions. As the real estate investment trust market constitutes a relatively research-intensive asset, it is a suitable example for further analysis. This, of course, is due to the fact that information-gathering on REITs is considered to be more comprehensive, because both capital and space markets have to be analyzed (Roulac, 1988). Generally, a rational investor analyzes a REIT portfolio thoroughly before making a decision about whether the stock is currently fairly priced. Besides fundamental equity market analysis, this mainly includes appraising the relevant (property) markets and the future development of rental income and yields. This relates to the discussion among researchers on the extent to which REITs behave more similarly to property or stock markets. In order to investigate this issue, as well as the relationship between information demand and the US REIT market in general, we apply a methodology similar to Preis *et al.* (2013a). For this purpose, two groups of keywords are compiled, one containing real-estate-related search terms, the other one (rather general) finance-related search terms. Accordingly, each search term constitutes an individual information-demand-based investment strategy whose trading signals derive from weekly changes in the underlying Google SV. Subsequently, the overall performance of the investment strategies from the two keyword subsets is compared with one another in order to gain knowledge about what kind of information demand predicts the REIT market more successfully. Also, since the past decade was characterized by turbulent markets, the time-specific dynamics of the relevance of information demand are of particular interest. Additionally, if Google search interest is linked to stock trading volume, as found by Preis *et al.* (2010), information demand should be a particularly good predictor during phases of exceptionally high returns or losses, which is why we also determine the strategies' predictive ability for the 40 most extreme upward and downward market movements between 2006 and 2013. All tested investment strategies are benchmarked against a buy-and-hold strategy for the Morgan Stanley Capital International (MSCI) US REIT Index, as well as the DJIA.

This paper contributes in several ways to the literature on the information demand of real estate (equity) investors, real estate equity markets and online search behavior. First and foremost, we find that search query data serve as a successful predictor for the US REIT market. Moreover, the results suggest that asset-specific (real-estate specific) search terms are better predictors for the US REIT market than finance-related search queries. Also, this is the first paper to examine the dynamics of Google Trends investment strategies' (GTIS) investment performance over time. The findings reveal that particularly during the crisis of 2008-2011, a period of substantial investor uncertainty and increased information demand, investment strategies based on the Google data set predict the market very successfully. This is supported by the fact that

GTIS have a much higher hit rate in predicting the 40 most extreme market movements (up to 75 percent). In terms of practical implications, a substantial number of information-demand-based investment strategies would have outperformed the market (MSCI US REIT Index).

The paper is structured as follows. The next section provides a brief overview of the relevant literature, while Section 3 describes the data set and its procurement. Section 4 outlines the methodology by which trading signals for the GTIS are generated, and explains the benchmark strategies, which are then compared against the Google strategies in Section 5. Here, a number of performance and risk measures are introduced on which the analysis is based, and the most important findings are presented. Section 6 concludes.

2. Literature review

In recent years, much research has been conducted on online data from search engines, social networks, internet encyclopedias, microblogging services and image hosting web sites (e.g. Moat *et al.*, 2013; Bordino *et al.*, 2012; Gilbert and Karahalios, 2010; Mao *et al.*, 2011; Bollen *et al.*, 2011; Preis *et al.*, 2013b). As this extensive data set reflects in an unparalleled manner the everyday activities of an increasingly complex society, it enables researchers to make very prompt and well educated guesses from online search activities about the future behavior of users. Particularly Google Trends (GT) data have recently been used in a number of ways since Ginsberg *et al.* (2009) published a groundbreaking article, in which they show how flu-symptom-related internet searches track the spread of the flu virus across the USA in a timely manner. A range of more recent studies focusses on the relationship between Google data and financial markets. It has been shown that the volume of search queries is positively related to trading activity, stock liquidity and volatility (Preis *et al.*, 2010; Bank *et al.*, 2011; Vlastakis and Markellos, 2012; Dimpfl and Jank, 2012; Latoeiro *et al.*, 2013). Da *et al.* (2011a, b) find that an increased SV for stock tickers predicts an increase of the Russell 3,000 stock index within the next two weeks and an eventual price reversal within the same year. Furthermore, they find that specific searches for firms' products predict positive (negative) revenue and earnings surprises. Drake *et al.* (2012) find that information demand for specific stocks starts to increase two weeks prior to earnings announcements. In a more recent article, Da *et al.* (2013) construct an index based on economy-related search terms and find a positive correlation with market volatility in the short-term, as well as return increases over the next few days. Preis *et al.* (2013a) show that stock trading strategies based on Google SV changes achieve greater profits than a random or buy-and-hold strategy. Finally, Kristoufek (2013) posits that more frequently searched stocks are riskier, and demonstrates how this specific feature can contribute to lowering the overall risk of a portfolio by assigning lower portfolio weights to risky stocks. In addition to financial markets, Dietzel *et al.* (2014), Hohenstatt *et al.* (2011), Hohenstatt and Kaesbauer (2014), Beracha and Wintoki (2013) and Wu and Brynjolfsson (2014) demonstrate Google's predictive abilities for the real estate markets at both national and state levels for commercial and residential real estate markets.

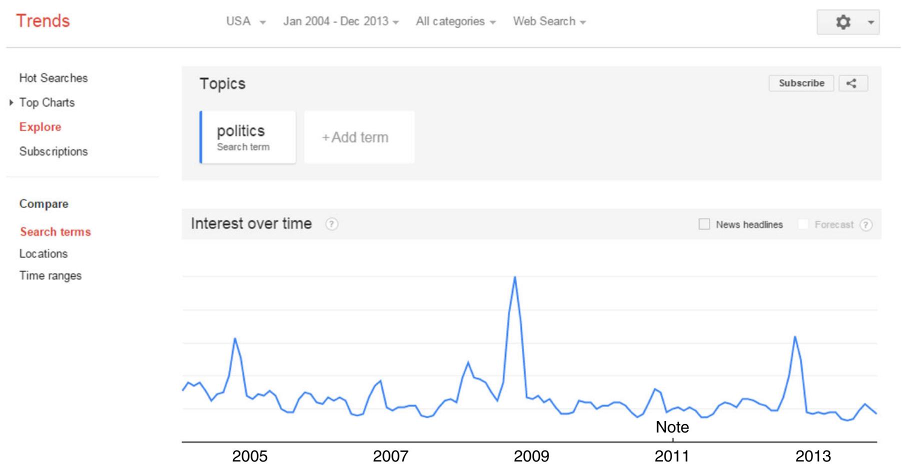
The existing literature has established various relationships between Google search data and financial, as well as property markets. As outlined above, the present article makes a contribution to this field of research by focussing on the relationship between online search data and the REIT market. By its very nature, the REIT market is an ongoing subject of debate with regard to the question of whether it is more similar to the stock or direct real estate market (e.g. Geltner and Kluger, 1998; Giliberto, 1990; Ross and Zisler, 1991; Myer and Webb, 1994; Brounen and Eichholtz, 2003; MacKinnon and Al Zaman, 2009; Sebastian and Schaetz, 2009; Oikarinen *et al.*, 2011; Hoesli and Oikarinen, 2012). We approach this issue by exploring the nature of information demand by real estate (equity) investors. First, it is crucial to establish to what extent the REIT market differs from the stock market in terms of its information transparency and efficiency. Wang *et al.* (1995) conclude that REIT markets behave differently from general stock markets, because they do not provide the same level of information dissemination and monitoring activities. This is mainly observed in terms of lower stock turnover ratios, a lower proportion of professional shareholders and relatively low financial analyst coverage. This supports the notion that REIT investors require a greater amount of information for pricing the stock. Tsai and Chiang (2013) analyze the relationship between six Asian/Pacific REIT and general stock markets and provide evidence that previous information about stocks lead changes in REIT markets. Furthermore, they find that price adjustment efficiency in the examined stock markets is greater than in the REIT markets, as disequilibrium occurs.

3. Data

3.1 Internet search query data

Typical information sources an investor might use to satisfy his information demand are financial service providers (Bloomberg, Google Finance, Yahoo Finance, etc.), newspaper web sites or news searches in general (e.g. *Financial Times*, *New York Times*), market reports (e.g. Jones Lang Lasalle office market report) and so on. Search engines enter this picture as either a mediator between the user and the final information source, or as the primary information source itself. Internet users enjoy the convenience of simply "googling" a web site, rather than having to type the full web address into their browser. This trend is intensifying, as most browsers have a separate search tool bar that is often linked to one of the large search engine providers. It is, for example, considerably easier to google "JLL Market Report Office New York 2014" than to visit the Jones Lang LaSalle web site, enter the research section and find the required market report manually. For these reasons, search engine providers have unique potential to capture a very broad share of interest.

