1. BUBBLE OR NOT? A TIME SERIES CLASSIFICATION PROBLEM USING MACHINE LEARNING

Radu LUPU[1, 2]
Adrian Cantemir CĂLIN[1, 3*]
Iulia LUPU[3]
Laura Andreea IANCU[1, 2]
Andreea Elena CROICU[1, 2]

Abstract

Anticipating the occurrence of financial bubbles holds paramount significance as it empowers investors to make judicious decisions and navigate potential losses adeptly. Additionally, the prediction and identification of bubbles play a pivotal role in achieving financial stability objectives. In light of these considerations, this research paper endeavors to address the challenge of predicting financial bubbles through a methodology that integrates the BSADF test with machine learning algorithms. The initial phase involves the identification of bubbles within the stock prices of all entities comprising the STOXX 600 index, followed by the application of a machine learning framework to forecast bubble values. The study aims to discern and incorporate all relevant features for the prediction of bubbles, employing a diverse array of neural network algorithms to formulate forecasts. Subsequently, the research evaluates the out-of-sample prediction accuracy of these algorithms.

Keywords: financial bubbles forecasting, BSADF, MLP, NBEATS, NHITS

JEL Classification: G12, G17

Introduction

The contemplation of bubbles within stock or financial markets has garnered significant attention and discourse among economists, investors, and analysts. Bubbles, denoted by a rapid surge in asset valuations succeeded by an abrupt downturn, possess the potential to exert profound influences on the economy and individual investment portfolios. The Global Financial Crisis of 2007-2008 and the ensuing Great Recession, often ascribed to the rupture of housing bubbles in several nations, have demonstrated the deleterious consequences that a decline in asset prices can inflict upon the actual economy (Gali et al, 2021).

[1] *Bucharest University of Economic Studies, Bucharest (Romania)*

[2] *Institute for Economic Forecasting, Romanian Academy, Bucharest (Romania)*

[3] *"Victor Slăvescu" Centre for Financial and Monetary Research, Romanian Academy, Bucharest (Romania)*

[*] *Corresponding author: adrian.calin@rei.ase.ro*

Historically, recurrent instances of bubbles have frequently manifested during epochs of heightened productivity, as exemplified by the railway boom, electricity boom, chemistry boom, and the more contemporary internet and telecommunications boom (Abreu and Brunnermeier, 2003). Throughout these intervals, there is a surge in profitability within certain industries, ultimately manifested in a rising trajectory of dividends. Investors' temporal perspectives contract, leading to investment strategies increasingly predicated on pursuing short-term capital gains. Consequently, these dynamics diminish the relevance of fundamentals in determining stock prices (Cerruti and Lombardini, 2022)). Notwithstanding the recurrent manifestation of this phenomenon, a universally acknowledged theory elucidating the genesis of bubbles remains absent from both theoretical and empirical literature. Nevertheless, consensus exists regarding the characterization of an asset bubble as a departure of the asset's market value from its fundamental value. Throughout economic history, there have been recurring instances of profound financial crises coinciding with the precipitous decline in asset prices within contemporary monetary and financial systems (Hashimoto et al, 2020). In theoretical terms, a bubble is deemed undesirable as it undermines both economic stability and efficiency (Wan, 2024). Such phenomena have adverse effects on the balance sheets of corporations, financial institutions (Sakuragawa, 2021), and households (Wan, 2021).

Predicting bubbles is crucial as it enables investors to make well-informed choices and sidestep probable losses. Early identification of a bubble permits investors to offload their assets before its collapse, thereby protecting their investments from significant price drops and reducing losses. However, forecasting bubbles in financial or stock markets is a complex task. It demands an extensive comprehension of market trends, economic indicators, and human psychology. Experts often point out the difficulty in predicting bubbles due to their basis in emotional responses and group dynamics, as opposed to logical decision-making. Furthermore, recent scholarly literature indicates that bubbles associated with tangible assets are inherently nonstationary phenomena linked to imbalanced growth. This suggests that comprehending the nature of asset price bubbles requires a departure from stationary models with a stable state to nonstationary models lacking such equilibrium.

Technical analysis is a prevalent method for bubble prediction. It examines past price movements and market behaviors to spot potential bubbles. For instance, a sharp and rapid price surge, known as a parabolic rise, might indicate an impending bubble. Fundamental analysis is another technique, focusing on evaluating a company's or market's financial stability and performance. This helps in spotting assets or markets that are overpriced due to speculation rather than genuine economic progress.

Additionally, macroeconomic factors like interest rates, inflation, and consumer spending are utilized by some experts to forecast bubbles. A significant rise in interest rates, for instance, could lead to falling asset prices, signaling a potential bubble. Despite these varied approaches, it's crucial to acknowledge the limitations of bubble prediction. Bubbles can persist longer than anticipated, and pinpointing their exact burst timing is almost impossible. Hence, portfolio diversification and not relying exclusively on bubble predictions is vital for investors.

Notwithstanding these limitations, and in light of the significance of the subject matter, a substantial body of literature has been devoted to the scrutiny of diverse inquiries concerning financial bubbles across various financial assets, with a primary focus on bubble identification. Regarding the identification of bubbles, Phillips et al. (2011, 2015a, 2015b) have proposed the employment of the SADF method. Subsequently, they extended their investigation to encompass the generalized sup Augmented Dickey-Fuller and BSADF test methods, building upon the foundational SADF methodology. This approach not only facilitates the detection of periodic explosive bubbles but also enables the identification of the inception point of such bubbles.

Driven by the pertinence of bubbles to financial stability and the typical dynamics of the broader economy, our endeavor seeks to address the void in the realm of bubble predictions through the following contributions.

Initially, we identify bubbles in the stock prices of all corporations constituting the STOXX 600 index during the period spanning January 2015 to October 2023. Our approach is grounded in the methodological framework established by Brunnermeier et al. (2020), and we employ the methodology proposed by Phillips, Shi, and Yu (2015a,b) to estimate episodes of bubbles. This involves the application of the backward sup augmented Dickey-Fuller (BSADF) approach, allowing us to discern the temporal extent of the bubble, delineate its expansionary phase, and pinpoint its subsequent contraction phase.

Secondly, we implemented the aforementioned methodology for all entities within the STOXX 600, scrutinizing their statistical characteristics. Across all companies, we successfully identified 3351 instances of bubbles during the specified period. Subsequently, detailed descriptive statistics for these identified bubbles are presented in the subsequent section. It is observed that a majority of companies experienced approximately six instances of bubbles, a figure consistent with both the mean and the median. Given our focus on predicting bubbles, we refined the dataset by filtering companies that exhibited precisely six bubbles. Subsequently, we developed our predictive algorithm specifically tailored to these instances. This selection process resulted in a database comprising 93 companies.

We employ a Machine Learning framework to generate forecasts for these variables. Our methodology involves a two-step process: ***1.*** Identifying all potential features deemed pertinent for predicting bubbles; ***2.*** Employing a suite of neural network algorithms that leverage these to formulate forecasts, subsequently assessing their out-of-sample prediction accuracy.

Literature review

The examination and quantification of extreme financial perils, such as the threat of market bubbles, have gained escalating significance within financial markets. Bubbles denote instances of swift escalation in asset valuations succeeded by precipitous declines, culminating in substantial monetary setbacks for investors and the potential destabilization of the market. It is imperative for market participants to grasp the concept of bubble risk to facilitate judicious investment decisions and to mitigate exposure to potential losses.

