DOI: 10.1111/eufm.12325

EUROPEAN FINANCIAL MANAGEMENT WILEY

ORIGINAL ARTICLE

Leading indicators for US house prices: New evidence and implications for EU financial risk managers

Tobias Basse[1,2] | **Frederik Kunze**[1] | **Miguel Rodriguez Gonzalez**[3] | **Danilo Saft**[1,4]

[1]Norddeutsche Landesbank Girozentrale, Hannover, Germany

[2]Touro College Berlin, Berlin, Germany

[3]Centre for Risk and Insurance, Leibniz University Hannover, Hannover, Germany

[4]Department of Economic Sciences and Media, Group for Service Information Systems Engineering, Ilmenau Technical University, Ilmenau, Germany

Correspondence
Miguel Rodriguez Gonzalez, Centre for Risk and Insurance, Leibniz University Hannover, Otto-Brenner-Straße 7, 30159 Hannover, Germany.
Email: mr@ivbl.uni-hannover.de

Abstract

This study draws on machine learning as a means to causal inference for econometric investigation. We utilize the concept of transfer entropy to examine the relationship between the US National Association of Home Builders Index and the S&P CoreLogic Case-Shiller 20 City Composite Home Price Index (SPCS20). The empirical evidence implies that the survey data can help to predict US house prices. This finding extends the results of Granger causality tests performed by Rodriguez Gonzalez et al. in 2018 using a new machine learning approach that methodologically differs from traditional methods in empirical financial research.

KEYWORDS

financial risk management, leading indicators, machine learning, transfer entropy, US house prices

JEL CLASSIFICATION

C58, G01, G11, R30

We would like to thank the editor John A. Doukas, an anonymous associated editor and an anonymous referee for extremely helpful comments. Their efforts have been essential in significantly improving this paper. We also have benefited greatly from comments made by Tristan Hirsch, Bernd Krampen, J.-Matthias v. d. Schulenburg, Christoph Schwarzbach, Günter Vornholz, and Christoph Wegener.

1 | INTRODUCTION

Risk management frameworks today are supposed to be forward looking (see e.g., Breden, 2008; Jorion, 2009). More specifically, Rochette (2009) has argued convincingly that all organisations should try to not forgo the advantages and opportunities that an adequate risk management programme can deliver, and that in order to ensure this a strong risk culture has to be implemented, that creates an environment where risk managers are not just waiting for bad things to happen. This could be of special importance for the financial services industry. As a matter of fact, Vazquez and Federico (2015) have noted that the global financial crisis has casted some doubts on the quality of bank risk management practices. These concerns have also been felt by bank regulators (see e.g., Liu et al., 2011; Vazquez & Federico, 2015). Moreover, it has also been questioned how the subprime crisis —a problem in a rather small segment of overall US financial markets—was able to hurt the global economy so badly (see Bullard et al., 2009; Eichengreen et al., 2012). This important question still is not answered in an adequate way yet. In any case, US house prices without a doubt have played a key role in this crisis (see e.g., Bullard et al., 2009; Wegener et al., 2019). As will be discussed in more detail, the European financial services industry also seems to be of some importance in this context (see e.g., Mizen, 2008; Noeth & Sengupta, 2012). Similarily, Rodriguez Gonzalez et al. (2018) have argued convincingly that risk managers should try to develop an early warning indicator system for real estate prices in the United States. Financial market prices that are determined in highly liquid markets should reflect information about the future. Stock prices, for example, ought to be helpful in forecasting the development of corporate earnings at the level of individual firms and the business activity in the economy as a whole (see e.g., Aylward & Glen, 2000; Goddard et al., 2006). Therefore, it could be argued that the stock market should be a leading indicator of corporate profitability and economic growth. Assuming a certain degree of market efficiency other prices that are determined by the activities of rational buyers and sellers in the financial sphere should also provide relevant information about future developments. However, with regard to house prices it has been argued convincingly that the property market could be inefficient due to its heterogeneity (see, most importantly, Clayton et al., 2009; Dietzel et al., 2014). Moreover, Hausler et al. (2018) have noted that real estate investors may be especially sensitive to changes in sentiment due to the specific characteristics of real estate markets such as the relatively low market transparency or long transaction periods. Therefore, sentiment indicators are likely to be perfect candidates on which to base an effective early warning system. Thus, financial time series that belong to the two categories house prices and sentiment indicators seem to be particularly suitable candidates when searching for use cases trying to implement forward looking financial risk management approaches that use the concepts of Granger causality or entropy. In this context, the concept of Granger causality obviously is of special importance. This is more or less true by definition because a certain variable is said to be Granger causing a second variable when it is useful in forecasting future values of the other time series (see, most importantly, Granger, 1969, 1988). One possibility is that a variable Granger causes another variable and that in the same time this second time series also Granger causes the first time series. In this case there are feedback effects among the two variables examined and consequently there is bidirectional Granger causality (see e.g., Hiemstra & Jones, 1994; Xie & Chen, 2014). Unidirectional Granger causality, on the contrary, is said to exist when one variable Granger causes the other variable but not vice versa.

The concept of Granger causality is of some importance in the field of real estate economics. This approach has, for example, been used to analyze the lead–lag relationship between house prices in different neighbouring regions (see e.g., Blake & Gharleghi, 2018; Teye et al., 2017) and between house prices and macroeconomic variables or the stock market (see, amongst others, Green, 2002; Luo et al., 2007). Moreover, this technique is also quite popular in the macroeconomic literature. In fact, Granger causality is an important concept for econometricians searching for leading indicators of economic activity (see e.g., Breitung & Candelon, 2006; Huh, 2002). Rodriguez Gonzalez et al. (2018) have employed this empirical approach to assess whether the National Association of Home Builders (NAHB) Market Index—an important sentiment indicator for economic activity in the US real estate sector—respectively its subcomponents can be a suitable leading indicator for property prices. The empirical evidence reported in this study seems to indicate that the NAHB data indeed can help to forecast house prices in the United States.

The tools developed in the field of machine learning right now are starting to have an impact on the real estate economics literature. Pioneering work in this area was done by Hausler et al. (2018). This very important study uses machine learning techniques to construct sentiment indicators for real estate markets. Our paper however adopts a completely different approach. We employ techniques of machine learning in combination with the concept of transfer entropy to improve early warning systems that are based on sentiment indicators and have been constructed using the tools of traditional time series analysis. Even though econometrics has a long-standing background in the application of Granger causality (e.g., from Cheng, 1979 and Geweke, 1984 to Luu Duc Huynh, 2019 and Osiobe, 2020), we further will consider emerging approaches from the field of machine learning and also will use the technique of transfer entropy (see e.g., Behrendt & Prange, 2021; Dimpfl & Peter, 2014). Here, we wish to highlight machine learning workflows based on transfer entropy parameter studies as promising frameworks for causal inference to aid the development of econometric models. More specifically, we use the concept of transfer entropy in combination with some tools from the field of machine learning to validate the results of the important study by Rodriguez Gonzalez et al. (2018) by employing a completely different methodology. As will be discussed later on in more detail the empirical research strategy used here can be helpful to cope with potential problems due to nonlinearities. Additionally, employing the transfer entropy approach in combination with techniques of machine learning might also be interesting because it will provide additional information about the relationship under investigation (especially with regard to the selection of the number of time lags to be considered in empirical models). At this point it has to be noted that the empirical evidence presented below just shows one possible application of the concept of transfer entropy in the field of financial economics. Obviously, this technique can be employed to analyze numerous other questions that are of relevance in economics and finance. As a matter fact, some ideas for future research will be discussed in the conclusion. However, it also has to be noted that the use case examined here—namely, the search for an appropriate leading indicator for house prices that can be helpful implementing forward-looking risk management approaches—is very important. The results of our empirical study, therefore, are highly relevant by themselves. This fact has clearly been demonstrated by the major economic problems that have been caused by US subprime crisis—a disaster that also had major consequences for financial risk managers in Europe and other parts of the world. In sum, our empirical findings seem to make a valuable contribution to the literature in the fields of real estate economics, financial risk management and machine learning.

The paper is structured as follows: Section 2 briefly surveys the relevant literature focusing on empirical evidence from the United States and the United Kingdom. Section 3 then examines some general machine learning issues that are of special importance for this study. In Section 4, the concept of transfer entropy is discussed. Section 5 provides some information about the data examined. Moreover, some first empirical evidence with regard to the time series properties of the variables examined are presented here. Section 6 then reports and also discusses the empirical evidence that was obtained using the transfer entropy methodology in some detail. Before concluding in Section 8, Section 7 examines why the results of our empirical investigations are important for risk managers working in the European financial industry.