In its role as the market-leading search engine with a share of 67.6 percent (Comscore, 2014) Google data can be seen as largely representative of the market as a whole. In 2008, Google extended its services with the publicly available web tool GT, which provides data about the popularity of a specific search query over time[1]. After the merging of "Google Insights for Search" and GT in 2012, the web facility appeared in a new interface offering various new features (see Figure 1). The requested data are presented as a graph in the form of a SV index, comprising data starting in 2004. The volume of search queries for a specific term is not given in absolute numbers, but in normalized and scaled values between 0 and 100. The value 100 represents the peak of a search query volume over the observed time span. Due to this normalization procedure, the SVI for the same term can change as soon as the volume reaches a new high. To extract the data for a specified search term, GT analyzes a sample of all search data, excluding data like repeated queries by the same user over a short period of time and terms of only limited interest[2]. Besides the depiction of the SVI as a graph, GT provides the user with a free download file containing a time series with weekly volumes of search queries. The weekly data covers search queries conducted from Sunday to Saturday. GT makes the newest weekly data available with an approximate two-day delay. Hence, an SVI, downloaded on Monday, already contains data for the previous week.



Notes: This screenshot shows the dashboard of Google Trends, available under www.google.de/trends/. By way of example, we present the web search interest for the term "politics" in the USA for the period January 2004 - December 2013

Figure 1.
Google trends' dashboard for the search term "politics"

While GT is set up by default for search queries conducted on a global level ranging from 2004 to the present, the tool offers a variety of ways to filter the SVI. The US regional filter can be applied on a national, state or MSA level to filter out untargeted information. The time frame filter can either be set to a fixed period on a monthly basis (e.g. January 2004-June 2011) or from 2004 to the present. Furthermore, GT provides a filter for categories and subcategories covering all kinds of subjects. This offers two exceptional advantages. First, terms affiliated with specific fields of interest can be allocated definitively to the appropriate area. In order to place searches in the right category, Google also analyzes preceding and following queries to gain a better understanding of the user's objectives. Second, GT not only provides SV within a (sub-) category, but also for entire categories or subcategories. For example, using the real estate category will yield a mix of search terms representing the demand for real-estate-related information, for example "for rent," "apartments" or simply "real estate." GT has recently launched an updated beta version of a filter option extension. Besides categories, a recently added feature enables the measurement of research interest in "topics," so as to capture overall search interest. As a consequence, GT's algorithms categorize various different search queries that are related to the same topic. Topics may refer to a company name, quotation or literature subject and so forth. One of the benefits of this new feature is the option to capture all different spellings of a search term (e.g. Munich, Muenchen, München). In practice, this means that when, for example, typing the term "capital market" in the search box, GT suggests either the SVI for the specific search term or the literature subject which includes all content-related terms. Moreover, different company names comprising the words "capital" and "market" are suggested.

For this study, we make use of the different filters to optimally capture the desired request for information. As this study covers the US REIT market, the regional filter for all terms and (sub-) categories is set to queries submitted in the United States only[3]. We focus mainly on individual search terms without any category filter. For search terms that may lead to misunderstandings, we obtain queries from within categories (_cat). In addition, we incorporate entire categories (_Cat), subcategories (_Subcat) and a number of different topics. For an overview of all searches (see Appendices 1 and 2).

3.2 Sampling noise

In one of the first studies to analyze GT data, Choi and Varian (2009) address the inconsistency in SVI data. They explain these issues in terms of GT's data sampling and extraction method. This inconsistency is manifest in slight changes in the SVI, when downloaded for the same time range, but on different occasions. Carrière-Swallow and Labbé (2013) analyze the sampling noise in greater detail, in order to identify the exact historical SVI. However, the sampling error in the data used corresponds with the deviation observed by Da *et al.* (2011a, 2013). When downloaded on different independent days, the correlation between the SVIs for the total time period usually lies above 97 percent (predominately above 98 percent). Moreover, we find a larger discrepancy between search terms with a lower number of total search requests and a stronger correlation in the more recent part of the SVI. However, in spite of this minimal deviation, different methods have recently been employed to minimize the noise. Baker and Fradkin (2011) use the average of four SVIs resampled during four different weeks. Preis *et al.* (2013a) average the SVI over three independently requested SVIs during consecutive weeks. To ensure the reliability of the results and to find clear underlying signals, we increase the number of SVIs. Thus, we use the arithmetic mean of five unique SVIs to decrease the prevailing variation among the individual SVIs. Consequently, we compute one solid and therefore more reliable SVI for each investment strategy:

$$SVI_{mean} = \frac{1}{5}\sum_{t=1}^{5} SVI_t \tag{1}$$

Furthermore, we exclude search terms of limited interest, for which the greater part of their weekly values is measured as 0. Another difficulty in doing research with GT data arises from terms whose relevance has increased tremendously over the time span. As a direct consequence of this substantial change in volume and the normalization procedure, weekly data, especially in the starting years of GT data (2004, 2005) are valued as 0. This could either be due to the fact that a term becomes more crucial after an unanticipated event, or due to the rapid increase in the number of Google users over recent years. Since we aim to ensure explicit signals, but also avoid excluding relevant terms which are affected by this issue, the observed time frame of the study extends from January 2006 to December 2013.

3.3 Search terms

In contrast to the existing literature (e.g. Da *et al.*, 2011a; Vlastakis and Markellos, 2012; Bank *et al.*, 2011), who specifically choose company-related search terms (e.g. company names, stock tickers) for their analysis, we focus on search terms that reflect the interest of market participants more broadly and on a market level. Thus, to measure the information demand of Google users regarding real estate in general and REITs in particular, we form a subset with real-estate-related search terms. Following the literature (Chauvet *et al.*, 2013; Baker and Fradkin, 2011; Preis *et al.*, 2013a) and in order to avoid an arbitrary selection process, three methods are applied to identify appropriate search terms for our real estate subset. In doing so, we start with logical keywords and synonyms for the terms “real estate” and “reit.” Subsequently, we use Google Sets, a Google Labs project which generates a list of similar terms, based on two or three starting values/terms (e.g. “real estate,” “properties,” “reit”). Further search terms are chosen by adding the top “related [search] terms” suggested by GT. In addition, we make use of entire categories and subcategories provided by GT, analogous to Choi and Varian (2012) or Hohenstatt *et al.* (2011). Besides single search terms, we also examine combined terms to include plurals and related terms with similar underlying search interest.

To examine the particular ability of “real-estate”-related terms to predict changes in the US REIT market, in comparison to other terms, we create three subsets. The first comprises real-estate-related search terms, (sub-)categories and topics. While the second subset covers finance-related search requests, the third subset serves as a control subset and contains the 50 most popular names in the USA[4].

3.4 Capital market data

The capital market data for the study derive from Thomson Reuters Datastream. To capture a wide range of US REITs, we extract the MSCI US REIT Index which represents approximately 85 percent of the US REIT universe, with exposure to all investment and property sectors[5]. Eligible for inclusion in the MSCI US REIT Index are “Diversified REITs,” “Industrial REITs,” “Mortgage REITs,” “Office REITs,” “Residential REITs,” “Retail REITs,” and “Specialized REITs,” all generating a majority of their revenue and income from real estate rental and leasing operations. The MSCI US REIT Index is a free float-adjusted market capitalization index. For the investment strategy, we obtain the weekly index prices starting on Monday, January 2, 2006 and ending on Monday, December 30, 2013. Similarly, we extract the DJIA to compare the results to a broad market index.

Since investing by following weekly GTIS signals requires frequent trading and is shaped by numerous changes in long and short positions, transaction costs are considerable. Hence, to make the results more reliable, we account for these costs when calculating the overall performance of the strategies. In order to determine the transaction costs for buying and selling the MSCI US REIT Index, we take a look at the Vanguard US REIT exchange-traded fund[6] (ETF) by way of example. When investing in an ETF, the main drivers of transaction costs are commissions, account service fees, expense ratios, and daily bid-ask spreads. As commissions and account service fees vary considerably among investors, and converge to practically zero with an increasing level of professionalism and portfolio size, we ignore this part of transaction costs. The expense ratio for ETFs is typically low and accounts for 0.10 percent p.a. (Vanguard REIT ETF), while the average bid-ask spread in the observed time span amounts to 0.18 percent. To examine the transaction costs of the DJIA, we use the Standard & Poor’s Depositary Receipts (SPDR) DJIA ETF with an expense ratio of 0.17 percent p.a. and an average bid-ask spread of 0.05 percent. However, in order to assess the results properly, we do not employ the average of bid-ask spreads, but rather the respective observed daily spreads. All results presented in this paper account for transaction costs and therefore reflect net performance.

