JEBI

Journal of Economics and Business Issues

Forecasting House Prices in Albania with the Deep Learning LSTM Network#

Blerina Vika [1], and Ilir Vika [2,*]

[1] Department of Statistics and Applied Informatics, Faculty of Economy, University of Tirana; blerina.vika@unitir.edu.al
[2] Research Department, Bank of Albania; ivika@bankofalbania.org
* Correspondence: blerina.vika@unitir.edu.al

#) *Disclaim: This research article is considered as a preliminary work that aims at stimulating debate and critical comments. It expresses the views of the authors and does not necessarily represent those of the institutions where they work. The authors are grateful to comments and suggestions received from participants at the "Istanbul Annual International Multidisciplinary Conference on Economics, Business, Technology and Social Sciences - 2023", organized by the European Institute for Research and Development (EIRD), Istanbul, Türkiye, in May 13-14, 2023.*

Abstract: This article investigates the role of economic, financial and demographic indicators in forecasting house prices in Albania. The pool of variables is drawn from empirical studies for advanced and developing countries. To test their importance, we employ the long short-term memory network from the machine learning techniques. As the time span of observations is rather limited, the specification of models is maintained to be parsimonious and sufficient to capture the most important dynamics. As such, we compare the performance of a univariate network with models containing the most related variables such as GDP and actual rental, and then augment them with bank loans and interest rates, demand from non-residents, unemployment rate, urban population, cost of construction, and area of building permits. The forecast ability is evaluated during the 2018-2022 period for horizons at 1, 4, 8 and 12 quarters ahead. Preliminary results suggest that multivariate, theory-driven models can help improve upon forecasts generated from the univariate network. Apart from GDP and rentals, costs of construction and financial indicators are some additional variables in which forecasters may have confidence on when predicting residential house prices in Albania.

Keywords: House prices; Machine Learning; LSTM model; Albania

Citation: Vika, B.; & Vika, I. (**2023**) Forecasting House Prices in Albania with the Deep Learning LSTM Network. *Journal Of Economic and Business Issues, 4(1)*, 40-55.

Received: 30/05/2023
Accepted: 18/11/2023
Published: 29/02/2024

1. Introduction

Albania's house prices have been growing much faster than in Euro area as well as most Eastern European countries (EECs) (Figure 1). The rapid increase in prices, especially since 2018, has naturally raised the question of a speculative bubble in the housing market. The concern is mostly related to a fall in prices that could have potential adverse consequences on the broader economy. The flourishment in asset prices that preceded the Asian and US financial crises in 1997 and 2008, respectively, was largely attributed to speculative real estate price bubbles. These days, the fearsome housing market realignment looms at a time when many central banks have been raising their policy rates in order to fight inflation, thus reversing the easy financial conditions that were enjoyed for almost a decade. The higher costs of financing may impede economic growth and the recent boom in property sector and perhaps put downward pressure on house prices.

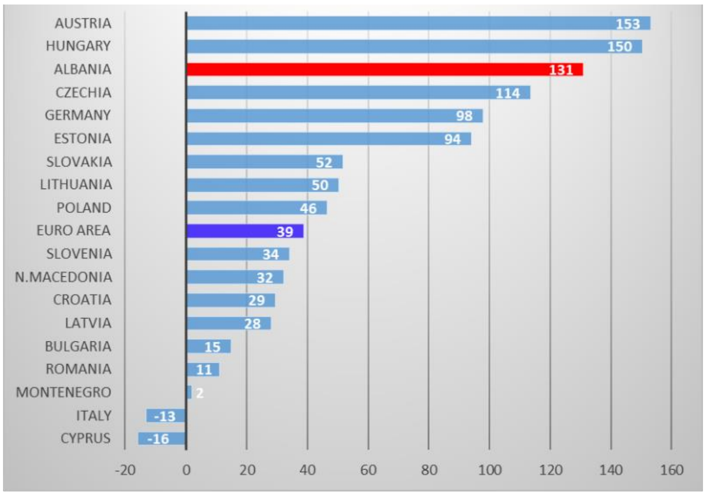


Figure 1. Cumulative growth (%) of house prices (2008-2022).