2 | SOME EMPIRICAL EVIDENCE FROM THE UNITED STATES AND THE UNITED KINGDOM

Meanwhile, numerous empirical studies have searched for a suitable leading indicator of housing activity in a number different countries (see e.g., Croce & Haurin, 2009; Rodriguez Gonzalez et al., 2018). As will be shown, the focus of the literature lies on data from the United States and from the United Kingdom. Without a doubt, sentiment indicators might be helpful in this context. As already noted, there is some empirical evidence suggesting that sentiment indicators can be helpful to predict changes in the market for real estate assets (see e.g., Dietzel et al., 2014; Tsolacos, 2012). In this context Tsolacos (2012) has argued convincingly that sentiment and confidence indicators could be particularly helpful when it comes to the identification of turning points in real estate markets. Given that that such phases are of crucial importance for investors and risk managers, the recent interest in this topic is certainly well understandable. Thus, it should come as no surprise that there is some related literature. As a matter of fact, there meanwhile are a number of relevant studies examining data from the United States. Moreover, some applied econometricians recently also have analyzed time series from the United Kingdom. As already noted, the NAHB housing market index is usually considered to be the most popular leading indicator for US real estate prices and other variables that are related to housing activity (see, amongst others, Goodman, 1994; Marcato & Nanda, 2016). The ability of this sentiment indicator to act as leading indicators of housing activity in the United States is discussed quite controversially in the literature. Marcato and Nanda (2016), for example, have shown that the NAHB housing market index can be helpful predicting house prices in the United States. To do so they have employed Granger causality tests. However, the empirical evidence presented by Croce and Haurin (2009) is less promising. They have examined the ability of the NAHB sentiment indicator to forecast US housing activity by also performing Granger causality tests. Though the results reported in this paper seem to imply that the NAHB data can indeed help to predict some important time series measuring real estate activity in the United States, there are still problems in a number of cases. Moreover, in the cases where Granger causality between the NAHB housing market index and other relevant time series from the US housing market has been detected Croce and Haurin (2009) have usually found evidence for the existence of bidirectional Granger causality. Thus, there seem to be, possibly nonlinear feedback effects. This would certainly be a problem using the NAHB housing market index as a leading indicator for the US real estate market activity employing the traditional techniques of time series econometrics. Additional research that can define the extent of causality and recognise possibly nonlinear feedback loops would be necessary to improve our understanding about the way the two time series are related to each other and about the predictive power of the sentiment indicator. Ideally, this would be an analysis that is updated and evolves with the time series themselves. This challenge revisits one of the major aspects of machine learning workflows with their ability to update analysis results on the fly and learn even nonlinear relationships from the data. Such relationships, as we will see in this article, can be captured and measured not only in existence, but also extent, by the concept of transfer entropy.

Given the importance of the question whether the NAHB data constitutes a suitable basis for an early warning system for US real estate activity Rodriguez Gonzalez et al. (2018) have tested for Granger causality employing the approach suggested by Toda and Yamamoto (1995). The main advantage of this approach is that it does not require major pretesting efforts that can cause problems (e.g., cointegration tests). Only the order of integration of the time series under investigation has to be determined. There is no need to examine variables in differences. Moreover, the test procedure is also very popular because of a Monte Carlo study by Zapata and Rambaldi (1997) that has shown some very favourable characteristics of this approach. This technique requires the estimation of an augmented vector autoregressive (VAR) model in levels. The empirical evidence reported by Rodriguez Gonzalez et al. (2018) shows that the VAR model should consider three to four time lags. Using this approach there are clear signs for Granger causality running from the NAHB housing market index to the S&P/Case-Shiller 20 city house price index. Consequently, the data compiled by the NAHB indeed seems to be helpful forecasting real estate prices in the United States. Another study of great importance for our empirical work is Hausler et al. (2018) because this paper uses techniques of machine learning to develop a news-based approach for prediction purposes examining data from the US real estate market. More specifically, the authors employ machine learning techniques to construct a text-based sentiment analysis tool that can be used to construct a useful leading indicator. The paper has reported very favourable empirical evidence. It certainly is an important door opener for the use of machine learning techniques in the field of real estate economics.

With regard to international evidence on the relationship between sentiment indicators and house prices there is some highly relevant recent empirical work from the United Kingdom. In fact, Wood (2003) and McLaren and Shanbhogue (2011) have suggested to utilize data from the Housing Market Survey which is compiled by the Royal Institution of Chartered Surveyors (RICS) to forecast house prices in this country. McLaren and Shanbhogue (2011) have shown that the combination of the RICS data with Internet search data (more specifically, people interested in both buying and selling properties) can be helpful to forecast house prices. Moreover, Kunze et al. (2020), meanwhile, have reported quite favourable empirical evidence testing for Granger causality between the RICS survey data and the level of house prices in the United Kingdom. Doing so they also have employed the approach that has been suggested by Toda and Yamamoto (1995). Therefore, Kunze et al. (2020)—which also builds on Rodriguez Gonzalez et al. (2018)—is of some importance for the empirical evidence to be presented later on.

3 | SOME GENERAL THOUGHTS ABOUT MACHINE LEARNING IN THE CONTEXT OF TRANSFER ENTROPY

The history of machine learning is closely linked to the development of computational capabilities and methodologies in computer science (see Athey, 2018). As key properties of data—such as variety, velocity and volume—grew over the past decades, traditional systems for handling and analysing data became infeasible. The exponential growth of what has been dubbed the big data trend was accompanied by an exponential growth of computational power. Computer science and software development—as original fields of growth—continuously created efficient algorithms and workflows to better work with multivariate, possibly incomplete, nonlinear, and unstructured data. Ease of utilization and flexible workflow adjustments have been at the heart of this development since. Machine learning best practices and frameworks emerged largely independent from established methodologies in fields such as econometrics. In this context, Athey (2018) has noted that there has not yet been much diffusion of econometric concepts into machine learning, but there is some overlap in common statistical assets. Above that, we argue that machine learning best practices will diffuse into the field of econometrics as it will become augmented by the tools and data science workflows codeveloped therein.

Machine learning applications and thorough methodologies have existed for a few decades already (see Mjolsness & Decoste, 2001). Nevertheless, especially breakthroughs in analysis of unstructured data (e.g., Hinton et al., 2012) and value creation associated with formerly underutilized data sets have led to accelerated attention, financing and development in all current disciplines of artificial intelligence, including traditional machine learning and data science approaches. As of yet, the focus of these disciplines has been seen in predictive analytics and pattern recognition (the latter centred around unsupervised clustering and dimensionality reduction) in big data sets (e.g., Athey, 2018; López de Prado, 2019). Basuchoudhary et al. (2017) underline the focus of prediction when utilizing machine learning for economic tasks. This seems natural from two perspectives: As it is hard from an epistemological perspective to understand large, frequently updated data sets and find their underlying structure with traditional or manually quickly traceable algorithms, a layer of abstraction for prediction and other tasks was built. Using this machine learning toolbox layer, researchers are able to efficiently and effectively provide higher predictive power at the cost of being able to apply models of complexity. Direct prediction on the contrary, offers many decision makers in the economy higher value-creating power than just deriving complex relationships and models from the data, at a minimal additional effort. However, we do not see the development to stop at prediction and reject efforts to categorise machine learning by the criterion of trying to predict something. As research continues, concepts well established in econometrics such as causal inference gain modern counterparts in machine learning. In fact, López de Prado (2019) argues that for every step in econometrics, machine learning already features a homologous step.

The interfaces of the toolboxes that were built in machine learning are easy to use for such tasks and offer tuning capabilities combined with high performance. This factor of automation of efficient workflows will be the driver of said continuous diffusion process of machine learning frameworks into different fields of science. The swiftness of the development in machine learning is underlined by the fact that the languages R (see R Core Team, 2019) and Python (see van Rossum & Drake, 2009) rank high in the IEEEs spectrum of programming languages (Cass, 2018). Python has been gaining wide popularity for its ease of use and flexibility, whereas Rs continued best practice support of data wrangling and analysis tasks will keep it relevant. The increasing amount of libraries and packages available to solve specific tasks enable quick and novel solution processes in both of these languages.

The focus on performance and being able to handle even complex data resulted in approaches that 'let the data speak for itself' (see Bzdok et al., 2017). After enabling the utilization of unstructured data in, for example, the medical sciences, we have not arrived at a point where we can demonstrate said homologous concepts in machine learning for original tasks in econometrics. We will demonstrate an efficient causal inference workflow using machine learning toolboxes not relying on Granger causality or the concept of cointegration. Moreover, we offer a methodology utilizing modern machine learning pipelines, modern big data handling technology stacks and tools. Importantly, this approach can be generalised to create machine learning supported causal inference pipelines in economic investigation to many structured and unstructured data set. Hence, though we demonstrate the bivariate, nonlinear numerical case as an introduction above what is possible with Granger causality measurements, this workflow can be extended further to multivariate nonlinear numeric, discrete and even textual data sets.