The correlation among bubbles across diverse financial assets, recognized as bubble risk transmission, is influenced by common factors such as information asymmetry. However, current research has overlooked two critical dimensions concerning the correlation of risks. Firstly, there is a dearth of exploration into bubble risk and its transmission within complex systems. Comprehending risk transmission necessitates a holistic perspective. The impact of a single asset's bubble risk can either be magnified or offset when assessed at the macroeconomic scale of an entire nation, posing challenges for singular asset risk analysis. Secondly, the prevailing employment of linear models to scrutinize risk impacts is inadequate. Bubbles themselves exhibit traits of sharp peak-fat tails, and their emergence is inherently unpredictable, rendering linear methodologies unsuitable for depicting the transmission processes effectively.

Bubbles represent a phenomenon characterized by economic distortions resulting in irrational fluctuations in asset valuations, which can profoundly impact market functionality. For example, Tokic (2010) dedicates considerable attention to the bubbles present during the 2008 crisis, with particular emphasis on the 2008 Oil Bubble. Various definitions of bubbles exist within scholarly discourse. West (1987) conceptualizes a bubble as the segment of asset value that strays from market fundamentals, either surpassing or falling short of the asset's intrinsic worth. The model undergoes empirical validation through the utilization of the dataset provided in Shiller's (1981).

Garber (1990) similarly posits that bubbles encompass price fluctuations that lack explanation by underlying economic fundamentals. Shiller (2000) argues that bubbles manifest due to irrational exuberance, whereby the escalating asset prices, propelled by investor trading activities, create a feedback loop amplifying prices further from fundamental values.

Regarding the etiology of bubbles, De Long et al. (1990) advance a noise trading model positing that irrational investment behaviors of noise traders, influenced by market noise, lead to deviations of asset prices from fundamental values, thereby engendering bubbles. Hott (2009) presents a model elucidating the emergence of price bubbles through the herding effect, offering a comprehensive examination of the bubble formation process.

Two predominant approaches for identifying bubble existence are direct and indirect methodologies. Direct testing methods typically utilize the Markov regime transition model, incorporating the West two-step detection method proposed by West (1987) and the Markov regime transition model (MRS). These methodologies presuppose that the dynamic adjustment process of bubbles mirrors the dynamic evolution of diverse states. Conversely, indirect detection methods discern bubbles based on price distributional characteristics. Representative methodologies encompass the variance boundary test advocated by Shiller (1981), the structural mutation point test proposed by Homm and Breitung (2012), and the unit root and cointegration test articulated by Diba and Grossman (1988). Nonetheless, these methodologies exhibit limitations. For instance, the variance boundary detection method overlooks the exponential growth of rational bubbles, the structural mutation test method relies on subjective elements, and the unit root and cointegration test methods encounter challenges in identifying periodic bubbles.

In response to these limitations, Phillips et al. (2015) introduce the supremum ADF test and the backward sup ADF test, which have gained widespread adoption as bubble detection methodologies. Chen and Xie (2017) employ the GSADF method to examine stock market bubbles across 10 European nations and 9 pan-Pacific countries. Furthermore, they utilize two alternative measures, namely MTAR and ESTAR, and uncover substantial evidence supporting the presence of periodically collapsing bubbles across multiple countries within their sample.

A body of scholarly literature, as evidenced by Khan, Su, Umar, and Yue (2021), Khan, Su, and Rehman (2021), Wang and Da Gao (2022), Khan et al. (2022), among others, examines bubble phenomena across multiple assets including coal prices, energy prices, and crude oil. These investigations are motivated by distinct political and financial market dynamics, employing the GSADF methodology. The prevailing consensus within this literature tends towards the recognition of bubbles within the aforementioned commodities. Li et al. (2020) employ GSADF to investigate the inception and zenith values of bubbles in major natural gas markets across Europe, Asia, and the Americas, alongside conducting metric analyses of these bubbles. The paper makes four primary contributions. Firstly, it provides documentation of the occurrences of bubble episodes within the examined regions. Secondly, the author observes that bubbles manifest simultaneously across all three markets, attributable to rapid and pronounced price fluctuations. Thirdly, discernible heterogeneity arises concerning the occurrence and duration of bubbles across the three markets, as mentioned earlier. Bubbles within the European Union (EU) context tend to persist for longer durations. Lastly, it is noted that global economic events can exert a concurrent influence.

Additionally, Pavlidis et al. (2019), Yao and Li (2021), among others, apply this category of bubble detection methodology across various domains. The initial investigation underscores the observation that aggregation diminishes the effectiveness of both the SADF and GSADF tests, whereas the subsequent study also employs a LPPLS sequence.

Numerous occurrences of price bubbles have been documented in both crude oil and natural gas markets (Gronwald, 2016; Caspi et al., 2018; Sharma and Escobari, 2018; Zhang et al., 2018, Akcora, and Kocaaslan 2023, Chang 2024). In the case of crude oil markets, the principal factors contributing to these price bubbles encompass global economic expansion, supply-demand imbalances, depreciation of the US dollar, heightened shale oil production in the US, and burgeoning demand in emerging economies (Khan et al., 2021). Conversely, the origins of price bubbles in the natural gas markets of the European Union (EU), Japan, and the United States (US) differ. The EU is primarily influenced by political factors, Japan grapples with supply-demand disparities, and the US contends with short-term financial market volatility (Li et al., 2020).

Caraiani and Călin (2019) conducted an investigation into the repercussions of monetary policy shocks, including unconventional measures, on energy sector bubbles in the United States. Employing a time-varying Bayesian VAR model, they quantified the effects of monetary policy shocks on asset prices and bubbles. The study unveiled noteworthy disparities in the impact of monetary policy shocks on the broader economy compared to the energy sector.

Teti and Maroni (2021) evaluated the emergence of contemporary bubbles in the technology industry, positing that market multiples contribute to a partially positive outlook. They contend that the current valuation appears relatively rational in comparison to the dot-com bubble of 2000, notwithstanding the price-earnings ratio persisting above the market average.

Geuder, Kinateder, and Wagner (2019) directed their focus towards cryptocurrency bubbles, emphasizing the recurrent nature of bubble behaviors in Bitcoin prices. They discerned no ongoing evidence of bubble behaviors post December 6, 2017. Kyriazis, Papadamou, and Corbet (2020) unearthed that Bitcoin had entered a bubble phase since June 2015, whereas other Cryptocurrencies exhibited bubble features from September 2015. However, since early 2018, there has been scant academic substantiation supporting the presence of bubbles in this latter group. Furthermore, Chowdhury, Damianov, and Elsayed (2022) scrutinized the epochs of bubble emergence and collapse, revealing interdependence and contagion effects. They observed Bitcoin acting as a net transmitter during market downturns and vice versa.

Housing bubbles have emerged as a significant subject of inquiry within the housing market, particularly in the aftermath of the subprime mortgage crisis of 2007. Numerous scholarly investigations have scrutinized the housing bubble and subsequent collapse in the United States (Goodman and Thibodeau (2008); Kivedal (2013); Shi (2017) or Evgenidis and Malliaris (2023)). Despite the considerable attention devoted to the performance of housing markets in the Eurozone, there exists a relative paucity of studies delineating housing bubble-like phenomena within this geographical region.

Recent scholarly endeavors have directed attention towards the interrelated movements of housing bubbles in response to monetary policy interventions. Dynamic factor models have been employed to probe into the global and nation-specific determinants driving these co-movements, alongside their reactions to monetary policy shocks. For instance, Caraiani and Călin (2019) discerned such patterns across a subset of OECD housing markets.