Real estate is regarded as a potential investment opportunity to government securities, especially in small countries like Albania where stock markets or alternative investment forms don't exist or are still underdeveloped. A reversal of the housing boom could have significant implications for households and policymakers, as housing is a crucial household asset and mortgages contribute substantially to household liabilities. Also, soaring housing prices might make homeownership unaffordable for many low-income households. Further increases in mortgage rates in domestic or foreign currencies might exacerbate affordability concerns. For these reasons, the link of housing market with other macroeconomic indicators and its direct and indirect effects on financial stability has propelled the need to monitor the developments of, and predict the future course of house prices.

This article aims to shed light on which theory-driven indicators are important in forecasting residential property prices in Albania. Traditionally, the relationship between house price growth and its economic fundamentals has been studied by using standard econometric techniques such as cointegration methods. However, our analysis explores the potential of the innovative long short-term memory (LSTM) neural network, which belong to the deep learning methods within the innovative Artificial Intelligence techniques. Neural networks allow for flexible mapping of variables, enabling accurate approximation of highly non-linear functions. Unlike standard econometric methods, neural networks do not require the specification of regression parameters.

Neural networks are becoming especially popular in economic analyses where traditional statistical techniques are insufficient. The process of establishing a neural network's architecture is similar to curve fitting, with the number of hidden layers and neurons determining the model's complexity. Using too few layers or neurons leads to poor fit and prediction, while using too many can result in overfitting. Neural networks possess inductive capabilities, allowing them to model complex systems even without precise knowledge of underlying rules. This article utilizes neural networks to address barriers faced in accurately forecasting house prices in Albania, such as the absence of a consensus model, disputes over included variables, and data measurement challenges.

The study begins by taking a glance at the literature on housing price dynamics. It further investigates through graphical inspection the drivers and implications of fluctuating housing prices, in order to assess factors behind price surges, particularly in the last five years. The next two sections highlight concerns about data measurement to proxy for house price fundamentals and briefly explain the model and forecasting procedure. Empirical results are discussed in Section 6, followed by concluding remarks in Section 7.

2. A Brief Literature Review

The relationship between house prices and macroeconomic fundamentals has been relatively overlooked in academic research, although housing is a significant component of household wealth, especially in developing countries, and it played a crucial role in driving the global financial crisis through the financial accelerator mechanism. Nevertheless, the literature has expanded after the global financial crisis.

A survey paper by Duca et al. [1] on the wide international literature on housing market dynamics draws five important lessons. Firstly, conventional theories based on efficient market rational expectations appear insufficient to explain housing market dynamics. Heterogeneity, trading costs, asymmetric information, and credit constraints contribute to slower price adjustments. Expectations-driven dynamics play a substantial role in housing booms and busts. Secondly, simplistic house price-to-rent arbitrage models are inadequate, as they fail to capture the variation in house prices due to factors like imperfect substitution, sticky rents, transaction costs, risk aversion, and price volatility. Credit constraints and the shadow price of credit also affect the relationship between house prices and rents. Thirdly, credit conditions are found to be a key driver of house prices, with shifts in credit availability significantly impacting price movements. Fourthly, differences in land supply responses explain much of the observed spatial variation in house prices. Areas with less elastic land supply experience more frequent and larger price bubbles. Housing supply and credit availability also influence national-level house prices. Finally, during the Covid-19 pandemic, house prices behaved differently than in previous downturns due to factors such as different shocks, a better-capitalized financial system, and government interventions.

These findings caution against putting too much weight on the role of real interest rates, as other factors like changing demand for detached housing or reduced supply contribute to the house price booms and busts. The connection between macro-financial factors and residential real estate prices has been examined in the literature for advanced as well as emerging market economies. Demand-side factors like income, wealth, and financial conditions, along with supply-side factors such as housing availability, play a role in determining long-term housing prices. However, the impact of these factors varies depending on the countries, time periods, and methodologies used in the studies (see for e.g., Tsatsaronis and Zhu [2]; Égert and Mihaljek [3]; Agnello and Schuknecht [4]; Cerutti, Dagher, and Dell'Ariccia [5]).