As a first principle in this workflow, we utilize a different empirical approach—namely transfer entropy—to examine the relationship between the time series analyzed by Rodriguez Gonzalez et al. (2018) with toolboxes and methods available from the machine learning realm. This concept recently has become quite popular in financial economics (see, among others, Behrendt & Prange, 2021; Dimpfl & Peter, 2014).

4 | TRANSFER ENTROPY AND AVAILABLE MACHINE LEARNING LIBRARIES FOR R AND PYTHON

Shannon entropy (see Shannon, 1948) is a measure of uncertainty of description of a random variable X with a distribution $P(x_i)$ by a certain number of base 2 bits, that is, it is possible to measure the quantity of information contained in X in bits by defining the average number of bits needed to encode draws from $P(x_i)$ via

$$H(X) = -\sum_{i=1}^{n}[P(x_i)\log P(x_i)].$$

Other measures of entropy exist that have a more parametrised approach to measuring information content, for example, Rényi entropy (see Rényi, 1970). Shannon entropy is the most widely used criterion to measure information content within a discrete variable.

Transfer entropy was introduced in Schreiber (2000), relying on Shannon entropy and the Kullback-Leibler distance, as a quantitative measure of statistically significant transfer of information in time series that is able to distinguish between driving and responding elements in such systems (see Bossomaier et al., 2016b; Simon et al., 2019). It, therefore, aims at detecting dynamic causation links between a paired time series (see Syczewska & Struzik, 2014). Simon et al. (2019) summarise accordingly, that information flow from a process J to a process I can be measured by quantifying the deviation from the generalised Markov property $p(i_{t+1} \mid i_t^{(k)}) = p(i_{t+1} \mid i_t^{(k)}, j_t^{(l)})$ given that I_{t+1} is conditional on the k previous observations (Markov process of order k) and J_{t+1} is conditional on the l previous observations of J (Markov process of order l) as well as relying on the Kullback–Leibler distance (see Schreiber, 2000). Therefore, transfer entropy based on Shannon entropy determines an information flow $T_{J\rightarrow I}$ by calculating

$$T_{J\to I}(k,l) = \sum_{i,j=0}^{n} p\left(i_{t+1}, i_t^{(k)}, j_t^{(l)}\right) \times \log\left(\frac{p\left(i_{t+1}, i_t^{(k)}, j_t^{(l)}\right)}{p\left(i_{t+1}|i_t^{(k)}\right)}\right).$$

Transfer entropy is a model free measure of information flow from one time series to another (see Bossomaier et al., 2016a; Vicente et al., 2011). This greatly distinguishes it from Granger causality. However, according to Bossomaier et al. (2016b), transfer entropy may be considered to be a generalisation of Granger causality and is in this respect, able to answer the question how much information is transferred at a certain timestep from the past of one time series to the current state of another time series. Bossomaier et al. (2016b) emphasise that transfer entropy is an asynchronous measure of information flow and, therefore, able to quantify differing amounts of information flow from a time series X to a time series Y opposed to the flow from Y to X. Previous entropy based measures (e.g., mutual information) did not expose this directional characteristic.

Syczewska and Struzik (2014) argue that financial time series often show autoregressive conditional heteroscedasticity and show non-Gaussian statistics alongside nonlinear correlations. In this respect, they give an overview of Granger causality tests for nonstationary financial time series and refer to the method published by Toda and Yamamoto (1995) as well as concepts for nonlinear Granger causality that may be applied in financial analysis, yet require more complicated steps to prepare and analyze such data. Dimpfl and Peter (2012) state that Granger causality has been a predominant measure to detect relationships between time series, however, its insights may often only be used to interpret the existence, and possibly compare statistics, rather than measure the exact quantity of information flow in financial time series as several assumptions about the underlying statistics and dynamics must be met for a quantitative interpretation of Granger causality. Transfer entropy on the contrary, according to Dimpfl and Peter (2012), is not limited to the assumptions made by the predominantly applied measures of Granger causality, especially regarding linear dynamics. Other methods, such as the Hasbrouck information share, assume cointegration between time series, whereas transfer entropy does again not have such prerequisites. Therefore, Dimpfl and Peter (2012) state that transfer entropy is applicable even if one cannot be sure about whether the assumptions required by the standard models are met by the data. Considering the abovementioned discussions, transfer entropy is a promising generalized measure for quantifying the extent and direction of information flow between financial time series.

Moreover, a derivative of transfer entropy, called 'effective transfer entropy', has been introduced by Marschinski and Kantz (2002) to account for bias effects from small sample sizes. Towards measuring bivariate transfer entropy and effective transfer entropy, an established library for the programming language R exists (Simon et al., 2019) and similar approaches to counter bias effects exist in the well-established approach developed by Wollstadt et al. (2013) that can quantify bi- and multivariate transfer entropy.

5 | DATA AND INITIAL ANALYSIS

This empirical study tries to validate the results reported by Rodriguez Gonzalez et al. (2018) employing a completely different methodological approach. Therefore, the same data set (variables and sample) is also examined here. As already noted, the NAHB housing market index is a very popular leading indicator for US real estate prices and other variables that are related to housing activity that is widely observed by investors in different segment of the global financial market. This time series is based on the results of a monthly survey among the members of the National Association of Home Builders asking the participants for their attitudes and expectations with regard to the demand for single-family homes and house market conditions in general (see e.g., Rodriguez Gonzalez et al., 2018; Wilcox, 2015). The measure of US real estate prices analyzed by Rodriguez Gonzalez et al. (2018)—and, therefore, also here—is the S&P/Case-Shiller CoreLogic 20 City Composite Home Price Index which reflects the development of house prices in 20 metropolitan areas of the United States. This real estate price index is quite popular among financial economists and consequently is often used in applied empirical studies (see e.g., the recent studies by Huang, 2019; Ramirez, 2019). The sample examined here is from January 1995 to April 2018. The data that is used for all calculations is taken from Bloomberg.

The time series properties (order of integration) of the variables have been one important reason for Rodriguez Gonzalez et al. (2018) to use the procedure suggested by Toda and Yamamoto (1995). Nonstationarity in single time series replications needs consideration using the transfer entropy approach, too (see e.g., Behrendt & Prange, 2021). Stationarity requirements of transfer entropy measurements are usually considered and evaluated in a strict sense (stationarity in mean, variance, covariance) and, therefore mitigated, for example, by taking differences. This may result in excluding possibly important information. It is, however, still under discussion whether a strict interpretation of stationarity is necessary or whether weaker assumptions may apply for such causal inference, for example, under the presence of a confounding driving factor for the nonstationarity (see Runge, 2018). In any case, the transfer entropy measurement libraries employed here offer ways to deal with nonstationary time series under certain additional assumptions. One way is to provide replications of the nonstationary time series process to infer significance of the measured causal relations. Typically, these ensemble methods allow for taking approximately stationary cyclic repetitions under similar conditions (e.g., for different subjects in neuroscientific experiments under similar experimental setup). Taking approximately stationary subsamples of the nonstationary time series to attain the required number of repetitions, under the assumption of an only slowly changing nonstationary regime, has been mentioned as a viable method outside domains such as neuroscience—where fast changes in time series are more common (see Gómez-Herrero, 2015).

As we deem the number of samples in the considered time series long enough, we wish to include both, the nonstationary original series, as well as the stationary derivative in a hyperparameter study on transfer entropy behaviour. First of all, the results of unit root tests are reported in the Tables 1 and 2. Here, the approach suggested by Elliott et al. (1996) is employed. This testing procedure is known for its improved power compared to more traditional unit root

TABLE 1 Unit root test results for the National Association of Home Builders (NAHB) Housing market index

This table reports the results of Elliott–Rothenberg–Stock unit root test and the appropriate critical values (5% error level) examining the NAHB Housing market index (in levels and first differences). Null hypothesis: Time series has a unit root; Exogenous: Constant.

	Data in levels	Data in first differences
Elliott–Rothenberg–Stock test statistic	21.554	0.225
5% critical value	3.179	3.179

TABLE 2 Unit root test results for the S&P/Case-Shiller CoreLogic 20 City Composite Home Price Index (SPCS20)

This table reports the results of Elliott–Rothenberg–Stock unit root test and the appropriate critical values (5% error level) examining the S&P/Case-Shiller CoreLogic 20 City Composite Home Price Index (in levels and first differences). Null hypothesis: Time series has a unit root; Exogenous: Constant.