Several scholarly inquiries have investigated the determinants of housing prices in the Europe (Bago et al., 2021; Martin et al., 2021, Tsai and Lin, 2022), alongside exploring housing price correlation and contagion (Tsai, 2018, Tsai, 2019). Nonetheless, only a limited number of studies have discerned housing bubbles within specific European regions. For example, Ott (2014) scrutinized the long-term equilibrium and short-term dynamics of housing markets in the Eurozone, identifying a protracted boom period from 1997 to 2007, succeeded by a corrective phase leading to a reversion to equilibrium by 2014. Zhu et al. (2017) delved into the impact of monetary policy and mortgage market structure on non-fundamental house price trends across 11 Eurozone nations, revealing that a singular monetary easing shock exerted a notable influence on house price escalations, notably in Ireland and Spain. Moons and Hellinckx (2019) examined the European Central Bank's interest rate policy's effects on the Irish financial-economic crisis, revealing that the ECB's monetary policy stipulations could forestall significant house price hikes, with average house prices experiencing a 25-30% downturn following the Irish housing bubble.

André et al. (2022) employed a BVAR framework to scrutinize the repercussions of monetary policy shocks on housing prices in the United States, the United Kingdom, and Canada. Additionally, they investigated the role of housing bubbles in shaping housing price responses to monetary policy tightening. Their findings indicated a negative reaction of housing prices to monetary tightening, implying that monetary policy might aid in mitigating housing price bubbles.

Data and Methodology

As delineated earlier, our objective is to identify bubbles in the stock prices of all corporations comprising the STOXX 600 index within the timeframe of January 2015 to October 2023, utilizing the (BSADF) approach.

We successfully isolate 3351 instances of bubbles during this period across all companies, and Table 1 provides the descriptive statistics pertaining to these identified bubbles.

Table 1. Descriptive statistics[4]

	Start	Peak	End	Duration	Boom Phase	Burst Phase
count	3351	3351	3351	3351	3351	3351
mean	10/31/2019	12/18/2019	1/14/2020	53.84632	33.94599	19.90033
std	551.661691	550.72659	552.714931	80.82855	67.85041	28.8328
min	6/3/2015	6/18/2015	6/25/2015	7	1	1
25%	2/1/2018	4/6/2018	4/27/2018	12	3	5
50%	1/8/2020	2/18/2020	3/9/2020	24	9	10
75%	7/16/2021	10/19/2021	11/23/2021	58	31	22
max	10/5/2023	10/16/2023	10/16/2023	771	724	321

Source: Authors' computation

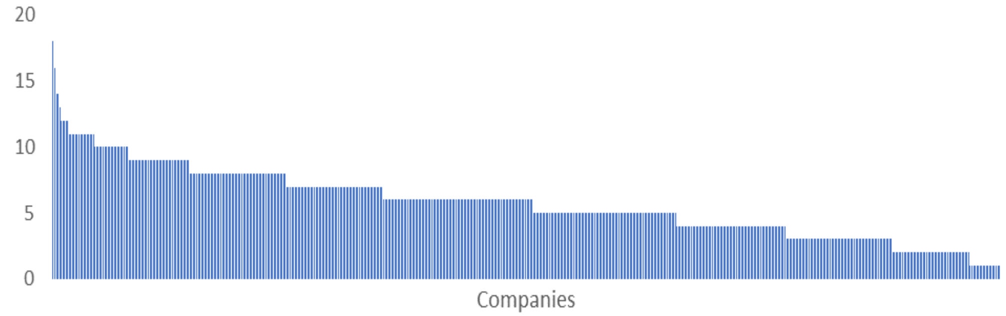
We notice the distribution of bubbles and their peak and end moments across all assets as observed in time. In this table the mean of 10/31/2019 for the start of the bubbles show that on average the bubbles across all assets tend to fluctuate around this moment. Same interpretation is valid for the peak and the end.

The preliminary outcomes of our bubble identification procedure are synthetically depicted in Figure 1, illustrating the frequency distribution of bubbles per company. Additionally, we offer a histogram portraying the distribution of the number of bubbles across all assets in our sample (Figure 2), accompanied by the statistical properties of the number of bubbles across assets as presented in Table 2.

We notice that the larger number of companies had about 6 bubbles. This is equal to the mean and the median. Our objective is to predict the bubbles, so we filter the data to select only the companies that prompted six bubbles and design our prediction algorithm for these situations. This selection procedure yielded 93 companies, which will represent our database.

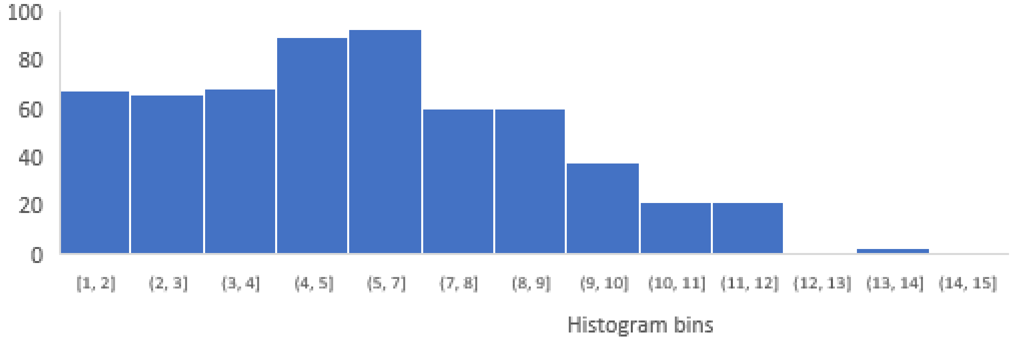
[4] *Start, Peak and End show the time when bubbles started, peaked and respectively ended. The others show the number of days for each category.*

Figure 1. Number of bubbles per company



Source: Authors' computation

Figure 2. Histogram of the number of bubbles across all assets in our sample



Source: Authors' computation

Table 2. Statistical properties of the number of bubbles across assets

mean	std	min	25%	50%	75%	max
5.69898	2.650486	1	4	6	7	18

Source: Authors' computation

Furthermore, our attention is directed towards discerning the duration of bubbles, the expansionary phase (boom), and the contraction phase (burst). Each of these facets is represented as Boolean variables, denoted by 0 for regular moments and 1 when bubbles are identified. A sample of these variables is provided below (Table 3).

Table 3. An example of values for the Boolean variables showing bubbles for PSP Swiss Property AG

	Bubble - all period	Boom phase	Burst phase
01/03/2019	0	0	0
04/03/2019	0	0	0
05/03/2019	0	0	0
06/03/2019	0	0	0
07/03/2019	0	0	0
08/03/2019	0	0	0
11/03/2019	0	0	0
12/03/2019	0	0	0
13/03/2019	0	0	0
14/03/2019	1	1	0
15/03/2019	1	1	0
18/03/2019	1	1	0
19/03/2019	1	1	0
20/03/2019	1	1	0
21/03/2019	1	1	0
22/03/2019	1	1	0
25/03/2019	1	1	0
26/03/2019	1	1	0
27/03/2019	1	1	0
28/03/2019	1	1	0
29/03/2019	1	0	1
01/04/2019	1	0	1
02/04/2019	1	0	1
03/04/2019	1	0	1
04/04/2019	0	0	0
05/04/2019	0	0	0

Source: Authors' computation

We are using a Machine Learning framework to create forecasts for these variables. Our approach relies on a two-step procedure: 1. Identify all possible features that could be useful for prediction of bubbles; 2. Use a set of neural network algorithms that make use of these features to create forecasts and analyze their prediction accuracy out-of-sample.

Feature extraction and selection

We employ the TSFRESH[5] Python library, which is designed to excel in the extraction of statistical features from time series data. This versatile toolkit offers numerous capabilities for calculating a comprehensive range of descriptive statistics[6], encompassing both the entire dataset and dynamic time series windows. As showcased in Christ et. al (2018) the toolbox provides 63 time series characterization methods, which are employed to estimate 794 features.

Since our aim is to create forecasts for the Boolean variables, we create rolling windows of 100 observations for each company and generate the statistical feature for each window based on the price dynamics. In the initial setup, this yielded 794 new explanatory variables each of them characterizing the rolling windows, for each company in our set of 93 assets selected for the analysis.