4. Methodology

In order to investigate the existence of a relationship between information demand and the US REIT market, we apply a methodology similar to Preis *et al.* (2013a) and formulate hypothetical REIT investment strategies with GT data by quantifying relative volume changes in the SVI on a weekly basis. On this basis, we first compare the SV in week *t* with the average volume of the three previous weeks. Additionally, a crucial adjustment regarding the timing and influence of information demand and its impact on the market is made. Here, we follow Da *et al.* (2011a), who conclude that search queries conducted two weeks previously, have predictive ability for the capital market. Furthermore, in terms of timing, Drake *et al.* (2012) find that information demand through the internet starts increasing, on average, about two weeks prior to earnings announcements. These studies underline the delay in information demand with regard to an actual impact on the market. We assume that there is a specific time frame between the research process and the final transaction for considering or promoting an investment. While for professional investors, this time delay can be caused, for example, by board meetings, investment committees, pending management approvals or risk management assessments (see e.g. Emmanuel *et al.*, 2010), for private investors, it can be presumed that the investment decision is held back by factors such as the ongoing process of collecting information or consulting investment advisors. Hence, we allow for a lag of two weeks. This is associated with the practical advantage that the two-week time frame provides sufficient latitude for downloading 5 differentiating SVIs so as to acquire the SVI_{mean} in order to deal with sampling noise. The fact that GT updates its data within 48 hours at the longest, ensures the applicability of GTIS. The trading signals are derived as follows:

$$\Delta SV_{(week_t)} = SV_{(week_{t-k})} - \frac{\sum_{i=1}^{T} SV_{(week_{t-k-i})}}{T} \tag{2}$$

where T is the number of periods, the $SV_{(week_{t+1})}$ is compared to ($T = 3$)[7]. k is the number of lags needed to incorporate the theory that an increased/decreased information demand has a lagged impact on market movements ($k = 2$).

Up to this point, it is unclear whether an increase in SV is also directly related to a subsequent increase in the price of the US REIT Index, or whether there is a reverse relationship. In the literature (e.g. Bank *et al.*, 2011; Da *et al.*, 2011a; Preis *et al.*, 2013a), this matter of whether there is a positive or negative relationship between SV and market movement has been the subject of controversy. Da *et al.* (2011a) refer to the attention theory of Barber and Odean (2008), which argues that investors are net buyers of "attention grabbing" stocks. Thus, in the short-term, an abnormal interest in certain stocks results in temporary positive price pressure. Preis *et al.* (2013a), on the other hand, base their theoretical framework on Herbert Simon's (1955) model of decision making and argue that people tend to gather more information about the state of the market during times of concern and uncertainty. They find that this has been reflected historically in an increase in Google SV for keywords of financial relevance. This implies that an increase in abnormal search interest predicts temporary downward market-pressure. The findings of Merton (1987) and Fang and Peress (2009) indicate a similar relationship, as they argue that stocks with low media coverage and investor attention tend to perform better. Considering both arguments in this literature, we take an objective and somewhat more data-driven approach to finding the relationship between a search term and market behavior. This stems from the fact that it is difficult to state whether a query for a term like "property" was conducted because of buying interest (in REITs for example) or rather due to concern (that the market is about to drop). To do so, we use the observation period extending from 2006 to the

Lehman Brothers collapse (15th September 2008) to determine the correlations between each individual SVI_{mean} and the MSCI US REIT Index[8]. The total sample of all SVI_{mean} is then divided into positively and negatively correlated sample groups. According to the respective correlation, we formulate two opposing strategies by looking at the relative changes in SVI_{mean} $\left(\Delta SV_{(week_t)}\right)$.
Positively correlated:

$$Trading\ Signal_{(week_t)} \begin{cases} 0, if\, \Delta SV_{(week_t)} > 0 \\ 1, if\, \Delta SV_{(week_t)} < 0 \end{cases} \tag{3}$$

Here, we define 0 as a buy signal and 1 as a sell signal[9]. If an SVI_{mean} is positively correlated to the MSCI US REIT Index, a positive value of (2) indicates an upward trend of the index (buy signal). Therefore, we take a long position and invest in the REIT Index. If (2), for a positively correlated SVI_{mean} is lower than zero, we take a short position (sell signal) in the MSCI US REIT Index[10]:

Negatively correlated:

$$Trading\ Signal_{(week_t)} \begin{cases} 1, if\, \Delta SV_{(week_t)} > 0 \\ 0, if\, \Delta SV_{(week_t)} < 0 \end{cases} \tag{4}$$

Here, we define 0 as a buy signal and 1 as a sell signal[11]. For a negatively correlated SVI_{mean}, we apply the same strategy as above, only vice versa. Hence, if $\Delta SV_{(week_t)}$ exceeds zero (sell signal), a short position is taken. The opposite applies for a value below zero (buy signal).

If the signal changes from buy to sell, we sell the index (ETF) and in an immediate second step, short-sell the index (ETF)[12]. In a case of vice versa, the short position is cleared and the index is bought. This means that a changing signal requires two transactions and thus causes twice the transaction costs[13]. For an unchanged signal, the position is held. The first trade according to GTIS is executed on Monday, February 13th, 2006.

To put the GTIS performance into perspective, the results are compared against three benchmarks. The main benchmark is the MSCI US REIT Index, since the strategies' profits and losses derive directly from the weekly index. We stipulate that the buy-and-hold strategy is based solely on two contrary transactions in terms of buying into and selling the index at the beginning and the end of the period under consideration, respectively. Similarly, we introduce the buy-and-hold strategy for the DJIA to compare the results to a broad market index. The third benchmark ("random strategy") is defined as the average of 10,00,000 variations of an investment strategy that follows purely random buy and sell signals for the MSCI US REIT Index in each week of the period. In addition to the three benchmarks and as another way of showing that the proposed investment strategies are not due to chance, they are compared to investment strategies based on changes in the SV of the 50 most popular given names in the USA.

5. Analysis and findings

5.1 Performance and risk measures

In order to examine the symmetric impacts of buy and sell actions we determine the investment performance in log returns for all GTIS and compare those results to the buy-and-hold strategies described in Section 4. As a second measure for quantifying the prediction abilities of GTIS, we calculate hit rates, which are defined by the number of a strategy's correct predictions about whether the index is going to rise (fall) divided by its total number of predictions. Since the observed time period is characterized by extreme market movements, the overall hit rate is extended by a second measure which particularly takes extreme market movements into account. Accordingly, we analyze the correct predictions that were made by the respective GTIS about the MSCI US REIT Index' 20 best, as well as the 20 worst performing weeks.

As a risk indicator, we analyze the probability of an investor losing money if he had applied a proposed investment strategy for a period of six months. We roll this time period through the entire observation period from 2006 to 2013 and take the average over all six-month periods. In addition, we measure the risk exposure of the GTIS by applying Jensen's α which provides insight into the strategies' abnormal returns in excess of a theoretically expected return (Jensen, 1967) and constitutes a risk-adjusted performance measure. Furthermore, the β's demonstrate how the investment strategies' returns move, in comparison to the benchmark returns. In light of the capital asset pricing model, we calculate Jensen's α as:

$$\alpha_J = R_{GTIS} - \left[R_f + \beta_{GTIS,M} \cdot \left(R_M - R_f\right)\right] \tag{5}$$

where R_{GTIS} is the returns from trading strategy, R_f the risk-free rate, determined from daily three-Month US- Treasury Bill yields (2006-2013), R_M the average of weekly DJIA returns, $\beta_{GTIS,M}$ the beta of investment strategy returns against weekly DJIA returns.

5.2 Empirical results

Table I provides an overview of the performance and risk measure results for the top 20 GTIS, as well as the three benchmarks, sorted according to their average annual log return[14]. The first and most crucial result is that, measured by the annual log returns from 2006-2013, 40 GTIS would have outperformed a buy-and-hold strategy for the MSCI US REIT Index with an absolute performance of 0.1 percent p.a., measured in log returns. The results are largely in line with the findings of Preis *et al.* (2013a) who state that changes in the search activity of Google users give an indication of short-term movements of financial markets.