	Data in levels	Data in first differences
Elliott–Rothenberg–Stock test statistic	20.751	0.645
5% critical value	3.178	3.178

tests (see, amongst others, Cooray & Wickremasinghe, 2007; Maddala & Wu, 1999). Cooray and Wickremasinghe (2007), for example, have stressed that the approach that has been suggested by Elliott et al. (1996) dominates other commonly used unit root tests when a time series has an unknown mean or a linear trend.

The empirical findings obtained using this approach seem to imply that both time series are nonstationary variables integrated of order one. Consequently, the results reported by Rodriguez Gonzalez et al. (2018) are confirmed using a different unit root testing procedure.

For calculating transfer entropy, different estimators of the mutual information distribution of the processes can be employed (see Lizier, 2014). A Gaussian estimator assumes approximately a pairwise jointly Gaussian distribution of all processes. Figure 1 shows that for diff(NAHB) versus diff(SPCS20), we observe an approximately jointly Gaussian pattern. For NAHB versus SPCS20, strong deviations from a joint Gaussian distribution pattern are exposed. The Gaussian estimator further only exposes linear relations in the data. That being said, this estimator is far less computationally expensive than others and may be a good start for a quick coarse overview when traversing large parameter spaces (as suggest by Lizier, 2014). Because of the limitations of the Gaussian estimator, we will concentrate on the Kraskov–Stögbauer–Grassberger mutual information estimator (KSG) (see Kraskov et al., 2004) for narrow interpretations, but will compare the behaviour to the Gaussian estimator for completeness of the parameter study.

6 | MACHINE LEARNING WORKFLOW AND DERIVED EMPIRICAL EVIDENCE

The machine learning workflow we developed constitutes advances over traditional econometric approaches as we were able to perform a hyperparameter grid search in a parallel manner. Scanning the hyperparameter space and drawing conclusions from the robustness of the observed phenomena is a scheme typical for machine learning pipelines that we deem advantageous for consideration in econometric investigations—especially in a situation such as the one at hand where some empirical studies arrive at slightly differing results depending on the methodology employed. Hence, a hyperparameter grid search traverses the meta-space of results dependent on methodological settings such as the chosen lags for causal inference, strictly stationary versus not strictly stationary processes considered, conditional mutual information estimator used for transfer entropy, and so forth. The parameter space we wish to traverse is indicated by combinations of the following value sets:

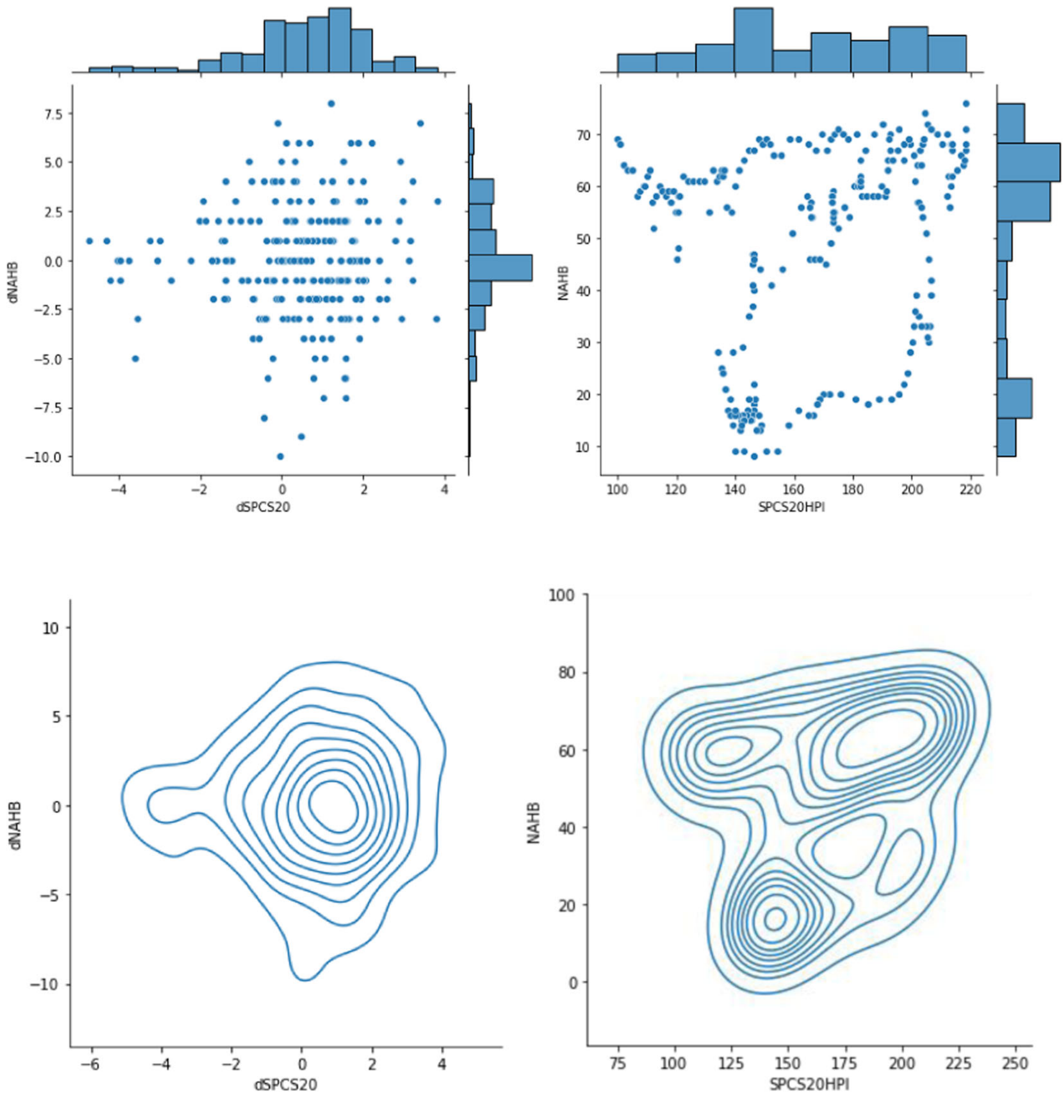

FIGURE 1 Bivariate scatter plot, marginal histograms (top row) and derived contour plots of the joint distributions (bottom row) for the differences processes (left column) and original processes (right column). These figures are plotted to gain insights about the data examined here. Processes that are considered 'jointly Gaussian' will show a circular accumulation and concentration of scatterpoints. Here, the differenced processes show an approximately jointly Gaussian distribution of their values, whereas the original time series deviate largely from this assumption

- Time series analyzed: (original, differenced)
- Subsampling method:
  - o Eighteen chunks (replications) of 13 data points in each subsample of the original full series available (239 observations)
  - o A single long series, with different start and end months to account for possible sample bias. Start and end of sample were varied randomly within a window of 40 months at the start and end of the original series. Thirty runs were completed and included as box plot

representations of the measurements to gain insights on the robustness of the results under varying subsample sets.

- Conditional Mutual Information Estimator: (JidtGaussianCMI, JidtKraskovCMI)
- Minimum Lag considered: 0 (no variation)
- Maximum Lag considered depending on the subsampling method above:
  - o 1, ..., 12
  - o 1, ..., 20

From a technical perspective, as the KSG estimation is computationally very expensive, we set up a multiprocessing pool using the *multiprocessing* library for Python. All pipelines ran on a modern 12 core CPU and a machine with 64 GB of RAM. The pipeline is optimised for using as much CPU power as possible, but adhering to the RAM limits. We observed strong RAM consumption with larger permutation settings as large amounts of surrogate data sets of the smaller original time series data are created and temporary objects for mutual information estimation and transfer entropy measurements required memory space in each parallel run.

All settings can also be run with machine learning scheduling software (see e.g., Apache Airflow, 2020) to reduce downtime and necessity of manual reruns. This can be important when analysing transfer entropy as each analysis run takes several hours or even days depending highly on the permutation settings. If constant monitoring is not possible, resources may go unused for significant amounts of time after completion of a fork if no scheduling tool is utilized. Each run of each fork saved its results using the *pickle* library.

For all experiments, we created a standard *conda* environment (see Anaconda Software Distribution, 2020). This environment was set up using Python 3.7 with its libraries *pandas*, *matplotlib* version 2.2 (for compatibility with *idtxl*), *networkx* version 2.4 (for compatibility with *idtxl*), *statsmodels* and *idtxl*. We ingested the original time series data from an Excel export directly from Bloomberg into a *pandas* data frame. We added the stationary derivatives as new columns to the dataframe for selection based on the desired current hyperparameter for the time series to be used (original or differenced). A *numpy* array of the two time series processes was created for further data processing. This included, for example, reshaping the *numpy* array into different subsets to create the abovementioned replications. Graphical plots confirmed the validity of the data. The reshaped *numpy* array was used to initialise an *IDTxl Data* object to subject it to further transfer entropy analysis in our hyperparameter study.