To increase the forecast accuracy, we used Christ et al. (2017) to filter the explanatory variables and select only the relevant ones for each Boolean variable corresponding to each company. The authors explain that the algorithm follows the following steps:

The Feature Extraction based on Scalable Hypotesis tests (FRESH) for a given parameter $q \in [0,1]$ uses the following procedure:

1. We perform a set of n_ϕ univariate feature mappings on $m \cdot n$ time series to produce the feature vectors X_ϕ with$\phi = 1, \dots, n_\phi$.
2. For each feature vector formed, $X_1, \dots, X_{n_\phi}$ we conduct a H_0^ϕ hypothesis test. In this step we employ the corresponding feature significance test and calculate the p-values $p_1, \dots, p_{n_\phi}$.
3. Finally, we apply the Benjamini-Yekutieli procedure to correct for dependent hypotheses for a FDR level of 1 on all the collected p-values $p_1, \dots, p_{n_\phi}$ in order to determine which null hypothesis needs to be rejected.

We collect all relevant features as time series corresponding to each company and attach the values for each of the Boolean variables. Table A1 in the appendix shows an excerpt of these features extracted for Aeroports de Paris SA.

Prediction algorithms

Given the large number of features and the interest to obtain a good accuracy of prediction, we decided to use neural network algorithms designed as classifiers. The models are Multilayer Perceptron (MLP) as in Zhang (2003), the Neural basis expansion analysis for interpretable time series (NBEATS) as in Oreshkin et al (2019) and the Neural hierarchical interpolation for time series (NHITS) as in Challu et al (2023).

The predictions provided by these models will be shaped as Boolean variables, hence we designed them as classifiers by specifying a binary classification loss function shaped as a likelihood function that assumes a Bernoulli distribution.

MLP is a simple neural network setup and we use it just for comparison with the more advanced models. A noteworthy advantage inherent to Artificial Neural Network (ANN) models, as opposed to other categories of nonlinear models, is their status as universal approximators. ANNs exhibit the capability to accurately approximate a broad class of functions, owing to their adeptness in

[5] *"Time Series Feature extraction based on scalable hypothesis tests".*

[6] *For an overview of the extracted features, consult: https://tsfresh.readthedocs.io/en/latest/text/list_of_features.html*

parallel processing information derived from the data. The model-building process does not necessitate any a priori assumption regarding the form of the model. Rather, the architecture of the network model is predominantly shaped by the inherent characteristics of the data.

The single hidden layer feedforward network stands as the prevailing model form extensively employed in time series modeling and forecasting. This model is distinguished by a network comprising three layers of elementary processing units connected through acyclic links. The mathematical representation capturing the relationship between the output and the inputs is as follows:

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-1}\right) + \varepsilon_t$$

where:

y_t represents the output

y_{t-1}, $y_{t-2}, \ldots, y_{t-p}$ are inputs

$\alpha_j (j = 0,1,2,\ldots,q)$ and $\beta_{ij} (i = 0,1,2,\ldots,p; j = 1,2,\ldots,q)$ are model parameters

and p is number of input nodes while q is the number of hidden nodes.

The logistic function is frequently employed as the transfer function for the hidden layer:

$$g(x) = \frac{1}{1 + \exp{(-x)}}$$

Oreshkin et al. (2020) introduce a modeling architecture denoted as the " ℓ -block" that is the backbone of the NBEATS construction (Neural Basis Expansion Analysis for Interpretable Time Series Forecasting) The architecture takes an input denoted as x_l and produces two vectors as output, namely $\hat{y}_l$ and $\hat{y}_l$. The input $x_l$represents a historical data window of a specified size, concluding with the most recent observation. Oreshkin et al. (2020) stipulate the size of this window as a multiple of the forecast horizon H, with window sizes ranging from 2H to 7H.

For subsequent blocks, the inputs consist of the residual outputs derived from the preceding blocks. Each block generates two distinct outputs: $\hat{y}_l$, which corresponds to the forecast of the block with a size of H, and $\hat{y}_l$, representing the optimal estimate of the block for x_l. This latter element is recognized in the literature as a "backcast."

Internally, the fundamental structural unit comprises two components. The initial component entails a fully connected network responsible for generating the forward θ_l^f and backward θ_l^b predictors for expansion coefficients. The second component encompasses the backward g_l^b and forward g_l^f basis layers, which receive the corresponding forward θ_l^f and backward θ_l^b expansion coefficients. Subsequently, these layers internally project these coefficients onto the set of basis functions, yielding both the backcast x and the forecast outputs $\hat{y}_l$.

The functioning of the initial segment of the l-th block is delineated by the subsequent equations:

$$h_{l,1} = FC_{l,1}(x_l)$$

$$h_{l,2} = FC_{l,2}(h_{l,1})$$

$$h_{l,3} = FC_{l,3}(h_{l,2})$$

$$h_{l,4} = FC_{l,3}(h_{l,3})$$

$$\theta_l^b = LINEAR_l^b(h_{l,4})$$

$$\theta_l^f = LINEAR_l^f(h_{l,4})$$

Here the LINEAR layer is just a linear projection layer such that:

$$\theta_l^f = W_l^f h_{l,4}$$

The fully connected (FC) layer is a conventional component characterized by a standard fully connected structure incorporating Rectified Linear Unit (RELU) non-linearity $h_{l,1} = RELU(W_{l,1}x_l + b_{l,1})$.

In this area of the model the forecast of expansion coefficients (θ_l^f) is conducted in order to optimize forecasting accuracy. The second segment of the network translates expansion coefficients to output through the utilization of the subsequent basis layers:

$$\hat{y}_l = g_l^f(\theta_l^f)$$

$$\hat{x}_l = g_l^b(\theta_l^b)$$

The corresponding equations are the following:

$$\hat{y}_l = \sum_{i=1}^{dim(\theta_l^f)} \theta_{l,i}^f v_i^f$$

$$\hat{y}_l = \sum_{i=1}^{dim(\theta_l^b)} \theta_{l,i}^b v_i^b$$

Challu et al. (2023) present the NHITS approach, an extension of the N-BEATS procedure. Analogous to N-BEATS, NHITS conducts local nonlinear projections onto basis functions across various blocks. Each block comprises a multilayer perceptron (MLP) tasked with acquiring the capability to generate coefficients for both the backcast and forecast outputs associated with its basis.

The backcast output is employed to refine the inputs for subsequent blocks, whereas the forecasts are aggregated to formulate the ultimate prediction. The blocks are organized into stacks, each specializing in learning distinct characteristics of the data by employing a unique set of basis functions.

At the initiation of each block ℓ, Challu et al. (2023) recommend the integration of a MaxPool layer with a kernel size of k_ℓ. This addition assists the block in focusing its analysis on components of the input associated with a specific scale.

Given the input to block ℓ, denoted as $y_{t-L:t,\ell}$ (where the input to the initial block, ℓ = 1, corresponds to the network-wide input, $y_{t-L:t,1} \equiv y_{t-L:t}$, this operation can be formalized as follows:

$$y_{t-L:t,\ell}^{(p)} = MaxPool(y_{t-L:t,\ell}, k_\ell)$$

Subsequent to subsampling, block ℓ examines its input and nonlinearly estimates the forward θ_ℓ^f and backward θ_ℓ^b interpolation MLP coefficients. These coefficients facilitate the learning of a hidden vector $h_\ell \in \mathbb{R}^{N_h}$, which is subsequently subjected to linear projection:

$$h_\ell = MLP_\ell\left(y_{t-L:t,\ell}^{(p)}\right)$$

$$\theta_\ell^f = LINEAR^f(h_\ell)$$

$$\theta_\ell^b = LINEAR^b(h_\ell)$$

Challu et al. (2023) employ temporal interpolation through the delineation of interpolation coefficients, formulated in relation to an expressiveness ratio that governs the allocation of parameters per unit of output time. The employed methodology integrates exponentially escalating expressiveness ratios, thereby accommodating a diverse spectrum of frequency bands while maintaining parameter control. Subsequently, the residual resulting from backcasting at the preceding hierarchical scale is subtracted from the input at the subsequent hierarchical level. This subtraction serves to augment the focus of the subsequent level's block on signals beyond the bandwidth previously addressed by the antecedent hierarchical members.

$$\hat{y}_{t+1:t+H} = \sum_{l=1}^{L} \hat{y}_{t+1:t+H,\ell}$$

$$y_{t-L:t,\ell+1} = y_{t-L:t,\ell} - \tilde{y}_{t-L:t,\ell}$$

Results

We generated predictions for each Boolean variable, including the one indicating the entire bubble, the one representing the boom phase, and the one indicating the burst phase of each bubble, utilizing three distinct deep neural models.