In the context of Albania, there are a number of studies that have investigated the property market developments and the house price predictive power by using various linear traditional models. Using certain macroeconomic indicators, Yzeiraj [6] tries to construct a "fundamental" house price index for Albania. Although the fundamental-based index varied from the actual HPI due to the indicator used as a proxy for rent, all measurements pointed to a certain degree of HPI overheating after 2006 and a period of "correction" starting from 2012. Suljoti [7] employs a vector error correction method (VECM) to investigate the role of mortgage loans on house price growth in Albania. Her findings reveal that house prices are strongly linked to financial leverage and the exchange rate but not so much with the interest rate on new euro loans. Still, the author hints that social demographic factors together with institutional obstacles (e.g. property rights and registration) might be important influencers of housing market developments. Three studies focusing on house prices in the capital city, Tirana, find that most explanatory variables are consistent with hedonic pricing theory (Kraja et al. [8]) and there is a statistically significant (long-run) relationship with mortgage loans, long-term lending rates and costs of construction (Marku et al. [9]) or remittances (Lleshaj and Korbi [10]). Yet, other attempts on country-wide prices conclude that it is difficult to establish a theoretically-relevant relationship with all of the explanatory variables (Koprencka et al. [11]) or the impact of macroeconomic variables seems to be time-varying, particularly with respect to the mortgage rate (Halili [12]).

As such, forecasting house prices remains a challenging task due to the complex interactions among the economic, financial, and demographic indicators. In recent years, machine learning (ML) techniques, such as the long short-term memory network (LSTM), have shown promise in improving the accuracy of house price forecasts. Several studies have applied ML models to predict house prices in various countries, including advanced and developing economies (see for e.g., Mora-Garcia et al. [13]; Wang and Li [14]; Park and Bae [15]; Banerjee et al. [16]; Kok et al. [17]; Ceh et al. [18]; Fan et al. [19]; Ho et al. [20]; Chatzidis [21]; Alfaro-Navarro et al. [22]; Hong [23]; Hacıevliyagil et al. [24]). Their results outperformed traditional time series models by achieving better forecast accuracy. To the best of our knowledge, there is no research article for Albania yet to have tried to take advantage of neural network tools. Hence, it added to our motivation to enrich the modest empirical literature for housing market in Albania with innovative and encouraging techniques that can deal with theoretical and statistical drawbacks, as well as data measurement challenges.

3. House Price Developments and Long-Run Stylized Facts

The housing price index (HPI) in Albania has increased significantly from 2000 to 2022, both in nominal and real terms (adjusted for consumer prices, CPI). In fact, the cumulative increase in the HPI is calculated to be about five times higher than that of the CPI. However, the performance of the HPI index has not been the same over time. A look at the HPI index developments in Figure 2 suggests that the 2000-22 period can be divided into i) a decade 2000-09 of rapid house price growth (with a cumulative increase of 146%); followed by ii) a period of price "stabilization", albeit fluctuating until 2017; and then iii) a strong upward HPI trend reappears during 2018-22, with prices increasing nearly 76% at a firm pace as reflected in its lower volatility.

It seems that in the long run the cumulative increase of HPI index (402%) reflects the performance of nominal gross domestic production (455%), and is not much related to the consumer price basket (78%) and even less to that of construction cost index, CCI (12%) (please see Table 1). This means that the galloping growth of the HPI index in the last five years may be the result of the need to adjust after a relatively long stagnation of prices during the period of high uncertainties related to, among other things, the shocks of the global crisis, the debt crisis in the Eurozone, and the gridlock situation in the domestic financial sector up the middle of the last decade.