Using our first approach of taking only small subsamples and subjecting them to analysis, we received the results displayed in Figure 2. Considering target process 1 (i.e., testing for a causal relationship from NAHB to SPCS20) the Gaussian estimator found significant causal relationships with several settings for the maximum number of lags considered. These, may however arise due to a non-Gaussian joint distribution of the subsamples and are reported here for the completeness of investigating the behaviour of the estimators on the time series at hand. Hence, we conclude that the results indicate false positive effects of Gaussian estimators under a not-jointly Gaussian distribution of the two time series, respectively, only considering linear effects between processes.

The KSG reported significant causal relations when considering lags 0 to 6. Looking in the other direction, it is important to note that no linear relations from SPCS20 to NAHB were found, as the Gaussian estimator did not measure any significant information transfers.

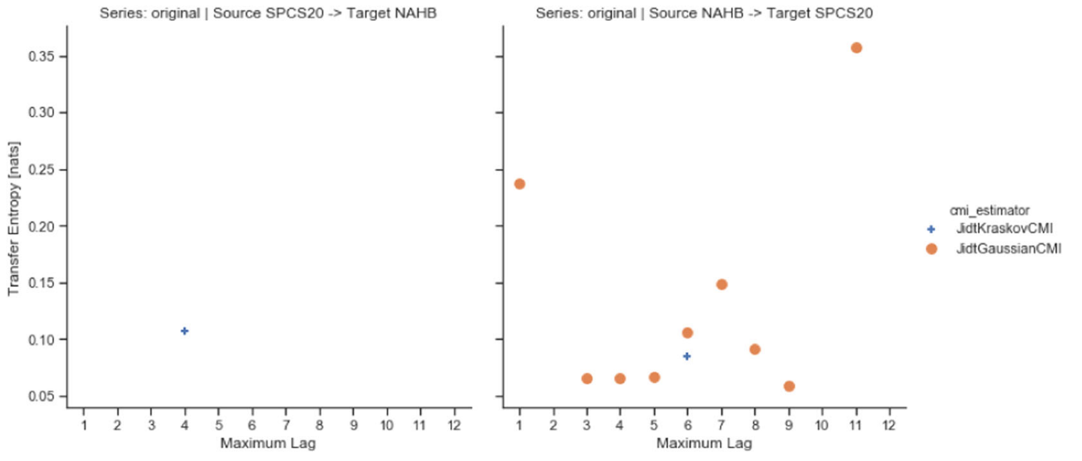

FIGURE 2 Joint transfer entropy results of a subset sampling study with 18 repetitions of 13 months. These figures show the results of the calculations searching for transfer entropy. Markers and colours indicate the CMI (conditional mutual information) estimator used. The x-axis shows the maximum lag considered in the measurement. Here, we observe several significant linear causal relations from National Association of Home Builders Index (NAHB) to SPCS20 using the Gaussian estimator. Nonlinear effects from NAHB to SPCS20 are present at lag 6. In the opposite direction, a nonlinear effect is observed at lag 4 only

However, the KSG estimator was able to uncover nonlinear relationships and reported a significant overall information transfer when considering lags 0 to 4. The results here are reported for completeness of methodological opportunities in a machine learning oriented hyperparameter search setting. The shortness of the time series subsamples is probably not suitable to uncover all causal relations, but the results are interesting in terms of comparison opportunities to other hyperparameter settings to gain meta-level knowledge about causal inference behaviour on these processes.

As we wish to include both the nonstationary original series as well as the stationary (i.e., differenced) derivative, we will now continue our hyperparameter study using stationary, that is, diff(NAHB) and diff(SPCS20), time series. We use differenced time series, as we will not need to rely on the presence of several replications of short approximately stationary subsamples to measure valid transfer entropy results. Hence, we will be able to observe transfer entropy for subsamples covering many years of the time series. By repeatedly taking small variations of the start and end points of the subsamples, we can observe the sensitivity of the transfer entropy measurement at each lag by doing several measurements with slight variations. Therefore, we constructed 30 runs with differing start and end points, each of guaranteed length of over 160 months. Figure 3 shows the results of the analysis. The left column of Figure 3 shows the boxplots for the measured information transfers from NAHB to SPCS20 in two subfigures. The right column shows the information transfer that we observed in the opposite direction. Each row indicates the estimator used. Each box plot features all significant transfer entropy measurements that we found amongst the 30 runs for each maximum lag indicated on the x-axis. Blue box-plots in the top row show the nonlinear results using the KSG estimator, whereas orange box-plots in the bottom row show the linear results using the Gaussian estimator.

Our empirical findings indicate robust results for information transfer from NAHB to SPCS20, as can be seen in the two subfigures in column 1 of Figure 3. Here, the analysis is very

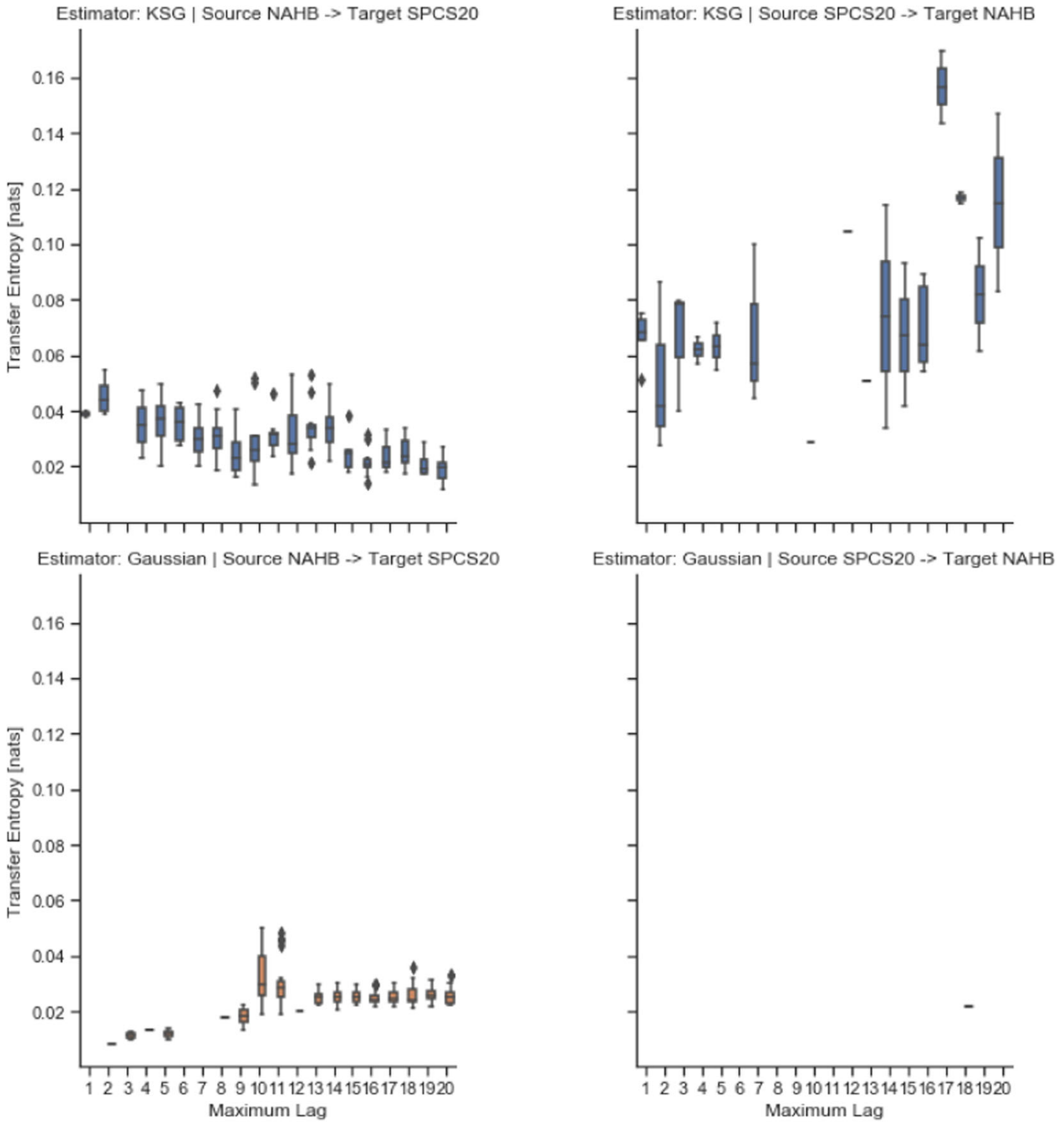

FIGURE 3 Joint transfer entropy results of considering 30 runs of long stationary subsamples of the complete time series. These figures examine the information flow between the variables under investigation here. The top row shows the results using the Kraskov–Stögbauer–Grassberger (KSG) estimator (including nonlinear effects), bottom row shows results using the Gaussian estimator (just including linear effects). Columns indicate the direction of the information flow as stated in the subfigure titles. The *x*-axis shows the maximum lag considered in the measurement. The *y*-axis shows the extent of the measured information transfer over 30 runs

prominent in that both, KSG and Gaussian estimators, converge to similar transfer entropy estimates with higher lags and the box plot bodies are very small in size. This is a plausible outcome for time series with jointly Gaussian distributions. The pattern of this information transfer at different lags for both estimators has also been reproducible with different sets of subsamples repeatedly. Measurements including nonlinear effects seem to differ only slightly from measures considering only linear effects. One interesting observation is that nonlinear information transfer from NAHB to SPCS20 is higher at lower lags than linear effects.