We segmented our sample into 100 sequential intervals, each comprising 100 consecutive observations for training purposes and 10 observations for making predictions. As previously noted, we incorporated exogenous (explanatory) variables, encompassing all relevant features extracted using the algorithm outlined by Christ et al. (2017), with the use of the first lag for these features.

Hence, leveraging the identification of patterns within the feature set from the preceding period, we generate sequential forecasts for the subsequent 10 observations in an out-of-sample fashion.

The table presented in Appendix A(Table A2) furnishes details on the precision of each model for every Boolean variable associated with each asset. Accuracy is calculated by tallying the instances in which the predicted value matches the actual value in a test sample of 10 observations and then dividing it by the total observations in a repeated exercise conducted 100 times.

Initially, our analysis is directed towards three dimensions. Our primary emphasis is placed on the comprehensive bubble, subsequently narrowing down to the boom and burst phases. A preliminary observation reveals that all models exhibit considerable efficacy in forecasting bubble phenomena. The precision of these predictions demonstrates variability among different assets and is also diverse with respect to the employed models.

In the comprehensive analysis of the entire bubble, it is observed that, although the predictive capacities of various models are comparable, NHITS and NBEATS consistently outperform MLP, particularly in instances characterized by lower accuracy. Additionally, instances arise wherein the accuracy of MLP significantly lags behind that of its counterparts.

Focusing on the precision of predictions during the boom phase unveils a reduced degree of volatility in contrast to the antecedent scenario. In this context, the performance of all three models demonstrates heightened similarity. While NHITS and NBEATS consistently produce the most accurate predictions, a discernible decline in MLP accuracy problems is evident compared to the earlier analysis. Nonetheless, instances are discerned where MLP outperforms the predictive capabilities of the other models.

Our final focal point pertains to the prediction accuracy during instances of identified bubble bursts. Observations reveal outcomes closely aligned with those obtained during the boom phase.

Overall, NHITS and NBEATS consistently exhibit heightened accuracy, although MLP attains its optimal overall performance in this context. The predictive accuracy of MLP closely aligns with that of the other models, with numerous instances demonstrating superior accuracy.

All findings are consistent with prior investigations undertaken in analogous domains. As an illustration, Ozgur et al. (2021) employ the Generalized Supremum Augmented Dickey-Fuller (GSADF) test for the identification of bubbles in metal prices. Subsequently, they apply a random forest algorithm to discern the pivotal factors contributing to the formation of price bubbles.

Ozgur et al. (2021) highlight that machine learning analyses reveal the potential manifestation of commodity market-specific characteristics in the occurrence of price bubbles. Moreover, the identification of threshold values for variables influencing the formation of bubbles in individual commodity prices can be leveraged to establish early warning indicators for bubble detection.

Furthermore, the utilization of deep learning algorithms for predicting cryptocurrency prices during bubble periods has been prevalent. Livieris et al. (2000) integrated Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Convolutional Neural Network (CNN) models to formulate a comprehensive model for hourly cryptocurrency price prediction and movement. This amalgamation resulted in enhanced accuracy in predicting cryptocurrency prices during crisis periods. Guarino et al. (2022) undertook an examination of the behavioral dynamics exhibited by Deep Reinforcement Learning (DRL)-based trading agents amidst the financial bubble, employing BTC and ETH datasets covering the period from January 1, 2015, to December 31, 2018. In a separate study, Sawhney et al. (2022) presented the CryptoBubbles model, specifically devised for a novel multispan recognition task with a primary emphasis on bubble detection. Their methodology encompassed a dataset comprising more than 400 cryptocurrencies sourced from 9 exchanges, introducing a distinctive task of span identification and dataset for detection grounded in the power-law dynamics inherent in cryptocurrencies and user interaction patterns on social media platforms. Finally, Montasser et al. (2022) conducted a test based on Dynamic Time Warping alongside clustering analysis of 18 cryptocurrencies to investigate their market efficiency parallels. Their findings elucidated COVID-19 as the central factor contributing to the genesis of price bubbles.

Conclusions

The study concludes that the integration of the BSADF test with neural network models offers a robust framework for predicting financial bubbles in the STOXX 600 index. The ability of the models to accurately forecast bubble phases highlights their potential as valuable tools for investors and policymakers to mitigate risks associated with financial bubbles. These findings underscore the importance of combining advanced statistical tests with machine learning techniques to enhance the predictability of market anomalies, paving the way for more informed investment decisions and policy formulations.

The findings of this study are crucial as they highlight the predictive capability of combining the BSADF test with neural network models for identifying financial bubbles within the STOXX 600 index. This approach not only enhances our understanding of market dynamics but also offers practical tools for risk management. By demonstrating the models' accuracy in forecasting bubble phases, this research supports the development of more sophisticated investment strategies and regulatory policies aimed at stabilizing financial markets. The implications extend beyond academia, affecting investors, policymakers, and the broader financial community, emphasizing the need for advanced analytics in financial decision-making.

The findings of this study offer policymakers the opportunity to craft more informed and forward-looking regulations aimed at preemptively mitigating the risks posed by financial bubbles. Utilizing predictive analytics enables the early detection of emerging bubbles, facilitating interventions to lessen their economic repercussions. These insights advocate for the creation of adaptive regulatory frameworks, better aligned with evolving market conditions, thereby bolstering financial stability and safeguarding the broader economy's welfare.

Future developments of this research could involve expanding the dataset to include a wider range of financial instruments and geographic markets to assess the universality of the proposed models. Additionally, exploring the integration of alternative data sources, such as social media sentiment or macroeconomic indicators, could enhance model accuracy. Advancements in machine learning techniques, such as deep learning and reinforcement learning, offer promising avenues to improve predictive capabilities further. Collaborations with financial institutions for real-world application and validation of these models could also provide practical insights and refine their effectiveness in detecting financial bubbles.

Acknowledgement

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS - UEFISCDI, project number PN-III-P1-1.1- TE-2021-1339, within PNCDI III.

References

Abreu, D. and Brunnermeier, M.K., 2003. Bubbles and crashes. *Econometrica*, 71(1), pp.173-204. https://doi.org/10.1111/1468-0262.00393.

Akcora, B. and Kocaaslan, O.K., 2023. Price bubbles in the European natural gas market between 2011 and 2020. *Resources Policy*, 80, 103186.

André, C., Caraiani, P., Călin, A.C. and Gupta, R., 2022. Can monetary policy lean against housing bubbles? *Economic Modelling*, 110, 105801. https://doi.org/10.1016/j.econmod.2022.105801.

Bago, J.L., Rherrad, I., Akakpo, K. and Ouédraogo, E., 2021. Real Estate Bubbles and Contagion: Evidence from Selected European Countries. *Review of Economic Analysis*, 13(4), pp.389-405.