Looking at the top performing search terms, one finding merits particular attention. The top six search terms are all strictly related to real estate. The picture becomes even clearer when looking at the top 20 GTIS, where only six searches from the finance subset contrasts with 14 real-estate-specific searches. This confirms the hypothesis that Google is able to capture the interest of potential real estate equity investors. However, the question arises as to what specifically causes real-estate-related searches to predict more accurately than, say, finance-related terms, even though the largest part of REITs is traded publicly? Graff (2001, p. 104) notes that a principle of basic economics states that "[…] the investment value of an asset equals the present value of future net cash flows expected from the asset." Because of REITs' unique legal requirements, they have only one kind of income-producing asset, namely properties. Accordingly, one can surely expect REIT investors to be aware of this feature. This suggests that investors gather information about potential future cash flows from their company shares by reviewing the property market as a whole. Apart from explicit management-related mistakes, a REIT's performance is inevitably highly correlated with the property

Table I. Performance results for the 20 top ranked information-demand-based investment strategies and the benchmark strategy

			Overall (2006-2013)						
Rank	Category	Search terms	Log returns (%) (p.a.)	t-stat	Rolling 6-month risk of loss (%)	β	Jensen's α (p.a.)	Hit rate	Hit rate (20 high/ 20 low) (p.a.) (%)
1	Real estate	properties + property	15.5	2.74	29.9	−0.13	14.0	57.6	75.0
2	Real estate	Real Estate Agencies_Cat	14.3	2.52	31.2	−0.17	12.9	53.2	70.0
3	Real estate	Realty trust + realty trusts	13.4	2.37	38.1	−0.06	11.9	54.2	67.
4	Real estate	home + homes	13.1	2.31	40.7	−0.27	11.7	54.2	72.5
5	Real estate	condos + condo + condom inium + condom iniums	12.6	2.23	35.0	−0.08	11.1	52.5	72.5
6	Real estate	Real Estate_Cat	11.5	2.03	43.5	−0.19	10.1	53.2	72.5
7	Finance	earnings + earning	10.8	1.90	31.5	−0.02	9.3	54.0	62.5
8	Real estate	real estate agent + real estate agents + real estate agencies	10.1	1.77	37.1	−0.28	8.8	54.4	62.5
9	Real estate	Real Estate Listings_Subcat	10.0	1.76	41.2	−0.18	8.6	52.8	75.0
10	Finance	politics	8.5	1.50	37.6	−0.34	7.3	50.4	62.5
11	Real estate	real estate	8.3	1.46	44.5	−0.44	7.1	52.8	70.0
12	Real estate	real estate company + real estate Companies	8.2	1.44	34.3	−0.19	6.8	56.6	60.0
13	Real estate	property tax	8.0	1.40	32.5	0.31	6.2	56.4	62.5
14	Real estate	Property Inspections & Appraisals_Subcat	7.6	1.34	33.5	0.04	6.1	54.7	65.0
15	Finance	profit+profits	7.2	1.26	36.1	−0.49	6.1	53.2	57.5
16	Real estate	commercial properties + commercial property	6.3	1.10	47.6	−0.20	4.9	52.0	72.5
17	Finance	News_Cat	6.1	1.07	31.2	−0.04	4.6	51.1	62.5
18	Real estate	report	5.1	0.89	35.8	−0.13	3.7	52.8	60.0
19	Finance	economy	5.1	0.89	41.2	−0.22	3.7	48.7	62.5
20	Finance	s&p 500	4.5	0.78	35.6	−0.08	3.0	54.0	60.0
–	Benchmark	MSCI US REIT Index (buy-and-hold strategy)	0.1	–	38.9	–	–	–	–
–	Benchmark	Dow Jones Industrial Average (buy-and-hold strategy)	2.1	–	30.7	–	–	–	–
–	Benchmark	Random strategy	0.0	–	0.0	–	–	50.0	50.0

Notes: This table depicts the main performance measures, log returns (p.a.), rolling 6-month risk of loss, the hit rates, as well as the hit rates for the 40 highest/lowest price movements (representing the most volatile market phases) of the 20 top performing information-demand-based investment strategies, ranked by log returns (p.a.). The rolling six-month risk of loss resembles the average probability an investor incurs, if he trades according to the respective GTIS signals for 6 months. At the bottom, we present the buy-and-hold strategy for the MSCI US REIT Index, as well as the Dow Jones Industrial Average and a random strategy (1,000,000 trials following purely random buy and sell signals for the MSCI US REIT Index in each week of the period under consideration). The results are significantly different from 0 for significance levels of 90 percent if t-stat > 1.28, 95 percent if t-stat > 1.65, 99 percent if t-stat > 2.34

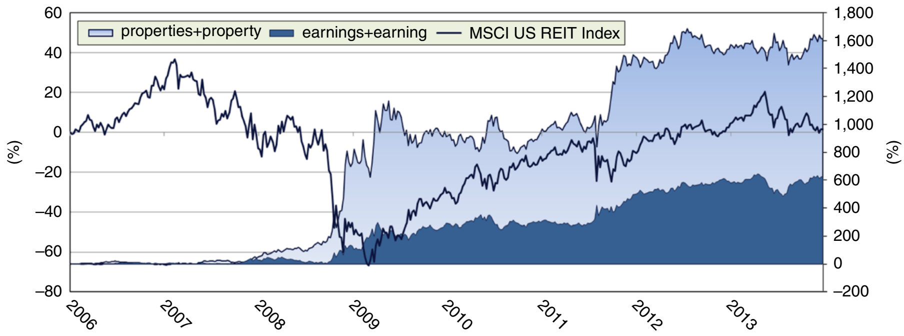
markets[15]. We believe that the GTIS that are based on real-estate-related searches are more able to capture this very process of information procurement by investors and thus perform better over the observation period as a whole.

The best performance is observed for the GTIS "properties + property" which would have achieved a remarkable log return performance of 15.5 percent p.a. (after accounting for transaction costs). This means that an investor would have made incomparably more than what could have been earned with a simple buy-and-hold strategy (0.1 percent p.a.). Abnormal returns, reflected by Jensen's α, are just as impressive, with a return of 14.0 percent p.a. for the GTIS "properties + property." The β's among the top 20 performing GTIS range between 0.31 and −0.49, showing that a predominant part of GTIS' movements are almost uncorrelated to the market portfolio. In other words, returns of the GTIS either tend to move marginally in the opposite direction or in the same direction of the market portfolio, proxied by the DJIA. The risk of losing money that an investor has to bear over a six-month holding period, yields interesting insights into the behavior of the portfolio returns over time. While the buy-and-hold strategy for the MSCI US REIT Index bears a risk of 38.9 percent of losing money, 16 GTIS have a lower exposure to risk, despite higher returns. The GTIS "properties + property" ranks first overall, with a loss risk of only 29.9 percent. This finding strongly suggests that a large number of GTIS are able to reduce the probability of losing, while gaining substantially higher returns.

Generally, one would assume that the best performing strategy is also the one with the highest hit rate, i.e. the best predictor of the direction of market movement. However, the results suggest a slightly different finding. The GTIS with the highest hit rates are not automatically the best performers. The second best predicting GTIS "real estate company + real estate companies" with a hit rate of 56.6 percent, for example, only ranks 12th in terms of absolute investment performance. Accordingly, we take a look at the hit rate for the 20 highs/lows. As presented in Table I, the best performing strategies have a considerably higher hit rate (up to 75 percent) during volatile weeks than on average. This suggests that the correct prediction of big jumps, i.e. very volatile market phases, is much more important than the absolute prediction accuracy. These results point toward the same direction as the notion that investors have a higher demand for information in times of either increased uncertainty or market attention. Hence, strategies that are more effective in capturing this uncertainty or interest-induced information demand, tend to be more profitable. Figure 2 depicts the performance of the GTIS "properties + property," as well as the finance-related term "earnings + earning," in comparison to a buy-and-hold strategy for the MSCI US REIT Index. The search term "earnings + earning" constitutes the best finance-related term and supports the findings of Gyamfi-Yeboah *et al.* (2012) who state that the earnings announcements have a significant impact on REIT returns.

5.3 Sub-periods

The results so far suggest that GTIS are able to outperform the benchmarks and yield consistent returns. However, since the period of observation (2006-2013) was shaped by instability, we take a closer look at certain sub-periods, in order to gain a better understanding of the time-specific dynamics of Google investment strategies, as can be seen in Table II. Hence, we define a "crisis period," which starts with the week of the Lehman Brothers collapse (September 15, 2008) and ends the week the MSCI US REIT Index, regains its pre-Lehman-collapse level for the first time (21st of February 2011). In



Notes: This exhibit graphs the absolute investment performance (reinvestment assumption) over time (2006-2013) for the GTIS "properties+property" (1,613.4 percent overall; 42.8 percent p.a.; reinvestment assumption), "earnings+earning" (625.2 percent overall; 28.2 percent p.a.; reinvestment assumption) and the MSCI US REIT Price Index (1.6 percent overall; 0.2 percent p.a.). Note that the MSCI is presented on a second scale (left scale), as otherwise, the curve would appear as a fairly flat line, due to the very large discrepancy in absolute investment performance

Figure 2. The absolute investment performance of the best performing GTIS of each subset ("properties + property" and "earnings + earning") in comparison to the MSCI US REIT Index

addition, we examine the years 2012-2013 ("recent years period"). While, during the crisis period, the MSCI US REIT Index has a standard deviation of 7.8 percent and can be characterized as quite turbulent, the second sub-period reflects economically more stable times with a standard deviation of only 2.2 percent. Table II shows the log returns (p.a.) for the entire period of observation, crisis and recent years.