It is also very interesting to see the results of measuring the opposite directional effect—pictured in the two subfigures in column two of Figure 3. Causal relations from SPCS20 to NAHB are only found with nonlinear estimators and seem to be more dependent on the chosen subsample. We observed moderate variation in the median and interquartile range of these KSG boxplots using different sets of start and end combinations, whereas this was not the case for the transfer from NAHB to SPCS20. Hence, the top right subfigure of Figure 3 features larger boxplots and more erratic measurements. Consequently, there could be a problem with the robustness of these results. It is also important to note that only nonlinear causal relations seem to be observable repeatedly when considering information flow from SPCS20 to NAHB. Hence, the bottom right subfigure of Figure 3 shows no prominent linear information transfer. We conclude that robust information flow at a stable extent can be observed with different estimators when considering causal effects from NAHB to SPCS20, whereas the opposite measures are not robust and only nonlinear, where present at all.

As already discussed, this empirical study uses the transfer entropy approach to confirm or invalidate the results from Granger causality tests reported by Rodriguez Gonzalez et al. (2018). Results of a hyperparameter study indicate that causal information flow from NAHB to SPCS20 seems not to be present in a very robust way when only considering lower lags, however, a joint information transfer is very prominent when considering an increasing number of lags. Further analysis showed that the information transfer from SPCS20 to NAHB seems to be nonlinear and fragile depending on the chosen time series subsample. Our hyperparameter study uncovered interesting behaviours regarding the ability to detect linear and nonlinear relations in different directions depending on the settings. We believe that further insights about the nature of the causal relationship can be derived from studying the results of a deeper hyperparameter analysis. The approach reported here uses machine learning workflows that can easily be extended to incorporate multivariate analysis of time series and further hyperparameter settings. We plan to employ the concept presented here to both, broaden and deepen our study to incorporate further time series to condition on, as well as finer granularities of further parameters. To do this, we can transfer the machine learning pipeline to a data centre cluster as it is able to run in typical big data and machine learning environments.

7 | IMPLICATIONS FOR THE EUROPEAN FINANCIAL SERVICES INDUSTRY

It certainly should not be questioned that European financial service firms with a direct exposure to the US real estate market (e.g., institutions owning commercial or residential real estate assets) ought to closely monitor housing prices in North America. Additionally, credit exposure to US real estate assets has to be kept in focus in this context. In fact, European banks played an important role in the subprime crisis because of their holdings of mortgage backed securities which created direct exposure to the US real estate market (see e.g., Hellwig, 2009; Noeth & Sengupta, 2012). Bullard et al. (2009) have stressed that as long as US house prices were rising most mortgage backed securities performed well (which means before the peak of the subprime crisis was reached) because borrowers were usually able to sell real estate assets without suffering losses when they were unable to make loan payments. But this changed with falling house prices. Consequently, investors that held large portfolios of mortgage backed securities all of a sudden had to cope with substantial losses. In this context Mizen (2008) has stressed the fact that some international investors had no experience with US real estate practices. In addition to this important problem there may also have been some difficulties with the risk management processes in many financial services firms back then. Lang and Jagtiani (2010), for example, have stressed the role of financial risk managers to improve our understanding of this crisis. As already noted, the crisis also created a challenge for banking regulation. Cherpack and Jones (2013), for instance, have argued convincingly that the subprime crises in the United States has triggered a strong response by regulators forcing the banks to improve their risk management systems. This change in the regulatory environment was not only observable in the United States. In any case, regulators clearly also played a role in the crisis (see e.g., Swan, 2009; Vazquez & Federico, 2015). After some very costly bank bailouts politicians and regulators wanted financial services firms to become more risk averse. There also have been discussions about possible linkages between the US subprime crisis and the European sovereign debt crisis (see e.g., Ureche-Rangau & Burietz, 2013; Wegener et al., 2019). The latter crisis obviously also has brought about major challenges for risk managers in the European financial services industry. More specifically, the events in the US mortgage market may have increased the level of risk aversion among investors in other countries. Therefore, the subprime crisis in the United States might have raised the awareness of asset managers that there could also be neglected risks buying government bonds issued by less fiscally prudent member countries of the European Monetary Union (e.g., Greece or Portugal). Chang and Leblond (2015), for example, have examined the behaviour of fixed income investors before, during and after the sovereign debt crisis in Europe in some detail. In this context, it has to be stressed that the costly bank rescue programmes also worsened the fiscal difficulties in many European countries and increased the premium for sovereign credit risk these issuers of government bonds had to offer to find investors (see e.g., Basse at al., 2012; Ejsing & Lemke, 2011). In any case, European financial services firms with an exposure to North American real estate assets should adopt appropriate measures to monitor the US property market. An early warning system for house prices in North America could definitely be a central component of such a risk management approach.

But there also is a more macroeconomic dimension. Fleming (1997) as well as Fleming and Remolona (1999), for example, have convincingly argued that announcements of surprising data for US key economic indicators can have strong effects on bond prices and interest rates in North America. Given that the real estate market is of high importance for the US economy (see e.g., Bouchouicha & Ftiti, 2012; Dogan & Topuz, 2020) it certainly does also make sense for risk managers working in financial institutions that have no direct or indirect exposure to the real estate market in North America to monitor US housing prices as soon as they hold some fixed income securities denominated in US dollars. In fact, Bouchouicha and Ftiti (2012) have noted real estate prices are considered to be one of the channels through which monetary policy affects the US economy. Therefore, all financial institutions with an exposure to the US bond market should at least in some way also analyze property prices in North America. Additionally, it has to be stressed that the US bond market is of global relevance. Most importantly, there is clear evidence for a rather strong influence of US interest rates on the behaviour of bond yields in Germany and other countries (see e.g., Bremnes et al., 2001; Monadjemi, 1997). Bremnes et al. (2001), for example, have reported convincing empirical evidence that US interest rates have a significant influence on both German and Norwegian interest rates, whereas the reverse effect at best seems to be modest. Moreover, US equity markets also seem to have a special importance for global share prices (see, amongst others, Gjerde & Sættem, 1995; Syriopoulos, 2007). Syriopoulos (2007), for instance, has noted that there is clear empirical evidence for the very important global role of the US equity market and has also shown that the introduction of the common European currency has not changed this special status. Therefore, even European financial service firms that do not hold US assets could have an incentive to closely monitor the real estate market in North America. Phrased somewhat differently, US house prices seem to matter for the North American bond and equity markets and, therefore, are also potentially relevant for interest rates and stock prices in Europe and other parts of the world. In this context Tiwari et al. (2020) have stressed the need to take into account potential tail events such as the US subprime crisis when analysing spillover-effects between the North American real estate market and financial asset returns. Schwert (2011), for example, has examined the link between financial markets and real economic activity in this crisis in some detail. An extensive literature overview of the causes of the global financial crisis and the European sovereign debt crisis focussing on financial regulation is provided, for example, by Meier et al. (2021). Moreover, Gorton (2009) has analyzed the origins of the crisis. In any case, there are a lot of good reasons for financial services firms (like banks, asset managers or insurance companies) in Europe to closely examine housing prices in the United States — even for those institutions that have no direct exposure to the North American real estate market.