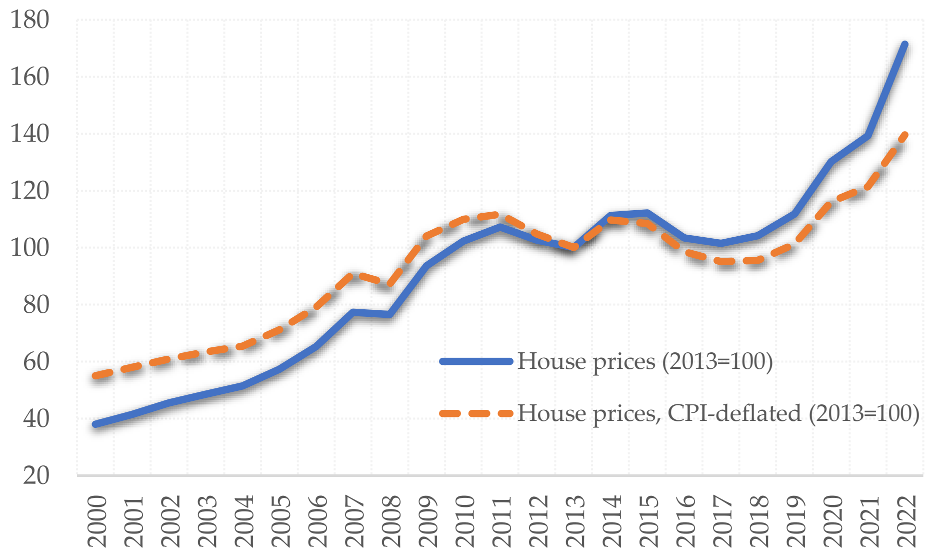


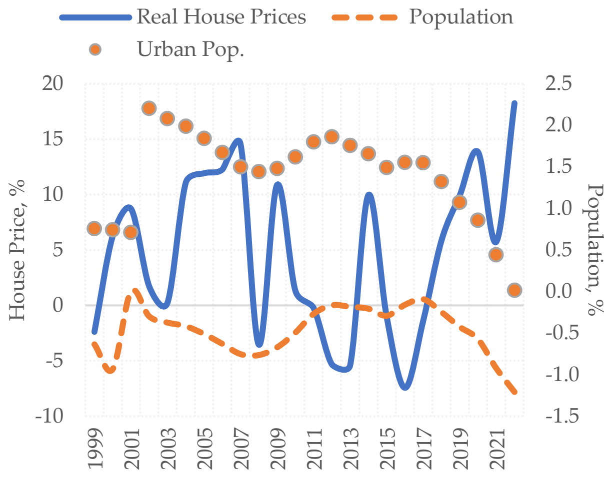
Figure 2. House price performance in Albania.

Table 1. Cumulative growth of house prices in the past two decades.

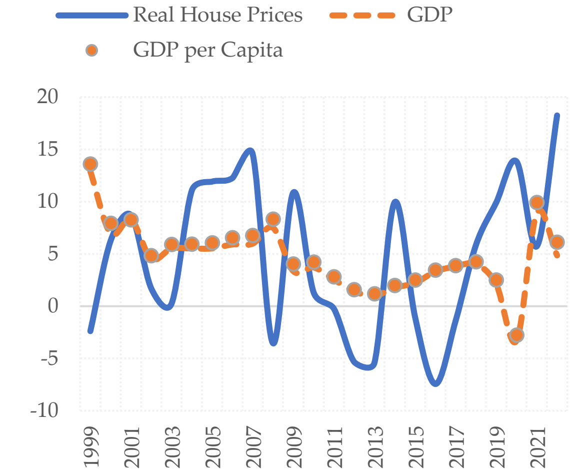
	Whole period 2000-2022	**Rapid HPI growth 2000-2009**	**Relatively calm period 2010-2017**	**Return of strong HPI growth 2018-2022**
HPI, cumulative growth (%).	401.7	146.5	1.9	76.3
HPI, st. dev. of yoy % changes.	8.4	7.9	6.7	8.3
CPI, cumulative growth (%).	77.8	30.1	14.8	12.7
CCI, cumulative growth (%).	6.0	-6.5	3.3	8.7
GDP, cumul. nom. growth (%)	454.9	181.1	39.1	41.9

A graphic analysis gives us an impression that the rapid growth of housing prices in recent years is related to the fundamental indicators (Figure 3). Despite the negative growth of the total population in recent decades, internal migration to urban areas in Albania has been constantly increasing. The favorable situation in the labor market, as implied by the decrease in the unemployment rate, as well as the wealth effect that is proxied here by the increase in GDP per capita, may have exerted pressure to raise housing prices. Non-resident interest in investing directly in real estate in Albania has grown significantly over the past five years, jumping from an average of 0.36% of GDP over 2013-17 to 1.59% in 2022. Similarly, the banking system appears to have further stimulated demand in the residential sector, as can be understood from the increase in the ratio of real estate loans to gross domestic product. In addition to lending, the continued decline in the real interest rate on new mortgage loans should have increased the affordability of housing purchases.