Caraiani, P. and Călin, A.C., 2019. Housing markets, monetary policy, and the international co-movement of housing bubbles. *Review of International Economics*, 28(2), pp.365–375. https://doi.org/10.1111/roie.12454.

Caraiani, P. and Călin, A.C., 2019. Monetary Policy Effects on Energy Sector Bubbles. *Energies*, 12(3), pp.472–472. https://doi.org/10.3390/en12030472.

Caspi, I., Katzke, N. and Gupta, R., 2018. Date stamping historical periods of oil price explosivity: 1876–2014. *Energy Economics*, 70, pp.582–587. https://doi.org/10.1016/j.eneco.2015.03.029.

Challu, C., Olivares, K.G., Oreshkin, B.N., Ramirez, F.G., Canseco, M.M. and Dubrawski, A., 2023. Nhits: Neural hierarchical interpolation for time series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(6), pp.6989-6997.

Chang, C-L., 2024. Extreme events, economic uncertainty and speculation on occurrences of price bubbles in crude oil futures. *Energy Economics*, 130, 107318. https://doi.org/10.1016/j.eneco.2024.107318

Chen, S.-W. and Xie, Z., 2017. Detecting speculative bubbles under considerations of the sign asymmetry and size non-linearity: New international evidence. *International Review of Economics & Finance*, 52, pp.188–209. https://doi.org/10.1016/j.iref.2017.09.008.

Chowdhury, S.R., Damianov, D.S. and Elsayed, A.H., 2022. Bubbles and crashes in cryptocurrencies: Interdependence, contagion, or asset rotation? *Finance Research Letters*, 46, 102494. https://doi.org/10.1016/j.frl.2021.102494.

Christ, M., Braun, N., Neuffer, J. and Kempa-Liehr, A.W., 2018. Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package). *Neurocomputing*, 307, pp.72-77. https://doi.org/10.1016/j.neucom.2018.03.067.

Christ, M., Kempa-Liehr, A.W. and Feindt, M., 2016. Distributed and parallel time series feature extraction for industrial big data applications. https://doi.org/10.48550/arXiv.1610.07717.

De Long, J.B., Shleifer, A., Summers, L.H. and Waldmann, R.J. 1990. Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, 45: pp.379-395. https://doi.org/10.1111/j.1540-6261.1990.tb03695.x

Diba, B.T. and Grossman, H.I., 1988. The Theory of Rational Bubbles in Stock Prices. *The Economic Journal*, 98(392), pp.746–746. https://doi.org/10.2307/2233912

Evgenidis, A. and Malliaris, A., 2023. House Bubbles, global imbalances and monetary policy in the US. *Journal of International Money and Finance*, 138, 102919. https://doi.org/10.1016/j.jimonfin.2023.102919

Galí J., Giusti, G. and Noussair, C.N., 2021. Monetary Policy and Asset Price Bubbles: A Laboratory Experiment. *Journal of Economic Dynamics and Control*, 130, 104184. https://doi.org/10.1016/j.jedc.2021.104184

Garber, P.M. 1990. Famous First Bubbles. Journal of Economic Perspectives, 4(2), pp.35–54. https://doi.org/10.1257/jep.4.2.35

Geuder, J., Kinateder, H. and Wagner, N.F., 2019. Cryptocurrencies as financial bubbles: The case of Bitcoin. *Finance Research Letters*, 31. https://doi.org/10.1016/j.frl.2018.11.011

Goodman, R.D. and West-Olatunji, C.A., 2008. Transgenerational Trauma and Resilience: Improving Mental Health Counseling for Survivors of Hurricane Katrina. *Journal of Mental Health Counseling*, 30(2), pp.121–136. https://doi.org/10.17744/mehc.30.2.q52260n242204r84

Gronwald, M., 2016. Explosive oil prices. *Energy Economics*, 60, pp.1–5. https://doi.org/10.1016/j.eneco.2016.09.012.

Guarino, A., Grilli, L., Santoro, D., Messina, F. and Zaccagnino, R., 2022. To learn or not to learn? Evaluating autonomous, adaptive, automated traders in cryptocurrencies financial bubbles. *Neural Computing and Applications*, 34(23), pp.20715-20756. https://doi.org/10.1007/s00521-022-07543-4.

Hashimoto K., Im R., Kunieda T. (2020), *Asset Bubbles, Unemployment, and a Financial Crisis, Journal of Macroeconomics*, Volume 65, 103212, ISSN 0164-0704, https://doi.org/10.1016/j.jmacro.2020.103212.

Hirano, T. and Toda, A.A., 2024. Bubble economics. *Journal of Mathematical Economics*, 102944. https://doi.org/10.1016/j.jmateco.2024.102944.

Homm, U. and Breitung, J., 2012. Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods. *Journal of Financial Econometrics*, 10(1), pp.198–231. https://doi.org/10.1093/jjfinec/nbr009

Hott, C., 2009. Herding behavior in asset markets. *Journal of Financial Stability*, 5(1), pp.35–56. https://doi.org/10.1016/j.jfs.2008.01.004.

Junmin, W., 2018. Prevention and landing of bubble. *International Review of Economics and Finance*, 56, pp.190-204. https://doi.org/10.1016/j.iref.2017.10.024.

Junmin, W., 2024. Transmission of housing bubbles among industrial sectors. *International Review of Economics and Finance,* 89, pp.692-701. https://doi.org/10.1016/j.iref.2023.08.001.

Khan, K., Chi Wei Su, and Khurshid, A., 2022. Do booms and busts identify bubbles in energy prices? *Resources Policy*, 76, 102556. https://doi.org/10.1016/j.resourpol.2022.102556.

Khan, K., Su, C.-W. and Rehman, A.U., 2021. Do multiple bubbles exist in coal price? *Resources Policy*, 73, 102232. https://doi.org/10.1016/j.resourpol.2021.102232.

Khan, K., Su, C.-W., Umar, M. and Yue, X.-G., 2021. Do crude oil price bubbles occur? *Resources Policy*, 71, 101936. https://doi.org/10.1016/j.resourpol.2020.101936.

Kivedal, B.K., 2013. Testing for rational bubbles in the US housing market. *Journal of Macroeconomics,* 38, pp.369–381. https://doi.org/10.1016/j.jmacro.2013.08.021.

Kyriazis, N., Papadamou, S. and Corbet, S., 2020. A systematic review of the bubble dynamics of cryptocurrency prices. *Research in International Business and Finance*, 54, 101254. https://doi.org/10.1016/j.ribaf.2020.101254.

Li, Y., Chevallier, J., Wei, Y. and Li, J., 2020. Identifying price bubbles in the US, European and Asian natural gas market: Evidence from a GSADF test approach. *Energy Economics*, 87, 104740. https://doi.org/10.1016/j.eneco.2020.104740.

Livieris, I.E., Pintelas, E., Stavroyiannis, S, and Pintelas, P., 2020. Ensemble deep learning models for forecasting cryptocurrency time series. Algorithms, 13(5), 121. https://doi.org/10.3390/a13050121.

Maas, D., Mayer, E. and Rüth, S.K., 2018. Current account dynamics and the housing cycle in Spain. *Journal of International Money and Finance*, 87, pp.22–43. https://doi.org/10.1016/j.jimonfin.2018.05.007.

Martin, C., Schmitt, N. and Westerhoff, F., 2021. Heterogeneous expectations, housing bubbles and tax policy. *Journal of Economic Behavior & Organization*, 183, pp.555-573. https://doi.org/10.1016/j.jebo.2020.12.033.

Masaya, S., 2021. An Economic Theory of Bubbles: Low Interest Rates, Secular Stagnation, and Financial Deterioration. Tokyo: Nikkei Business Publications, Inc.

Montasser, G.E., Charfeddine, L. and Benhamed, A., 2022. Covid-19, cryptocurrencies bubbles and digital market efficiency: sensitivity and similarity analysis. *Finance Research Letters*, 46, 102362 https://doi.org/10.1016/j.frl.2021.102362.