8 | CONCLUSION

The empirical evidence reported above seems to imply that the NAHB housing market index can help to forecast US house prices robustly, even with linear relations. An interesting outcome of our hyperparameter grid search is that only fragile nonlinear relations seem to exist in the opposite direction when considering stationary derivates. The machine learning pipeline presented here offers an easy way to further study these effects in even finer detail on a data centre cluster. Thus, the findings of Rodriguez Gonzalez et al. (2018) are validated and extended using a completely different methodological approach (namely transfer entropy) and were uncovered using a machine learning workflow. Consequently, there is additional empirical evidence for the ability of sentiment indicators to predict real estate prices in the United States. Moreover, the testing procedure used here is not based on the traditional Granger causality approach. This fact renders the results reported above particularly interesting. At this point, it has to be noted that the optimal number of time lags considered in the VAR estimated by Rodriguez Gonzalez et al. (2018) is lower than the optimal number of time lags considered here. In any case, the information provided by the NAHB housing market index certainly can be helpful for financial risk managers building forward-looking early warning system for US house prices. As already noted, this is very important because modern risk management approaches ought to be guided by the idea that the responsible personal in an organisation is not just waiting for bad things to happen. The empirical findings reported above, of course, are of special importance for US financial institutions with a strong exposure to real estate assets in North America. However, the experiences during the recent global financial crisis do show that the results should also matter for banks and other financial services firms that are located in Asia or Europe (see, e.g., Noeth & Sengupta, 2012; Wegener et al., 2019). Moreover, the empirical approach that is used in this paper has also great potential in the field of applied econometrics. In fact, this empirical study shows just one possible application where the concept of entropy can be usefully employed. In particular, this technique can be applied to check the results of Granger causality tests. With regard to real estate economics the concept of entropy could, for instance, be employed to examine the lead–lag relationship between the returns of real estate investment trusts and changes to house prices. This important research question already has been analyzed using Granger causality tests (see e.g., He, 2000; Myer & Webb, 1993). Of course there are also further potential applications in other fields of economics and finance. Future research in the area of energy economics, for example, could focus on the relationship between energy consumption and economic growth (see, amongst others, Belke et al., 2011; Tsani, 2010). With regard to macroeconomics the concept of entropy might, for instance, be used to search for appropriate leading indicators of economic growth (see e.g., Breitung & Candelon, 2006; Huh, 2002). Beyond that, this approach seems to be suitable to search for ripple effects among regional housing prices (see, amongst others, Lee & Chien, 2011; Shi et al., 2009). Furthermore, the concept of entropy could also be helpful in the corporate finance literature to test for dividend signalling or dividend smoothing examining time series data (see, for instance, Basse & Reddemann, 2011; Goddard et al., 2006).

REFERENCES

Anaconda Software Distribution. (2020). Anaconda documentation. *Anaconda Inc*. https://docs.anaconda.com/

Apache Airflow. (2020). Apache Software Foundation. https://airflow.apache.org/

Athey, S. (2018). The impact of machine learning on economics, *The economics of artificial intelligence: An agenda* (pp. 507–547). University of Chicago Press.

Aylward, A., & Glen, J. (2000). Some international evidence on stock prices as leading indicators of economic activity. *Applied Financial Economics*, *10*, 1–14.

Basse, T., Friedrich, M., & Kleffner, A. (2012). Italian government debt and sovereign credit risk: An empirical exploration and some thoughts about consequences for European insurers. *Zeitschrift für die gesamte Versicherungswissenschaft*, *101*, 571–579.

Basse, T., & Reddemann, S. (2011). Inflation and the dividend policy of US firms. *Managerial Finance*, *37*, 34–46.

Basuchoudhary, A., Bang, J. T., & Sen, T. (2017). *Machine-learning techniques in economics: New tools for predicting economic growth*. Springer.

Behrendt S., Prange P. (2021). What are you searching for? On the equivalence of proxies for online investor attention. *Finance Research Letters*, *38*. http://doi.org/10.1016/j.frl.2019.101401

Belke, A., Dobnik, F., & Dreger, C. (2011). Energy consumption and economic growth: New insights into the cointegration relationship. *Energy Economics*, *33*, 782–789.

Blake, J. P., & Gharleghi, B. (2018). The ripple effect at an inter-suburban level in the Sydney metropolitan area. *International Journal of Housing Markets and Analysis*, *11*, 2–33.

Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016a). Information flow in complex systems, *An introduction to transfer entropy* (pp. 1–10). Springer.

Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016b). Information flow in complex systems, *An introduction to transfer entropy* (pp. 65–95). Springer.

Bouchouicha, R., & Ftiti, Z. (2012). Real estate markets and the macroeconomy: A dynamic coherence framework. *Economic Modelling*, *29*, 1820–1829.

Breden, D. (2008). Monitoring the operational risk environment effectively. *Journal of Risk Management in Financial Institutions*, *1*, 156–164.

Breitung, J., & Candelon, B. (2006). Testing for short-and long-run causality: A frequency-domain approach. *Journal of Econometrics*, *132*, 363–378.

Bremnes, H., Gjerde, Ø., & Sættem, F. (2001). Linkages among interest rates in the United States, Germany and Norway. *Scandinavian Journal of Economics*, *103*, 127–145.

Bullard, J., Neely, C. J., & Wheelock, D. C. (2009). Systemic risk and the financial crisis: A primer. *Federal Reserve Bank of St. Louis Review*, *91*, 403–418.

Bzdok, D., Krzywinski, M., & Altman, N. (2017). Machine learning: A primer. *Nature Methods*, *14*, 1119–1120. https://doi.org/10.1038/nmeth.4526

Cass, S. (2018). The 2018 top programming languages—Python stays on top and Assembly enters the Top Ten. *IEEE Spectrum*. https://spectrum.ieee.org/at-work/innovation/the-2018-top-programming-languages

Chang, M., & Leblond, P. (2015). All in: Market expectations of eurozone integrity in the sovereign debt crisis. *Review of International Political Economy*, *22*, 626–655.

Cheng, H. (1979). Causality tests in econometrics. *Journal of Economic Dynamics and Control*, *1*, 321–346.

Cherpack, P. L., & Jones, B. W. (2013). Community bank proactive risk management: Concentration management, stress testing and capital planning. *Journal of Risk Management in Financial Institutions*, *6*, 411–432.

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*, 5–37.

Cooray, A., & Wickremasinghe, G. (2007). The efficiency of emerging stock markets: Empirical evidence from the South Asian region. *Journal Of Developing Areas*, *41*, 171–183.

Croce, R. M., & Haurin, D. R. (2009). Predicting turning points in the housing market. *Journal of Housing Economics*, *18*, 281–293.

Dietzel, M. A., Braun, N., & Schäfers, W. (2014). Sentiment-based commercial real estate forecasting with Google search volume data. *Journal of Property Investment and Finance*, *32*, 540–569.

Dimpfl, T., & Peter, F. J. (2013). Using transfer entropy to measure information flows between financial markets. *Studies in Nonlinear Dynamics and Econometrics*, *17*(1). http://doi.org/10.1515/snde-2012-0044

Dimpfl, T., & Peter, F. J. (2014). The impact of the financial crisis on transatlantic information flows: An intraday analysis. *Journal of International Financial Markets, Institutions and Money*, *31*, 1–13.

Dogan, C., & Topuz, J. C. (2020). Real effects of real estate: Evidence from unemployment rates. *Studies in Economics and Finance*, *37*, 605–623.

Eichengreen, B., Mody, A., Nedeljkovic, M., & Sarno, L. (2012). How the subprime crisis went global: Evidence from bank credit default swap spreads. *Journal of International Money and Finance*, *31*, 1299–1318.

Ejsing, J., & Lemke, W. (2011). The Janus-headed salvation: Sovereign and bank credit risk premia during 2008–2009. *Economics Letters*, *110*, 28–31.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64*, 813–836.

Fleming, M. J. (1997). The round-the-clock market for US Treasury securities. *Economic Policy Review*, *3*, 9–32.

Fleming, M. J., & Remolona, E. M. (1999). What moves bond prices? *Journal of Portfolio Management*, *25*, 28–38.

Geweke, J. (1984). Chapter 19. Inference and causality in economic time series models, *Handbook of econometrics* (Vol. 2, pp. 1101–1144). Elsevier.

Gjerde, Ø., & Sættem, F. (1995). Linkages among European and world stock markets. *European Journal of Finance*, *1*, 165–179.

Goddard, J., McMillan, D. G., & Wilson, J. O. (2006). Dividend smoothing vs dividend signalling: evidence from UK firms. *Managerial Finance*, *32*, 493–504.

Goodman, J. (1994). Using attitude data to forecast housing activity. *Journal of Real Estate Research*, *9*, 445–453.

Gorton, G. (2009). The subprime panic. *European Financial Management*, *15*, 10–46.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438.

Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of Econometrics*, *39*, 199–211.

Green, R. K. (2002). Stock prices and house prices in California: New evidence of a wealth effect? *Regional Science and Urban Economics*, *32*, 775–783.

Gómez-Herrero, G., Wu, W., Rutanen, K., Soriano, M. C., Pipa, G., & Vicente, R. (2015). Assessing coupling dynamics from an ensemble of time series. *Entropy*, *17*, 1958–1970.