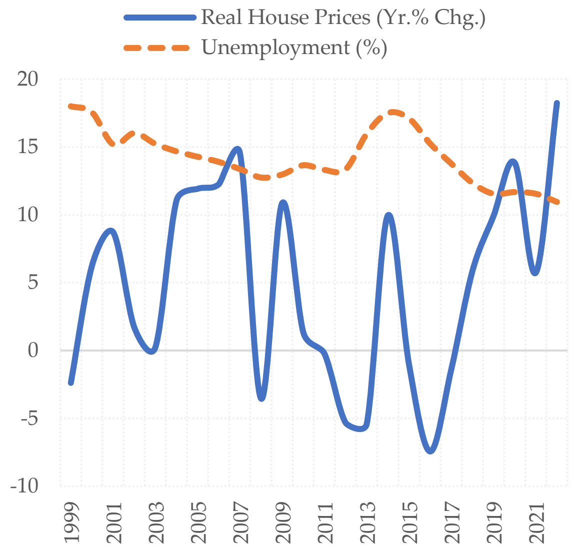
On the other hand, the influencing factors that affect the supply side such as costs and building permits for housing present a different picture. The construction cost index is constantly reported with a negative annual growth in real terms, and therefore incapable to influence the real price upsurge in the residential market. While the influence of construction permits is not very clear. The reduction of the area allowed for housing construction to only 3 square km during the years 2012-17 at a time when the urban population increased by over 159 thousand people may have created a synergy which has been somewhat reflected in the increase in housing prices in the subsequent period (see Table 2). However, the granting of building permits for dwellings measured in square km during 2018-22 should have eased the demand pressures resulting from the increase in the urban population, as is evident from the reduction of the ratio of these indicators to its lowest level since 2002.



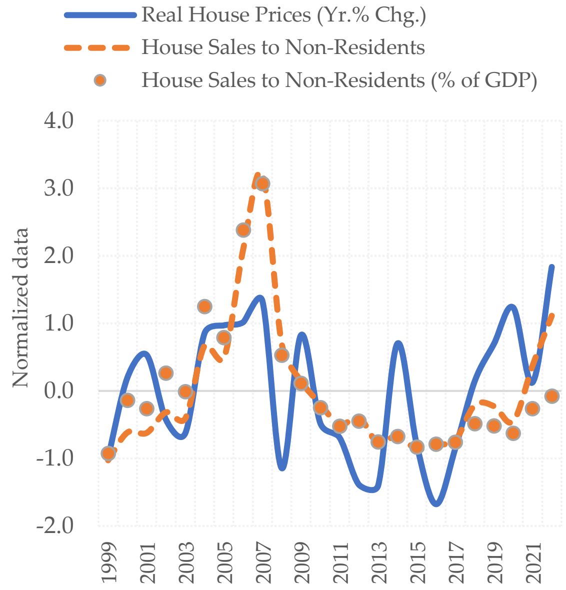
(a) House Prices and Population (YoY % change)



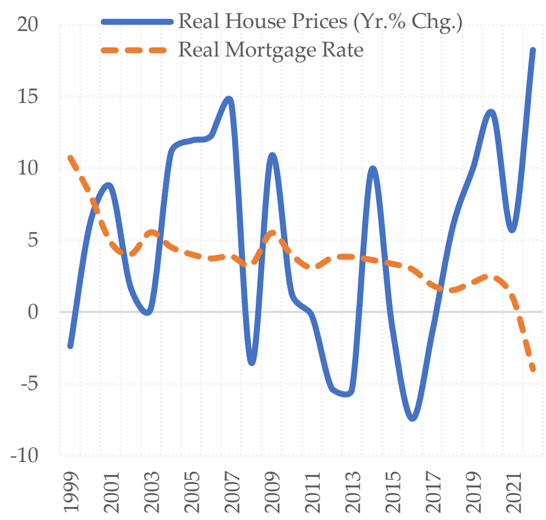
(b) House Prices and Income (YoY % change)