Moons, C. and Hellinckx, K., 2019. Did monetary policy fuel the housing bubble? An application to Ireland. *Journal of Policy Modeling*, 41(2), pp.294–315. https://doi.org/10.1016/j.jpolmod.2019.03.006.

Oreshkin, B.N., Carpov, D., Chapados, N. and Bengio, Y., 2019. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. https://doi.org/10.48550/arXiv.1905.10437.

Ott, H., 2014. Will euro area house prices sharply decrease? *Economic Modelling*, 42, pp.116–127. https://doi.org/10.1016/j.econmod.2014.06.004.

Ozgur, O., Yilanci, V., Ozbugday, F.C., 2021. Detecting speculative bubbles in metal prices: Evidence from GSADF test and machine learning approaches. *Resources Policy*, 74, 102306. https://doi.org/10.1016/j.resourpol.2021.102306.

Pavlidis, E., Martínez-García, E. and Grossman, V., 2019. Detecting periods of exuberance: A look at the role of aggregation with an application to house prices. *Economic Modelling,* 80, pp.87–102. https://doi.org/10.1016/j.econmod.2018.07.021.

Phillips, P.C.B., Shi, S. and Yu, J., 2015. Testing For Multiple Bubbles: Historical Episodes Of Exuberance And Collapse In The S&P 500. *International Economic Review*, 56, pp.1043-1078. https://doi.org/10.1111/iere.12132.

Rahal, C., 2016. Housing markets and unconventional monetary policy. *Journal of Housing Economics*, 32, pp.67–80. https://doi.org/10.1016/j.jhe.2016.04.005.

Risse, M. and Kern, M., 2016. Forecasting house-price growth in the Euro area with dynamic model averaging. *The North American Journal of Economics and Finance*, 38, pp.70–85. https://doi.org/10.1016/j.najef.2016.08.001.

Sawhney, R., Agarwal, S., Mittal, V., Rosso, P., Nanda, V. and Chava, S., 2022. Cryptocurrency Bubble Detection: A New Stock Market Dataset, Financial Task & Hyperbolic Models. https://doi.org/10.48550/arXiv.2206.06320.

Sharma, S. and Escobari, D., 2018. Identifying price bubble periods in the energy sector. *Energy Economics*, 69, pp.418–429. https://doi.org/10.1016/j.eneco.2017.12.007.

Shi, S., 2017. Speculative bubbles or market fundamentals? An investigation of US regional housing markets. *Economic Modelling*, 66, pp 101-111. https://doi.org/10.1016/j.econmod.2017.06.002

Shiller, R.C., 2000. Irrational Exuberance. *Philosophy and Public Policy Quarterly*, 20, pp.18-23.

Shiller, R.J., 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review*, 71(3), pp.421–436.

Teti, E. and Maroni, D., 2021. The new great bubble in the technology industry?. *Technology Analysis & Strategic Management*, 33(5), pp.520-534. https://doi.org/10.1080/09537325.2020.1828577.

Tokic, D., 2010. The 2008 oil bubble: Causes and consequences. *Energy Policy*, 38(10), pp.6009–6015. https://doi.org/10.1016/j.enpol.2010.05.056.

Tsai, I-C. and Lin C-C., 2022. A re-examination of housing bubbles: Evidence from European countries. *Economic Systems*, 46(2), 100971. https://doi.org/10.1016/j.ecosys.2022.100971.

Tsai, I.-C., 2019. European house price deviation: infectivity and the momentum effect. *Economic Research-Ekonomska Istraživanja*, 32(1), pp.1521-1541. https://doi.org/10.1080/1331677X.2019.1636698.

Tsai, I-C., 2018. House price convergence in euro zone and non-euro zone countries. *Economic Systems*, 42(2), pp.269–281. https://doi.org/10.1016/j.ecosys.2017.05.010

Wang, S., Feng, H. and Gao. D., 2023. Testing for short explosive bubbles: A case of Brent oil futures price. *Finance Research Letters*, 52, 103497. https://doi.org/10.1016/j.frl.2022.103497.

West, K.D. 1987. A Specification Test for Speculative Bubbles. *The Quarterly Journal of Economics*, 102(3), pp.553 – 580. https://doi.org/10.2307/1884217.

Yao, C.-Z. and Li, H.-Y., 2021. A study on the bursting point of Bitcoin based on the BSADF and LPPLS methods. *The North American Journal of Economics and Finance*, 55, 101280. https://doi.org/10.1016/j.najef.2020.101280.

Zekkari, K.B., 2024. Asset bubble and growth: Elastic labor supply with fiscal policy. *Journal of Mathematical Economics*, 110, 102930. https://doi.org/10.1016/j.jmateco.2023.102930.

Zhang, D., Shi, M. and Shi, X., 2018. Oil indexation, market fundamentals, and natural gas prices: An investigation of the Asian premium in natural gas trade. *Energy Economics*, 69, pp.33–41. https://doi.org/10.1016/j.eneco.2017.11.001.

Zhang, D., Wang, T., Shi, X. and Liu, J., 2018. Is hub-based pricing a better choice than oil indexation for natural gas? Evidence from a multiple bubble test. *Energy Economics*, 76, pp.495–503. https://doi.org/10.1016/j.eneco.2018.11.001.

Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, pp.159-175. https://doi.org/10.1016/S0925-2312(01)00702-0.

Zhu, B., Betzinger, M. and Sebastian, S., 2017. Housing market stability, mortgage market structure, and monetary policy: Evidence from the euro area. *Journal of Housing Economics*, 37, pp.1–21. https://doi.org/10.1016/j.jhe.2017.04.001.

Appendix A

Table A1. Excerpt of features extracted for rolling windows of 100 observations for Aeroports de Paris SA

	prices variance larger than standard deviation	prices has duplicate max	prices has duplicate min	prices has duplicate	prices sum values	prices abs energy	prices mean abs change	prices fft coefficient attr "real" coeff 47	prices fft coefficient attr "real" coeff 48	prices fft coefficient attr "real" coeff 49	prices fft coefficient attr "real" coeff 50
2015-05-22 00:00:00	1	0	0	1	10941.62	1198725.26	1.05	-2.43	-8.03	-7.03	1.94
2015-05-25 00:00:00	1	0	0	1	10957.54	1202130.23	1.05	-12.27	-8.77	-9.29	-17.86
2015-05-26 00:00:00	1	0	0	1	10972.70	1205355.37	1.04	-4.28	-5.52	-5.49	2.70
2015-05-27 00:00:00	1	0	0	1	10987.90	1208643.13	1.05	-9.42	-10.51	-10.09	-17.90
2015-05-28 00:00:00	1	0	0	1	11000.20	1211283.94	1.07	-4.50	-1.02	-1.85	5.60
2015-05-29 00:00:00	1	0	0	1	11010.50	1213498.44	1.07	-4.16	-9.98	-8.80	-15.90
2015-06-01 00:00:00	1	0	0	1	11019.30	1215392.20	1.07	-6.27	1.79	0.33	7.10
2015-06-02 00:00:00	1	0	0	1	11027.00	1217048.47	1.07	0.12	-10.01	-8.35	-14.80
2015-06-03 00:00:00	1	0	0	1	11035.75	1218935.40	1.05	-10.26	1.70	-0.10	6.05
2015-06-04 00:00:00	1	0	0	1	11040.35	1219930.38	1.06	6.76	-6.67	-4.79	-10.65
2015-06-05 00:00:00	1	0	0	1	11043.90	1220689.20	1.07	-11.11	3.41	1.51	7.10
2015-06-08 00:00:00	1	0	0	1	11044.70	1220859.28	1.08	10.74	-4.45	-2.59	-7.90
2015-06-09 00:00:00	1	0	0	1	11045.15	1220956.36	1.09	-11.22	4.17	2.40	7.45
2015-06-10 00:00:00	1	0	0	1	11045.80	1221097.06	1.09	10.22	-4.91	-3.30	-8.10
2015-06-11 00:00:00	1	0	0	1	11046.10	1221161.92	1.07	-9.81	4.63	3.24	7.80
2015-06-12 00:00:00	1	0	0	1	11043.40	1220571.43	1.06	11.39	-1.90	-0.77	-5.10