Hausler, J., Ruscheinsky, J., & Lang, M. (2018). News-based sentiment analysis in real estate: A machine learning approach. *Journal of Property Research*, *35*, 344–371.

He, L. (2000). Causal relationships between apartment REIT stock returns and unsecuritized residential real estate. *Journal of Real Estate Portfolio Management*, *6*, 365–372.

Hellwig, M. F. (2009). Systemic risk in the financial sector: An analysis of the subprime-mortgage financial crisis. *De Economist*, *157*, 129–207.

Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance*, *49*, 1639–1664.

Hinton, G. E., Krizhevsky, A., & Sutskever, I. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*, 1106–1114.

Huang, M. (2019). Markov-switching impacts of housing-market expectations on credit markets. *Managerial Finance*, *46*(3), 381–400. http://doi.org/10.1108/mf-08-2019-0391

Huh, H. S. (2002). GDP growth and the composite leading index: A nonlinear causality analysis for eleven countries. *Economics Letters*, *77*, 93–99.

Jorion, P. (2009). Risk management lessons from the credit crisis. *European Financial Management*, *15*, 923–933.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, *69*, 066138.

Kunze, F., Basse, T., Rodriguez Gonzalez, R., & Vornholz, G. (2020). Forward-looking financial risk management and the housing market in the United Kingdom: Is there a role for sentiment indicators? *Journal of Risk Finance*, *21*, 659–678.

Lang, W. W., & Jagtiani, J. A. (2010). The mortgage and financial crises: The role of credit risk management and corporate governance. *Atlantic Economic Journal*, *38*, 123–144.

Lee, C. C., & Chien, M. S. (2011). Empirical modelling of regional house prices and the ripple effect. *Urban Studies*, *48*, 2029–2047.

Liu, S., Sathye, M., Chang, C. L., Jiménez-Martín, J. Á., McAleer, M., & Pérez-Amaral, T. (2011). Risk management of risk under the Basel Accord: Forecasting value-at-risk of VIX futures. *Managerial Finance*, *37*, 1088–1106.

Lizier, J. T. (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI*, *1*, 11.

Luo, Z., Liu, C., & Picken, D. (2007). Granger causality among house price and macroeconomic variables in Victoria. *Pacific Rim Property Research Journal*, *13*, 234–256.

Luu Duc Huynh, T. (2019). Spillover risks on cryptocurrency markets: A look from VAR-SVAR Granger causality and Student's t copulas. *Journal of Risk and Financial Management*, *12*, 52.

López de Prado, M. (2019). Beyond econometrics: A roadmap towards financial machine learning. https://ssrn.com/abstract=3365282

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

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

Marschinski, R., & Kantz, H. (2002). Analysing the information flow between financial time series. *European Physical Journal B*, *30*, 275–281.

McLaren, N., & Shanbhogue, R. (2011). Using Internet search data as economic indicators. *Bank of England Quarterly Bulletin*, *51*, 134–140.

Meier, S., Rodriguez Gonzalez, M., & Kunze, F. (2021). The global financial crisis, the EMU sovereign debt crisis and international financial regulation: Lessons from a systematic literature review. *International Review of Law and Economics*, *65*, 105945. https://doi.org/10.1016/j.irle.2020.105945

Mizen, P. (2008). The credit crunch of 2007-2008: A discussion of the background, market reactions, and policy responses. *Federal Reserve Bank of St. Louis Review*, *90*, 531–567.

Mjolsness, E., & Decoste, D. (2001). Machine learning for science: State of the art and future prospects. *Science (New York, N.Y.)*, *293*, 2051–2055. https://doi.org/10.1126/science.293.5537.2051

Monadjemi, M. S. (1997). International interest rates linkages: evidence from OECD countries. *International Review of Financial Analysis*, *6*, 229–240.

Myer, F. N., & Webb, J. (1993). Return properties of equity REITs, common stocks, and commercial real estate: A comparison. *Journal of Real Estate Research*, *8*, 87–106.

Noeth, B., & Sengupta, R. (2012). Global European banks and the financial crisis. *Federal Reserve Bank of St. Louis Review*, *94*, 457–479.

Osiobe, E. U. (2020). Human capital, capital stock formation, and economic growth: A panel Granger causality analysis. *Journal of Economics and Business*, *3*, 569–582.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramirez, S. R. (2019). An examination of firm licensing behaviour after a payday-loan ban. *Applied Economics*, *51*, 5090–5103.

Rochette, M. (2009). From risk management to ERM. *Journal of Risk Management in Financial Institutions*, *2*, 394–408.

Rodriguez Gonzalez, M., Basse, T., Kunze, F., & Vornholz, G. (2018). Early warning indicator systems for real estate investments: Empirical evidence and some thoughts from the perspective of financial risk management. *Zeitschrift für die gesamte Versicherungswissenschaft*, *107*, 387–403.

van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*, 075310.

Rényi, A. (1970). *Probability theory*. North-Holland.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*, 461–464.

Schwert, G. W. (2011). Stock volatility during the recent financial crisis. *European Financial Management*, *17*, 789–805.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423.

Shi, S., Young, M., & Hargreaves, B. (2009). The ripple effect of local house price movements in New Zealand. *Journal of Property Research*, *26*, 1–24.

Simon, B., Thomas, D., Franziska, J. P., & David, J. Z. (2019). RTransferEntropy—Quantifying information flow between different time series using effective transfer entropy. *SoftwareX*, *10*(100265), 1–9.

Swan, P. L. (2009). The political economy of the subprime crisis: Why subprime was so attractive to its creators. *European Journal of Political Economy*, *25*, 124–132.

Syczewska, E. M., & Struzik, Z. R. (2014). Granger causality and transfer entropy for financial returns. *Acta Physica Polonica A*, *127*(3–A), 129–135.

Syriopoulos, T. (2007). Dynamic linkages between emerging European and developed stock markets: Has the EMU any impact? *International Review of Financial Analysis*, *16*, 41–60.

Teye, A. L., Knoppel, M., de Haan, J., & Elsinga, M. G. (2017). Amsterdam house price ripple effects in The Netherlands. *Journal of European Real Estate Research*, *10*, 331–345.

Tiwari, A. K., André, C., & Gupta, R. (2020). Spillovers between US real estate and financial assets in time and frequency domains. *Journal of Property Investment and Finance*, *38*, 525–537.

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*, 225–250.

Tsani, S. Z. (2010). Energy consumption and economic growth: A causality analysis for Greece. *Energy Economics*, *32*, 582–590.

Tsolacos, S. (2012). The role of sentiment indicators for real estate market forecasting. *Journal of European Real Estate Research*, *5*, 109–120.

Ureche-Rangau, L., & Burietz, A. (2013). One crisis, two crises… the subprime crisis and the European sovereign debt problems. *Economic Modelling*, *35*, 35–44.

Vazquez, F., & Federico, P. (2015). Bank funding structures and risk: Evidence from the global financial crisis. *Journal of Banking and Finance*, *61*, 1–14.

Vicente, R., Wibral, M., & Lindner, M. (2011). Transfer entropy—A model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, *30*, 45–67.

Wegener, C., Kruse, R., & Basse, T. (2019). The walking debt crisis. *Journal of Economic Behavior and Organization*, *157*, 382–402.

Wilcox, J. A. (2015). The home purchase sentiment index: A new housing indicator. *Business Economics*, *50*, 178–190.

Wollstadt, P., Lizier, J., Vicente, R., Finn, C., Martinez-Zarzuela, M., Mediano, P., Novelli, L., & Wibral, M. (2019). IDTxl: The Information Dynamics Toolkit xl: a Python package for the efficient analysis of multivariate information dynamics in networks. *Journal of Open Source Software*, *4*(34) http://doi.org/10.21105/joss.01081

Wood, R. (2003). The information content of regional house prices: can they be used to improve national house price forecasts? *Bank of England Quarterly Bulletin*, *43*, 304–314.

Xie, Z., & Chen, S. W. (2014). Untangling the causal relationship between government budget and current account deficits in OECD countries: Evidence from bootstrap panel Granger causality. *International Review of Economics and Finance*, *31*, 95–104.

Zapata, H. O., & Rambaldi, A. N. (1997). Monte Carlo evidence on cointegration and causation. *Oxford Bulletin of Economics and Statistics*, *59*, 285–298.

How to cite this article: Basse, T., Kunze, F., Rodriguez Gonzalez, M., & Saft, D. (2022). Leading indicators for US house prices: New evidence and implications for EU financial risk managers. *European Financial Management*, 28, 722–743. https://doi.org/10.1111/eufm.12325