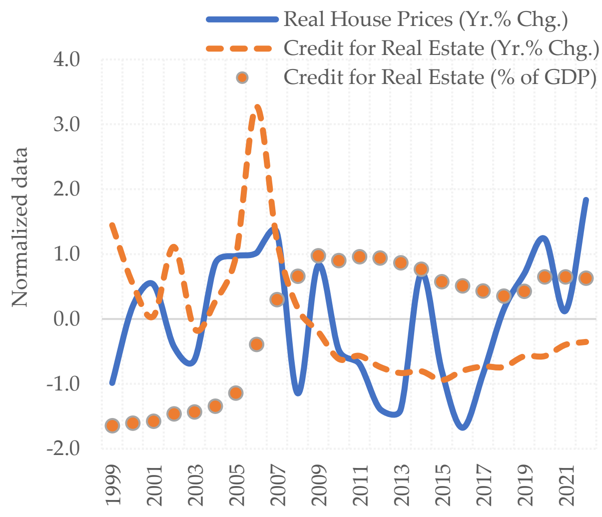
(c) House Prices and Unemployment Rate



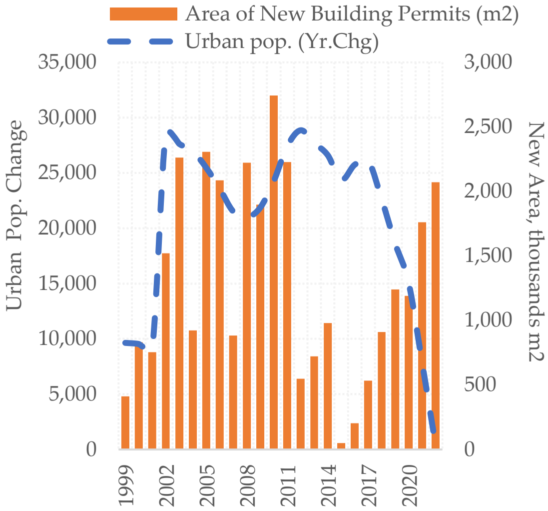
(d) House Prices and House Sales to Non-Residents



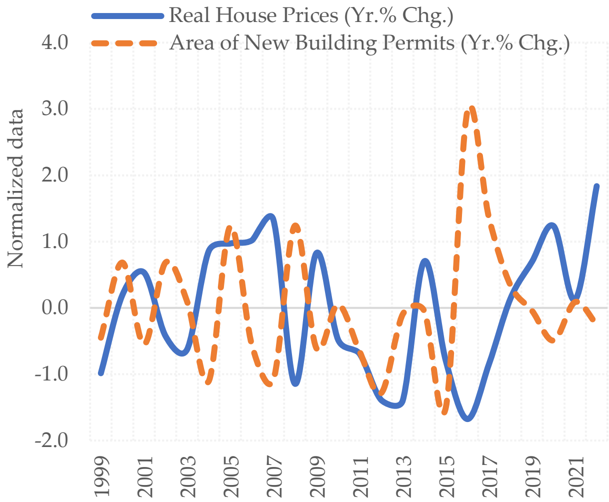
(e) House Prices and Mortgage Rate



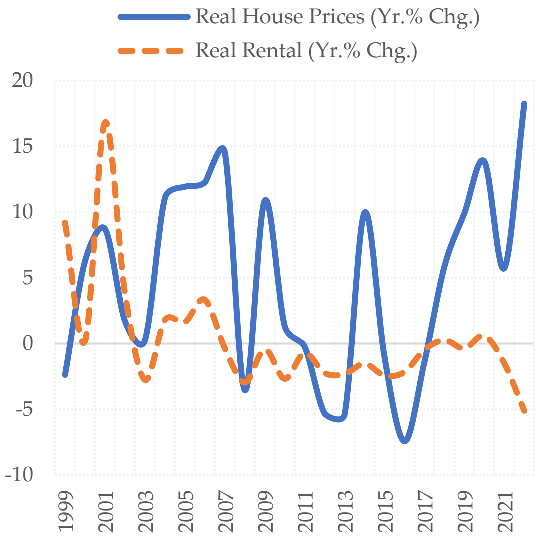
(f) House Prices and Credit for Real Estate



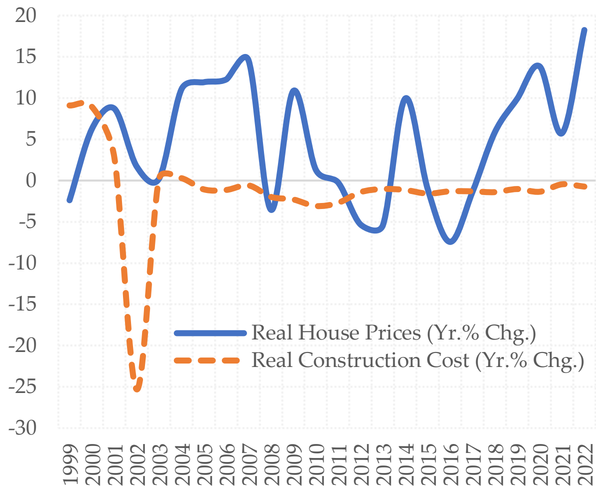
(g) House Demand and Supply



(h) House Prices and New Area for Build. Permits



(i) House Prices and Rentals



(j) House Prices and Construction Costs

Figure 3. Graphic analysis of house price determinants.