The results paint a clear picture. Many GTIS performed very strongly during the crisis. On the one hand, this is due to extreme market movements during that period [16]. On the other hand, it again suggests that especially in times of uncertainty, when investors have an increased appetite for information, GTIS work exceptionally well. Analogous to the statements above, we believe this is because investors investigate information more intensely and thereby display more and clearer signs of their investment behavior. This is in line with Vlastakis and Markellos (2012), who find that the effect of Google information demand becomes stronger during high-return market phases. The picture changes to a certain degree when considering the results of the recent years period 2012-2013. Even though the observation period is less than half a year shorter than the crisis period, the annual log returns are significantly lower. Similar to the crisis period, this is probably partly due to lower market volatility per se. Nevertheless, the MSCI US REIT Index buy-and-hold strategy (2.6 percent p.a.) would have been outperformed by 25 GTIS during this period of observation.

5.4 Testing for randomness

As described in the methodology section, in order to prove that the results are not only due to chance, we implement a random strategy. After 1,000,000 trials, as expected, the analysis suggests an average performance of 0.0 percent p.a., measured by log returns. In addition, we introduce a random sample with the 50 most popular first names in the USA (US Social Security Agency, 2014). Based on these searches, we generate short and long signals in the same manner as for the rest of the strategies. The results reveal an average performance of −2.5 percent p.a. (t-stat − 4.88; p-value 0.000), measured by log returns[17].

Table II.
Top 20 google trends investment strategies (GTIS) ranked by their respective log returns (p.a.) for the overall time period (2006-2013) and sub-periods (crisis, recent years)

Rank	Category	Search terms	Overall (2006-2013) Log returns (p.a.) (%)	Crisis (09/13/2008-02/19/2011) Log returns (p.a.) (%)	Recent years (2012-2013) Log returns (p.a.) (%)
1	Real estate	properties + property	15.5	25.8	1.7
2	Real estate	Real Estate Agencies_Subcat	14.3	25.9	8.8
3	Real estate	realty trust + realty trusts	13.4	19.4	−3.9
4	Real estate	home + homes	13.1	34.8	2.4
5	Real estate	condos + condo + condominium + condominiums	12.6	38.6	2.0
6	Real estate	Real Estate_Cat	11.5	36.2	5.5
7	Finance	earnings + earning	10.8	23.9	4.2
8	Real estate	real estate agent + real estate agents + real estate agencies	10.1	16.0	4.7
9	Real estate	Real Estate Listings_Subcat	10.0	28.1	−3.4
10	Finance	politics	8.5	21.8	8.1
11	Real estate	real estate	8.3	26.9	−1.0
12	Real estate	real estate company + real estate companies	8.2	1.7	2.7
13	Real estate	property tax	8.0	15.2	4.9
14	Real estate	Property Inspections & Appraisals_Subcat	7.6	22.4	2.4
15	Finance	profit+profits	7.2	27.3	2.9
16	Real estate	commercial properties + commercial property	6.3	22.0	5.1
17	Finance	News_Cat	6.1	19.6	1.2
18	Real estate	report	5.1	6.3	−2.2
19	Finance	economy	5.1	23.1	6.3
20	Finance	s&p 500	4.5	−1.0	5.6
–	Benchmark	MSCI US REIT Index (buy-and-hold strategy)	0.1	0.0	2.6
–	Benchmark	Dow Jones Industrial Average (buy-and-hold strategy)	2.1	2.2	6.4
–	Benchmark	Random Strategy	0.0	0.0	0.0

Notes: This table depicts the log returns (p.a.) for the entire observation period (2006-2013), the crisis period (09/13/2008-02/19/2011) and “recent years” period (2012-2013) for the 20 best performing strategies and the four benchmark strategies (random strategy, buy-and-hold strategy for the MSCI US REIT Index and the Dow Jones Industrial Average, momentum strategy). The standard deviations of the MSCI US REIT Index over the presented sub-periods are 5.2, 7.8 and 2.2 percent, respectively. This shows the varying volatility of the market during the different sub-periods

5.5 Search relevance and impact factor

In order to find out whether search relevance is positively correlated with greater investment performance, we identify the average relative SV of all real estate searches. Search relevance is defined as how often a certain term is searched for in comparison to others. Terms that are searched for more often are therefore more relevant than others. Since GT does not provide the absolute SV, we compute a relative impact factor by a pair-wise comparison of each individual SVI with a benchmark SVI (the "real estate category" supplied by GT). The search with the highest impact factor is set to 100. The ranking of all 40 searches is presented in Appendix 3. We find a positive correlation between search relevance and investment performance, by calculating the rank correlation coefficients of Kendall and Spearman (Spearman's $\rho = 0.30$, t-stat $= 1.90$; Kendall's $\tau = 0.24$, z-stat $= 2.22$). These results can be interpreted cautiously as meaning that for the most part, more frequently searched terms tend to perform better as investment strategies. In other words, investment performance is positively related to search relevance.

5.6 Comparison to broad market index

In order to investigate whether the GTIS perform differently for the MSCI US REIT index in comparison to a general equity index, we run the sample of real estate and finance-related search terms on the DJIA. By comparing the respective results, we find that real-estate-specific searches contain on average more valuable information about the behavior of the REIT market than finance-related searches. At the same time, finance-related search terms perform better for the DJIA. As presented in Table III, the results indicate that while asset-specific information play a significant role for predicting a one-asset-specific market, more general finance-related search queries are superior in predicting changes in the broader stock market.

5.7 Out-of-sample testing

So far in this current research, investment strategies have mainly been used as a methodology for examining the relationship between (online) information demand and the US REIT market in general. Nevertheless, the investment strategies themselves could also be of great interest for the investment industry. The analysis in the present paper is based predominantly on the greatest possible time period (overall period). But since the relationship between the SVI and the underlying index is determined between 2006 and the Lehman Brothers collapse (September 15, 2008), this period can be seen as a model training period according to trading strategy analysis. Consequently, the correlation period is now defined as an in-sample period, while the remaining time

Table III. Average performance results for the GTIS from the real estate and finance subsets based on the MSCI US REIT Index and the Dow Jones Industrial Average

	MSCI US REIT Index				Dow Jones Industrial Average			
	Average log returns (p.a.) (%)	t-stat	p-value	Average hit rate (%)	Average log returns (p.a.) (%)	t-stat	p-value	Average hit rate (%)
Category								
Real estate	2.4	2.22	0.016**	51.2	−2.1	−5.06	0.000***	49.0
Finance	−2.2	−2.78	0.004***	49.5	0.1	0.33	0.371	50.4

Notes: This table depicts the average log returns (p.a.) and the average hit rate for the GTIS from the real estate and finance subsets, respectively, applied to the MSCI US REIT index and the Dow Jones Industrial Average. Significant at $*p < 0.10$; $**p < 0.05$; $***p < 0.01$

period is considered as an out-of-sample period. To identify which investment strategies within the proposed set an investor would have chosen at this specific point in time, the GTIS are sorted according to their performance during the in-sample period. Table IV presents the five best performing (in-sample) investment strategies and their respective out-of-sample performance. Interestingly, apart from the term "fund + funds," all real-estate-related terms remain positive and outperform the benchmark (0.9 percent p.a.) for the rest of the observation period. The performance of the GTIS "properties + property" corresponds to the overall results and continues to yield very stable out-of-sample performance of 15.4 percent p.a. compared to 15.7 percent p.a. (in-sample).

In summary, it can be stated that there is indeed a relationship between information demand and the US REIT market. The findings moreover suggest that the methodology yields remarkable potential for application in information-demand-based investment strategies.

6. Conclusion

Social science research has become more and more popular over recent years, not least because of the emerging availability of "big data." The term big data is generally defined as any information collected and made available by internet service providers. Among this broader data set, SV data made available by the tool GT, is enjoying growing popularity among researchers. Existing research has found various promising ways to use the data for real estate and financial markets.