Source: Authors' computation

Table A2. Accuracy of prediction for all assets and models

	Whole Bubble			Boom Phase			Burst Phase		
	MLP	NHITS	NBEATS	MLP	NHITS	NBEATS	MLP	NHITS	NBEATS
ADPFP	0.945	0.945	0.942	0.962	0.961	0.961	0.962	0.961	0.961
AEDBB	0.826	0.837	0.837	0.915	0.919	0.919	0.915	0.919	0.919
AENASQ	1	1	1	1	1	1	1	1	1
AKRBPNO	1	1	1	1	1	1	1	1	1
ALVGY	1	1	1	1	1	1	1	1	1
ANDRAV	1	1	1	1	1	1	1	1	1
AXFOSS	1	1	1	1	1	1	1	1	1
BASGY	1	1	1	1	1	1	1	1	1
BATSLN	1	1	1	1	1	1	1	1	1
BIMFP	1	1	1	1	1	1	1	1	1
BRBYLN	1	1	1	1	1	1	1	1	1
BUCNSE	1	1	1	1	1	1	1	1	1
BWYLN	1	1	1	1	1	1	1	1	1
CABKSQ	1	1	1	1	1	1	1	1	1
CCHLN	0.937	0.944	0.955	0.913	0.994	0.992	0.913	0.994	0.992
CDIFP	0.926	0.919	0.915	0.988	0.996	0.988	0.988	0.996	0.988
CHRDC	1	1	1	1	1	1	1	1	1
CLNSE	0.842	0.846	0.804	0.909	0.933	0.92	0.909	0.933	0.92
CNALN	0.35	1	1	0.505	0.945	0.945	0.505	0.945	0.945
CNHIIM	1	1	1	1	1	1	1	1	1
COFBBB	0.601	0.62	0.603	0.788	0.856	0.862	0.788	0.856	0.862
COLSQ	1	1	1	1	1	1	1	1	1
CONGY	1	1	1	1	1	1	1	1	1
CPGLN	0.945	0.945	0.945	0.973	0.973	0.973	0.973	0.973	0.973
CRDALN	1	1	1	1	1	1	1	1	1
CSFP	1	1	1	1	1	1	1	1	1
DANSKEDC	1	1	1	1	1	1	1	1	1
DB1GY	1	1	1	1	1	1	1	1	1
DBKGY	1	1	1	1	1	1	1	1	1
EKTABSS	1	1	1	1	1	1	1	1	1
EQNRNO	1	1	1	1	1	1	1	1	1
ERFFP	1	1	1	1	1	1	1	1	1
ESSITYBSS	0.78	0.794	0.777	0.83	0.844	0.827	0.83	0.844	0.827

	Whole Bubble			Boom Phase			Burst Phase		
	MLP	NHITS	NBEATS	MLP	NHITS	NBEATS	MLP	NHITS	NBEATS
EVKGY	1	1	1	1	1	1	1	1	1
FDJFP	0.922	0.913	0.912	0.957	0.956	0.958	0.957	0.956	0.958
FORTUMFH	1	1	1	1	1	1	1	1	1
FRFP	0.964	0.964	0.964	0.979	0.979	0.979	0.979	0.979	0.979
GALESE	1	1	1	1	1	1	1	1	1
GAWLN	0.51	0.88	0.889	0.56	0.882	0.892	0.56	0.882	0.892
HEN3GY	0.974	0.9	0.909	0.997	0.999	0.997	0.997	0.999	0.997
HOLMBSS	1	1	1	1	1	1	1	1	1
HOLNSE	1	1	1	1	1	1	1	1	1
HSBALN	1	1	1	1	1	1	1	1	1
INCHLN	1	1	1	1	1	1	1	1	1
INDTSS	1	1	1	1	1	1	1	1	1
INDUCSS	1	1	1	1	1	1	1	1	1
IPNFP	1	1	1	1	1	1	1	1	1
ITVLN	1	1	1	1	1	1	1	1	1
JMATLN	1	1	1	1	1	1	1	1	1
KINDSDBSS	1	1	1	1	1	1	1	1	1
LONNSE	1	1	1	1	1	1	1	1	1
LRFP	1	1	1	1	1	1	1	1	1
LXILN	1	1	1	1	1	1	1	1	1
MKSLN	0.875	0.896	0.903	0.677	0.687	0.691	0.677	0.687	0.691
MTXGY	0.884	0.89	0.9	0.802	0.873	0.862	0.802	0.873	0.862
NEXIIM	0.93	0.93	0.93	0.932	0.932	0.932	0.932	0.932	0.932
OMVAV	1	1	1	1	1	1	1	1	1
PHIANA	1	1	1	1	1	1	1	1	1
PKNPW	1	1	1	1	1	1	1	1	1
PNNLN	0.818	0.792	0.804	0.865	0.863	0.854	0.865	0.863	0.854
PSNLN	1	1	1	1	1	1	1	1	1
PUBFP	0.995	0.982	0.98	1	1	1	1	1	1
REPSQ	1	1	1	1	1	1	1	1	1
RFFP	1	1	1	1	1	1	1	1	1
RHMGY	0.06	1	1	1	1	1	1	1	1
RILBADC	0	1	1	0.997	1	0.999	0.997	1	0.999
RNOFP	1	1	1	1	1	1	1	1	1

	Whole Bubble			Boom Phase			Burst Phase		
	MLP	**NHITS**	**NBEATS**	**MLP**	**NHITS**	**NBEATS**	**MLP**	**NHITS**	**NBEATS**
ROCKBDC	1	1	1	1	1	1	1	1	1
SAFFP	0.861	0.936	0.936	0.97	0.959	0.917	0.97	0.959	0.917
SALMNO	1	1	1	1	1	1	1	1	1
SAPGY	1	1	1	1	1	1	1	1	1
SBRYLN	1	1	1	1	1	1	1	1	1
SCABSS	1	1	1	1	1	1	1	1	1
SGROLN	1	1	1	1	1	1	1	1	1
SKFBSS	1	1	1	1	1	1	1	1	1
SLHNSE	1	1	1	1	1	1	1	1	1
SOIFP	0.412	1	1	1	1	1	1	1	1
SPLPW	0.719	0.741	0.726	0.679	0.694	0.599	0.679	0.694	0.599
SRT3GY	1	1	1	1	1	1	1	1	1
STLAMIM	1	1	1	1	1	1	1	1	1
SWECBSS	1	1	1	1	1	1	1	1	1
TEFSQ	1	1	1	1	1	1	1	1	1
TENIM	1	1	1	1	1	1	1	1	1
TIGOSS	1	1	1	0.993	0.981	0.986	0.993	0.981	0.986
TKAGY	1	1	1	1	1	1	1	1	1
TOPDC	1	1	1	1	1	1	1	1	1
TRELBSS	1	1	1	1	1	1	1	1	1
UPMFH	1	1	1	1	1	1	1	1	1
VALMTFH	1	1	1	1	1	1	1	1	1
VIVFP	1	1	1	1	1	1	1	1	1
VNAGY	0.667	0.915	0.917	1	1	1	1	1	1
VOW3GY	1	1	1	1	1	1	1	1	1
WEIRLN	1	1	1	1	1	1	1	1	1

Source: Authors' computation