Table 2. Urban population growth per sq. km of house building area permits.

Years	**Urban population, cumulative growth (1)**	**Sum of new house building area permits, in km² (2)**	**Ratio of column (1) over (2) (3)**
2002-22	472,243	29.28	16,131
2002-12	277,204	19.62	14,130
2012-22	223,898	10.21	21,936
2012-17	159,416	3.04	52,526
2018-22	64,482	7.17	8,991

4. Data Issues

As there is no consensus on a properly defined multivariate model that could fully explain and predict fluctuations of house prices, macroeconomic performance should provide at least in the long run, appropriate information for the tendency of prices in the house market. The economic data that is relevant for analyzing movements in house prices in Albania is not always available at quarterly frequency for the full period under investigation. Thus, many of them had to undergo certain manipulation in terms of interpolation from annual frequency and/or earlier period extrapolation by using supportive indicators. Below is the list of variables that could be finally constructed at quarterly frequency for the 1998Q1:2022Q4 period:

- House price index. It is constructed as a combination of the indices provided by the Financial Stability Department (2013Q1-2022Q4) and the Monetary Policy Department (1998Q1-2013Q4) at the Bank of Albania.
- Gross domestic product. Data is available from the Institute of Statistics (Instat) at quarterly frequency from 2009. Quarterly data that goes back to 1998Q1 is provided by the Research Department, Bank of Albania.
- Unemployment rate. Data source is Instat for population age of 15 years old and above.
- Rent. Rental is an item in the consumer price index, downloaded from the Institute of Statistics.
- Construction cost index. Data is available from Instat since 2002Q1. Data for the earlier period is calculated from authors with vintage data provided from the Statistics Department, Bank of Albania.
- Commercial banks' real estate loans. This indicator is obtained from assets of the banking system, which are available in the Bank of Albania's website.
- Interest rate on new housing loans. The composite indicator is constructed by authors using the simple average of interest rates on lek-, and euro-denominated loans for house purchases by households. The data is only available at monthly frequency from Bank of Albania since December 2015. Before that, data is extrapolated by means of regressions on the 12M Treasury bill rate, interest rates on euro loans and the 3M Euribor rate.
- House sales to non-residents. Data on non-residents' direct investments in houses is only available since 2014Q1 from the balance of payments statistics at the Bank of Albania. To extrapolate direct investment in real estate data series, various OLS regressions were run by authors with sample periods 2014-2022 or 2008-2013, using (different lags of) variables such as remittances, deposits of non-residents, and total inflows of FDIs. Each forecasts back to 1994 were compared with, and selected on the basis of extrapolations that are found closer to the survey results by Gedeshi and Uruci [25], which reveals that about 18 percent of remittances are used for reconstruction of purchasing new houses.
- Urban population. It is available from World Development Indicators. Annual data is interpolated into quarterly frequency using constant match average method.
- Area of construction permits issued for new residential buildings. Data is available from Instat at quarterly frequency. Earlier annual data is interpolated using constant match sum method.

5. The LSTM Technique and Modeling Strategy

Neural networks are being increasingly used as powerful tools for predicting time series data. Over time there have been developed a number of techniques that address internal challenges within these networks. Some popular examples include the multilayer perceptron (MLP), recurrent neural networks (RNN), generalized regression neural networks (GRNN), and long short-term memory (LSTM). These evolving neural network techniques belong to the family of deep learning methods, and have recently found extensive application in forecasting time series.

The hidden layers of conventional artificial neural networks consist of basic neuron configurations that overlook context. In addition to them, RNNs incorporate a directed cycle that enables past events to capture important information for future outcomes. RNNs are predominantly utilized for temporally correlated data, resulting in a neural network model that makes use of past data and the correlation between lagged and current data to forecast future data. Nonetheless, RNNs suffer from vanishing gradients when there is a substantial temporal gap between the relevant information and its utilization point. To address this issue, the so-called LSTM technique introduces cell states, acting as a conveyer belt, to the hidden states of an RNN. Through recursive operations, LSTM preserves both cell states and hidden states, mitigating the problem of vanishing gradients.