In this paper, investment strategies based on weekly Google SV are used to explore the relationship between online information demand and the US REIT market. To do so, buy (long) and sell (short) signals for the MSCI US REIT Index are being identified from changes in internet SV. Following this approach, the performance of individual information-based investment strategies is compared to a buy-and-hold strategy for the MSCI US REIT Index, the DJIA and a random strategy. The main conclusions are as follows. First, many GTIS substantially outperform the tested benchmarks. Hence, we draw the conclusion that there is a significant relationship between search interest and the performance of the US REIT market. Second, we find that real-estate-specific search terms are for the most part better predictors of the US REIT market than

Table IV. Performance results for the top five information-demand-based investment strategies within the in-sample and out-of-sample time period

Rank	Category	Search terms	In-Sample (01/06/2006-09/13/2008) Log returns (p.a.) (%)	Out-Of-Sample (09/13/2008-12/28/2013) Log returns (p.a.) (%)
1	Real estate	real estate company + real estate companies	21.0	1.6
2	Real estate	realty trust + realty trusts	18.5	10.8
3	Real estate	properties + property	15.7	15.4
4	Finance	fund + funds	10.3	−4.0
5	Real estate	real estate agent + real estate agents + real estate agencies	10.0	10.1
–	Benchmark	MSCI US REIT Index (buy and hold strategy)	−1.6	0.9

Notes: This table depicts the main performance measure log returns (p.a.). In addition, we list the benchmark buy-and-hold strategy for the MSCI US REIT Index. The results are sorted by their respective in-sample performance finance-related terms. On the other hand, finance-related terms tend to perform better with broader (less asset-specific) markets, such as the DJIA. We believe the main reason is that REIT investors mostly gather information about the underlying assets of their stocks, and that sentiment regarding real estate as a whole can be captured by Google. Property-specific GTIS exploit this information and are therefore better predictors of market movements than non-property-specific GTIS. Third, the best performing strategy "properties + property" would have gained a remarkable 15.5 percent p.a., measured by log returns, (compounding to 1,613.4 percent overall, reinvestment assumption) over an eight year investment horizon (2006-2013), after accounting for transaction costs, compared to 0.1 percent p.a. market index performance, measured by log returns. Fourth, GTIS perform better during volatile market phases, which is probably because of a greater demand for information induced by either investor uncertainty or increased interest in the market, so that people indicate their future behavior more clearly. This is also supported by the fact that the best market-movement-predicting strategies are not necessarily the best performing ones and that the most successful strategies have considerably better hit rates for predicting big jumps in the market.

It should be noted that if the advantages of search query data become known to a broader public and gain popularity among market participants, the benefits of this very data set could decrease to some extent (as information efficiency increases). This corresponds to Fama (1970), who states that an efficient market is one that fully reflects available information. Accordingly, investors are totally informed about the future cash flows of an asset. However, the given situation is slightly different. The Google data set does not provide more information about the assets' future cash flows, but about the behavior of investors who are trading this very asset. Thus, it does not per se improve market efficiency, but rather predicts the behavior of market participants and thereby more likely captures expectations (sentiment) than the fundamental element of market movements. In summary, it is difficult to foresee what is going to happen in the future, given the rapid dynamics of the internet and means of information supply. There is a chance that users will replace classic search engines over time with other web-services (e.g. social media, Q&A web sites) to gather their information. However, in our opinion, the World Wide Web, with its prevailing capacities, is not really challengeable and will therefore remain the most important source of information supply for a long time.

Global financial markets are presently quite vulnerable to new information entering the market, although they may have no direct impact on the firms underlying cash flows in the first place. As a logical consequence, it follows that between the availability of new information and a reaction on the stock market, there is an individual who has to assess the situation and make up his mind about whether to buy, hold or sell. In order to do so, he will seek information. We believe that this information-procurement process can be measured, quantified and exploited at least partly, in order to make short-term predictions about investor behavior. Combined with fundamentals of the REIT market or a single REIT company, the use of this information set could be utilized to further improve the understanding of real estate equity pricing and the predictive power of investment strategies for the listed real estate market. We believe that this field of research is still in its infancy as an approach for predicting the behavior of research-intensive financial market securities. Further research might also examine the application of the methodology used in this current paper toward other tradable asset classes that are available to the greater public.

Notes

1. See Google (About Trends Graph): https://support.google.com/trends/answer/4355164?hl=en&ref_topic=4365531
2. See Google (Where Trends data comes from): https://support.google.com/trends/answer/4355213?hl=en&ref_topic=4365599
3. We are aware that focusing on US internet searches only is a limitation, as some REIT stock holders are international. Using global search interest, however, would be too undirected, as global interest in real estate varies quite strongly, which would lead to distorted results.
4. Top names over the last 100 years: www.ssa.gov/oact/babynames/decades/century.html
5. www.msci.com/products/indices/country_and_regional/domestic_equity_indices/reit/msci_us_reit_index.html
6. The Vanguard REIT ETF is an ETF which closely tracks the return of the MSCI US REIT Index.
7. Following Preis *et al.* (2013a).
8. This specific time period was chosen to resemble an in-sample period for the investment strategy which will be discussed in more detail in the analysis section.
9. If the signal remains unchanged from one week to another or the following week, the current trading position is maintained.
10. We realize that the MSCI US REIT Index, by its very nature, is not tradable. However, there are investment vehicles, e.g. the Vanguard REIT ETF, for which such mechanisms (short and long position) can be applied.
11. If the signal remains unchanged from one week to another or the following week, the current trading position is maintained.
12. There are various ways to short an ETF. Depending on respective broker services and availability, the easiest way is to use an inverse ETF which is constructed to perform inversely to the movement of the underlying benchmark. A second option would be to short-sell the ETF. Particularly for broad market indices, there are various financial products (e.g. index certificates) offered by large banks that reflect the exact opposite of the underlying index's performance.
13. We are aware that trading costs for entering a long or short position can be different. However, for the sake of simplicity, we assume the same trading costs for both.
14. The ranking is based on the performance measured by log returns p.a. The total sample of results can be obtained from the authors upon request.
15. Fuess *et al.* (2014) point out that, due to their tax-exempt status, REITs are closely tied to the underlying (commercial) real estate market.
16. Anderson *et al.* (2012) state that the majority of high-variance price periods fall within the recent recession and real estate crises.
17. The results can be obtained upon request.

References

Anderson, R.I., Boney, V. and Guirguis, H. (2012), "The impact of switching regimes and monetary shocks: an empirical analysis of REITs", *Journal of Real Estate Research*, Vol. 34 No. 2, pp. 157-181.

Baker, S.R. and Fradkin, A. (2011), "What drives job search? Evidence from Google Search", SIEPR Discussion Paper No. 10-020, Stanford University, Stanford, CA, available at: www.siepr.stanford.edu/RePEc/sip/10-020.pdf

Bank, M., Larch, M. and Peter, G. (2011), "Google search volume and its influence on liquidity and returns of German stocks", *Financial Markets and Portfolio Management*, Vol. 25 No. 3, pp. 239-264.

Barber, B.M. and Odean, T. (2008), "All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors", *Review of Financial Studies*, Vol. 21 No. 2, pp. 785-818.

Beracha, E. and Wintoki, J. (2013), "Forecasting residential real estate price changes from online search activity", *Journal of Real Estate Research*, Vol. 35 No. 3, pp. 283-312.

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.

Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A. and Weber, I. (2012), "Web search queries can predict stock market volumes", *PLoS ONE*, Vol. 7 No. 7, pp. 1-17.

Brounen, D. and Eichholtz, P. (2003), "Property, common stock, and property shares", *Journal of Portfolio Management*, Vol. 29 No. 5, pp. 129-137.

Carrière-Swallow, Y. and Labbé, F. (2013), "Nowcasting with google trends in an emerging market", *Journal of Forecasting*, Vol. 32 No. 4, pp. 289-298.

Chauvet, M., Stuart, G. and Chandler, L. (2013), "Fear and loathing in the housing market: evidence from search query data", working paper, SSRN eLibrary.

Choi, H. and Varian, H. (2009), "Predicting the present with google trends", working paper, Google Inc., Cupertino.

Choi, H. and Varian, H. (2012), "Predicting the present with google trends", *The Economic Record*, Vol. 88 No. S1, pp. 2-9.

Comscore (2014), "US search engine rankings", April, available at: www.comscore.com/Insights/Market-Rankings/comScore-Releases-April-2014-US-Search-Engine-Rankings (accessed July 3, 2014).

Da, Z., Engelberg, J. and Gao, P. (2011a), "In search of attention", *The Journal of Finance*, Vol. 66 No. 5, pp. 1461-1499.

Da, Z., Engelberg, J. and Gao, P. (2011b), "In search of fundamentals", working paper, University of Notre Dame and University of North Carolina, Chapel Hill, NC.

Da, Z., Engelberg, J. and Gao, P. (2013), "The sum of all fears: investor sentiment and asset prices", working paper, SSRN eLibrary.

Dietzel, M., Braun, N. and Schaefers, W. (2014), "Sentiment-based commercial real estate forecasting with google search volume data", *Journal of Property Investment and Finance*, Vol. 32 No. 6, pp. 540-569.

Dimpfl, T. and Jank, S. (2012), "Can internet search queries help to predict stock market volatility?", *CFR Working Paper*, Vols 11-15, pp. 1-32.

Drake, M.S., Roulstone, D.T. and Thornock, J.R. (2012), "Investor information demand: evidence from google searches around earning announcements", *Journal of Accounting Research*, Vol. 50 No. 4, pp. 1001-1040.

Emmanuel, C., Harris, E. and Komakech, S. (2010), "Towards a better understanding of capital investment decisions", *Journal of Accounting & Organizational Change*, Vol. 6 No. 4, pp. 477-504.

Fama, E. (1970), "Efficient capital markets: a review of theory and empirical work", *The Journal of Finance*, Vol. 25 No. 2, pp. 383-417.

Fang, L. and Peress, J. (2009), "Media coverage and the cross-section of stock returns", *The Journal of Finance*, Vol. 64 No. 5, pp. 2023-2052.

Fuess, R., Mager, F. and Zaho, L. (2014), "The effect of macroeconomic news and monetary policy announcements on us reit and stock prices", *Real Estate Finance*, Vol. 30 No. 4, pp. 143-155.

Geltner, D. and Kluger, B. (1998), "REIT-based pure-play portfolios: the case of property types", *Real Estate Economics*, Vol. 26 No. 4, pp. 581-612.

Gilbert, E. and Karahalios, K. (2010), "Widespread worry and the stock market", *Proceedings of the International Conference on Weblogs and Social Media*, pp. 58-65.

Giliberto, S.M. (1990), "Equity real estate investment trusts and real estate returns", *Journal of Real Estate Research*, Vol. 5 No. 2, pp. 259-263.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L. and Zimmermann, K.F. (2009), "Detecting influenza epidemics using search engine query data", *Nature*, Vol. 457 No. 7232, pp. 1012-1014.

Graff, R.A. (2001), "Economic analysis suggests that REIT investment characteristics are not as advertised", *Journal of Real Estate Portfolio Management*, Vol. 7 No. 2, pp. 99-124.

Gyamfi-Yeboah, F., Ling, D.C. and Naranjo, A. (2012), "Information, uncertainty, and behavioral effects: evidence from abnormal returns around real estate investment trust earnings announcements", *Journal of International Money and Finance*, Vol. 31 No. 7, pp. 1930-1952.

Hoesli, M. and Oikarinen, E. (2012), "Are REITs real estate? Evidence from international sector level data", *Journal of International Money and Finance*, Vol. 31 No. 7, pp. 1823-1850.

Hohenstatt, R. and Kaesbauer, M. (2014), "Geco's weather forecast for the UK housing market: to what extent can we rely on google ECOnometrics?", *Journal of Real Estate Research*, Vol. 36 No. 2, pp. 253-281.

Hohenstatt, R., Kaesbauer, M. and Schaefers, 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.

Jensen, M.C. (1967), "The performance of mutual funds in the period 1945-1964", *The Journal of Finance*, Vol. 23 No. 2, pp. 389-416.

Kristoufek, L. (2013), "Can google trends search queries contribute to risk diversification?", *Nature – Scientific Reports*, Vol. 3 No. 2713, pp. 1-5.

Latoeiro, P., Ramos, S.B. and Veiga, H. (2013), "Predictability of stock market activity using Google search queries", Working Paper Nos 13-06, Statistics and Econometrics Series 05.

MacKinnon, G.H. and Al Zaman, A. (2009), "Real estate for the long term: the effect of return predictability on long-horizon allocations", *Real Estate Economics*, Vol. 37 No. 1, pp. 117-153.

Mao, H., Counts, S. and Bollen, J. (2011), "Predicting financial markets: comparing survey", News, Twitter and Search Engine Data", arXiv preprint arXiv, Quantitative Finance Papers 1112.1051.

Merton, R.C. (1987), "A simple model of capital market equilibrium with incomplete information", *The Journal of Finance*, Vol. 42 No. 3, pp. 483-510.

Moat, H.S., Curme, C., Avakian, A., Dror, Y.K., Stanley, H.E. and Preis, T. (2013), "Quantifiying wikipedia usage patterns before stock market moves", *Nature-Scientific Reports*, Vol. 3 No. 1801, pp. 1-5.

Myer, N.F.C. and Webb, J.R. (1994), "Retail stocks, retail REITs and retail real estate", *Journal of Real Estate Research*, Vol. 9 No. 1, pp. 65-84.

Oikarinen, E., Hoesli, M. and Serrano, C. (2011), "The long-run dynamics between direct and securitized real estate", *Journal of Real Estate Research*, Vol. 33 No. 1, pp. 73-104.

Preis, T., Reith, D. and Stanley, H.E. (2010), "Complex dynamics of our economic life on different scales: insights from search engine query data", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 368, pp. 5707-5719.

Preis, T., Moat, H.S. and Stanley, E. (2013a), "Quantifying trading behavior in financial markets using google trends", *Nature-Scientific Reports*, Vol. 3 No. 1684, pp. 1-6.

Preis, T., Moat, H.S., Bishop, S., Treleaven, P. and Stanley, H.E. (2013b), "Quantifying the digital traces of hurricane sandy on flickr", *Nature – Scientific Reports*, Vol. 3 No. 3141, pp. 1-6.

Ross, S.A. and Zisler, R.C. (1991), "Risk and return in real estate", *Journal of Real Estate Finance and Economics*, Vol. 4 No. 2, pp. 175-190.

Roulac, S. (1988), "How to value real estate securities", *The Journal of Portfolio Management*, Vol. 14 No. 3, pp. 35-39.

Sebastian, S. and Schaetz, A. (2009), "Real estate equities – real estate or equities?" EPRA Research, Brussels.

Simon, H.A. (1955), "A behavioral model of rational choice", *The Quarterly Journal of Economics*, Vol. 69 No. 1, pp. 99-118.

Tsai, M.-S. and Chiang, S.-L. (2013), "The asymmetric price adjustment between REIT and stock markets in Asia-Pacific markets", *Economic Modelling*, Vol. 32, pp. 91-99.

US Social Security Agency (2014), "Top names over the last 100 years", available at: www.ssa.gov/oact/babynames/decades/century.html (accessed June 28, 2014).

Vlastakis, N. and Markellos, R.N. (2012), "Information demand and stock market volatility", *Journal of Banking & Finance*, Vol. 36 No. 6, pp. 1808-1821.

Wang, K., Erickson, J. and Chan, S.H. (1995), "Does the REIT stock market resemble the general stock market?", *Journal of Real Estate Research*, Vol. 10 No. 4, pp. 445-460.

Wu, L. and Brynjolfsson, E. (2014), "The future of prediction: how google searches foreshadow housing prices and sales", working papers, Wharton School, University of Pennsylvania.

Further reading

Preis, T., Reith, D. and Stanley, H.E. (2011), "Complex dynamics of our economic life on different scales: insights from search engine query data", *Philosophical Transactions of the Royal Society A*, Vol. 368 No. 1933, pp. 5707-5719.

Corresponding author

Karim Rochdi can be contacted at: karim.rochdi@wiwi.uni-regensburg.de

(The appendix follows overleaf.)

Appendix 1

Table AI.
Overview of Google search volume indices (real estate subset)

Code	Search terms	Category	Subcategory	Topic
re_Apartments & Residential Rentals_Subcat	Apartments & Residential Rentals_Subcat	–	Apartments & Residential Rentals	–
re_Commercial & Investment Real Estate_Subcat	Commercial & Investment Real Estate_Subcat	–	Commercial & Investment Real Estate	–
re_commercial properties + commercial property	commercial properties + commercial property	–	–	–
re_commercial real estate	commercial real estate	–	–	–
re_condos + condo + condominium + condominiums	condos + condo + condominium + condominiums	–	–	–
re_foreclosure + foreclosures	foreclosure + foreclosures	–	–	–
re_health care reit + avalonbay communities + vornado realty + prologis + boston properties + equity residential + ventas + hcp + public storage + simon property group	health care reit + avalonbay communities + vornado realty + prologis + boston properties + equity residential + ventas + hcp + public storage + simon property_group	–	–	–
re_home + home_cat	home + homes	Real Estate	–	–
re_house + houses + housing_cat	house + houses + housing	Real Estate	–	–
re_jll + jones lang lasalle + cbre + cb richard ellis + colliers + dtz + cushman wakefield + knight frank + savills + grubb ellis + new mark grubb + marcus millichap	jll + jones lang lasalle + cbre + cb richard ellis + colliers + dtz + cushman wakefield + knight frank + savills + grubb ellis + new mark grubb + marcus milichap			
re_nareit	nareit	–	–	–
re_properties + property	properties + property	–	–	–
re_property agent + real estate agent	jproperty agent + real estate agent	–	–	–
re_Property Development_Subcat	Property Development_Subcat	–	Property Development	–

(*continued*)

Table AI.

Code	Search terms	Category	Subcategory	Topic
re_property fund + property funds	property fund + property funds	–	–	–
re_Property Inspections & Appraisals_Subcat	Property Inspections & Appraisals_Subcat	–	Property Inspections & Appraisals	–
re_Property Management_Subcat	Property Management_Subcat	–	Property Management	–
re_property tax	property tax	–	–	–
re_real estate	real estate	–	–	–
re_Real Estate Agencies_Subcat	Real Estate Agencies_Subcat	–	Real Estate Agencies	–
re_real estate agent + real estate agents + real estate agencies	real estate agent + real estate agents + real estate agencies			
re_real estate blog + real estate blogs	real estate blog + real estate blogs	–	–	–
re real estate company + real estate companies	real estate company + real estate companies	–	–	–
re_real estate development	real estate development	–	–	–
re_real estate funds + real estate fund + real estate investment fund + real estate investment funds	real estate funds + real estate fund + real estate investment fund + real estate investment funds	–	–	–
re_Real estate investment trust	Real estate investment trust	–	–	Real estate investment trust
re_real estate investment trust + real estate investment trusts + reit + reits	real estate investment trust + real estate investment trusts + reit + reits	–	–	–
re_real estate investment + real estate investments + investing in real estate	real estate investment + real estate investments + investing in real estate	–	–	–
re_real estate investor + real estate investors	real estate investor + real estate investors	–	–	–
re_Real Estate Listings_Subcat	Real Estate Listings_Subcat	–	Real Estate Listings	–

(*continued*)

Table AI.

Code	Search terms	Category	Subcategory	Topic
re_real estate market + real estate markets	real estate market + real estate markets	–	–	–
re_real estate news	real estate news	–	–	–
re_Real Estate_Cat	Real Estate Cat	Real Estate	–	–
re_realty investment + realty investments	realty investment + realty investments	–	–	–
re_realty trust + realty trusts	realty trust + realty trusts	–	–	–
re_reit index + reit indices + reit etf	reit index + reit indices + reit etf	–	–	–
re_reit investment + reit investments + reit investing + reits investing + investing in reit + investing in reits	reit investment + reit investments + reit investing + reits investing + investing in reit + investing in reits	–	–	–
re_rent + renting + lease + leasing	rent + renting + lease + leasing	–	–	–
re_report_cat	report	Real Estate	–	–
re_vacancy_cat	vacancy	Real Estate	–	–

Notes: This table shows the search volume indices used in this study. The regional filter for all terms and (sub-) categories is set to queries submitted in the USA. The main focus is on individual search terms without any category filter. For search terms that could lead to misunderstandings, we obtain queries from within categories (_cat). In addition, we incorporate entire categories (_Cat), subcategories (_Subcat) and a number of topics. Appendices 1 and 2 present the real estate and finance subset, respectively

Appendix 2

Table AII. Overview of google search volume indices (finance subset)

Code	Search Terms	Category	Subcategory	Topic
fi_bonds + bond	bonds + bond	–	–	–
fi_Business News Subcat	Business News_Subcat	–	Business News	-
fi_Capital market	Capital market	–	–	–
fi_Cash	Cash	–	–	–
fi_Commodities & Futures Trading_Subcat	Commodities & Futures Trading_Subcat	–	Commodities & Futures Trading	–
fi_crash	crash	–	–	–
fi_Credit & Lending_Subcat	Credit & Lending_Subcat	–	Credit & Lending	–
fi_crisis	crisis	–	–	–
fi_Currencies & Foreign Exchange_Subcat	Currencies & Foreign Exchange_Subcat	–	Currencies & Foreign Exchange	–
fi_debt	debt	–	–	–
fi_default	default	–	–	–
fi_dividend + dividends	dividend + dividends	–	–	–
fi_dow jones	dow jones	–	–	–
fi_earnings + earning	earnings + earning	–	–	–
fi_economic growth	economic growth	–	–	–
fi_economy	economy	–	–	–
fi_Employment	Employment	–	–	Employment
fi_fed	fed	–	–	–
fi_Finance_Cat	Finance Cat	Finance	–	–
fi_financial markets + financial market	financial markets + financial market	–	–	–
fi_fund + funds	fund + funds	–	–	–
fi_gains	gains	–	–	–
fi_gold	gold	–	–	–
fi_Gross domestic product	Gross domestic product	–	–	Gross domestic product
fi_income	income	–	–	–
fi_inflation	inflation	–	–	–
fi_interest rates + interest rate	interest rates + interest rate	–	–	–
fi_investor relations + investor relation	investor relations + investor relation	–	–	–

(*continued*)

Table AII.

Code	Search Terms	Category	Subcategory	Topic
fi_investor + investors	investor + investors	–	–	–
fi_losses + loss	losses + loss	–	–	–
fi_news	news	–	–	–
fi_News_Cat	News_Cat	News	–	–
fi_oil	oil	–	–	–
fi_optionshouse + etrade + schwab + interactive brokers + loyal three + scott trade + ameritrade + optionsxpress	optionshouse + etrade + schwab + interactive brokers + loyal three + scott trade + ameritrade + optionsxpress –	–	–	
fi_politics	politics	–	–	–
fi_profit + profits	profit + profits	–	–	–
fi_return + returns	return + returns	–	–	–
fi_revenue	revenue	–	–	–
fi_risk	risk	–	–	–
fi_s&p 500	s&p500	–	–	–
fi_Stock	Stock	–	–	Stock
fi_stock markets + stock market	stock markets + stock market	–	–	–
fi_trade	trade	–	–	–
fi_trusts + trust	trusts + trust	–	–	–
fi_war	war	–	–	–
fi_yahoo finance + google finance + bloomberg	yahoo finance + google finance + bloomberg	–	–	–
fi_yield + yields	yield + yields	–	–	–

Notes: This table shows the search volume indices used in this study. The regional filter for all terms and (sub-) categories is set to queries submitted in the United States. The main focus is on individual search terms without any category filter. For search terms that could lead to misunderstandings, we obtain queries from within categories (_cat). In addition, we incorporate entire categories (_Cat), subcategories (_Subcat) and a number of topics. Appendices 1 and 2 present the real estate and finance subset, respectively

Appendix 3

Rank	Search Terms	Impact Factor
1	Real Estate_Cat	100.00
2	Real Estate Listings	74.87
3	Apartments & Residential Rentals	46.98
4	home + homes	30.62
5	real estate	26.67
6	house + houses + housing	26.67
7	rent + renting + lease + leasing	20.25
8	properties + property	19.90
9	Real Estate Agencies_Cat	19.26
10	condos + condo + condominium + condominiums	7.02
11	Property Management_Cat	5.43
12	Property Development_Cat	4.44
13	Property Inspections & Appraisals_Cat	3.95
14	Commercial & Investment Real Estate_Cat	3.95
15	foreclosure + foreclosures	2.60
16	property tax	1.92
17	health care reit + avalonbay communities + vornado realty + prologis + boston properties + equity residential + ventas + hcp + public storage + simon property group	0.78
18	real estate agent + real estate agents + real estate agencies	0.75
19	commercial real estate	0.74
20	jll + jones lang lasalle + cbre + cb richard ellis + colliers + dtz + cushman wakefield + knight frank + savills + grubb ellis + new mark grubb + marcus millichap	0.61
21	property agent + real estate agent	0.41
22	real estate company + real estate companies	0.45
23	commercial properties + commercial property	0.37
24	Real estate investment trust	0.30
25	real estate investment + real estate investments + investing in real estate	0.28
26	real estate investment trust + real estate investment trusts + reit + reits	0.28
27	real estate market + real estate markets	0.23
28	real estate news	0.18
29	report	0.19
30	real estate development	0.14
31	real estate investor + real estate investors	0.11
32	realty trust + realty trusts	0.10
33	real estate funds + real estate fund + real estate investment fund + real estate investment funds	0.09
34	realty investment + realty investments	0.06
35	real estate blog + real estate blogs	0.06
36	property fund + property funds	0.05
37	vacancy	0.05
38	reit index + reit indices + reit etf	0.03
39	reit investment + reit investments + reit investing + reits investing + investing in reit + investing in reits	0.02
40	nareit	0.01

Notes: This table reports the ranking of all 40 real-estate-related searches sorted by their search volume, computed as a relative impact measure. Because Google Trends does not supply the absolute number of searches, the impact factor is determined by pair-wise comparisons of each individual SVI with the benchmark SVI "real estate." The findings show that the overall search volume of Google search terms is positively correlated with their investment performance (Kendall's $\tau = 0.24$, z-stat $=$ 2.22; Spearman's $\rho = 0.30$, t-stat $= 1.90$)

Table AIII. Google search relevance impact factor of real-estate-specific search terms

This article has been cited by:

1. Nicole Braun IREBS - International Real Estate Business School, University of Regensburg, Regensburg, Germany . 2016. Google search volume sentiment and its impact on REIT market movements. *Journal of Property Investment & Finance* **34**:3, 249-262. [Abstract] [Full Text] [PDF]
2. Marian Alexander Dietzel International Real Estate Business School (IRE|BS), University of Regensburg, Regensburg, Germany . 2016. Sentiment-based predictions of housing market turning points with Google trends. *International Journal of Housing Markets and Analysis* **9**:1, 108-136. [Abstract] [Full Text] [PDF]