As such, the LSTM method intends to enhance the long-term memory of RNNs by incorporating memory cells and gating mechanisms that control information flow. LSTM networks overcome the limitations of RNNs by storing information and reducing errors over extended periods. The fundamental structure of LSTM revolves around the network's ability to learn what to retain, discard what is considered unnecessary, and keep track of information over long time intervals. In this study, an LSTM model is employed to identify which of the fundamental indicators prove relevant for predicting house prices in Albania.

Table 3 displays the univariate and multivariate LSTM model specifications that are chosen in our analysis. The modeling forecasting procedure draws its variables from a pool of indicators that are found theoretically relevant in other empirical studies. A number of networks for the real house price function were used, where both real rental and real GDP appear as informative variables in almost all models. This basic network design is then augmented with a couple of the remaining variables, so the selection procedure may contain up to a handful of regressors that are considered sufficient to describe the most important dynamics given the limited number of observations. Following other studies, all variables enter the model in year-on-year growth rates, except for the real interest rate and the unemployment rate which are kept in levels.

The structure of neural networks is generally described by three "layers" and the number of "neurons", or nodes that transfer information among them. It starts with the input layer, whose nodes can be viewed as the model's explanatory variables. Information from input nodes is processed to the output layer (equivalently, the dependent variable) via a number of nodes in the hidden layer(s). Given the limitations of our data series, all LSTM model specifications take the dependent variable of real house prices in the current quarter to be a function of two input lags of all explanatory indicators including its own lags, thereby introducing two autoregressive terms.[1] The flexibility of the networks and model fit could be increased by adding more neurons and hidden layers (Ghysels and Marcellino [26]), therefore we consider two hidden layers with 200 and 100 nodes, respectively. There are a number of nonlinear functions needed to "activate" the hidden layers. We employ the Rectified Linear Unit (ReLU) function, which is found to be computationally efficient (Nair and Hinton [27]).

[1] A reduced-form model type with four autoregressive terms was tried, too, but they were found to provide inferior predictions when compared to models with two autoregressive terms. For sake of space the results are not shown here, but can be available from the authors on demand.

Table 3. LSTM network specifications.

Dependent variable	**Explanatory variables**				
		Fundamental indicators			
	(1)	(2)	(3)	(4)	(5)
HPI	LDV				
HPI	LDV	GDP			
HPI	LDV	YPC			
HPI	LDV	RENT			
HPI	LDV	RENT	GDP		
HPI	LDV	RENT	YPC		
HPI	LDV	RENT	YPC	CCI	
HPI	LDV	RENT	GDP	CCI	
HPI	LDV	RENT	GDP	BCRE	
HPI	LDV	RENT	GDP	RIR	
HPI	LDV	RENT	GDP	RIR	BCRE
HPI	LDV	RENT	GDP	HSNR	
HPI	LDV	RENT	GDP	UNR	
HPI	LDV	RENT	GDP	URB	
HPI	LDV	RENT	GDP	URB	CM2

Note: HPI = real house price index, CPI-deflated; LDV = lagged dependent variable; GDP = real GDP; YPC = real GDP per capita; RENT = real rental, CPI-deflated; CCI = construction cost index; BCRE = bank credit for real estate; RIR = real interest rate on new loans; HSNR = house sales to non-residents; UNR = unemployment rate; URB = urban population; CM2 = area of construction permits in thousand square meters.

The LSTM networks are trained by using the Adaptive Moment Estimation (Adam), which is an extension to the classical stochastic gradient-descent procedure used to update the network weights. Adam combines the advantages of certain popular extensions, and may in practice be preferred to other adaptive learning-method algorithms. One of the benefits of these optimization algorithms is that one does not need to tune the learning rate (Ruder [28]), therefore we count upon its default value of around 0.001. Many researchers monitor errors on a validation set during training and stop the parameter updating earlier if the forecast error does not improve enough. However, the parameters (biases and weights) in our analysis are updated by the Adam optimizer until the error in the training sample is reduced and becomes equal to the learning rate value, or a maximum of 400 epochs is reached.

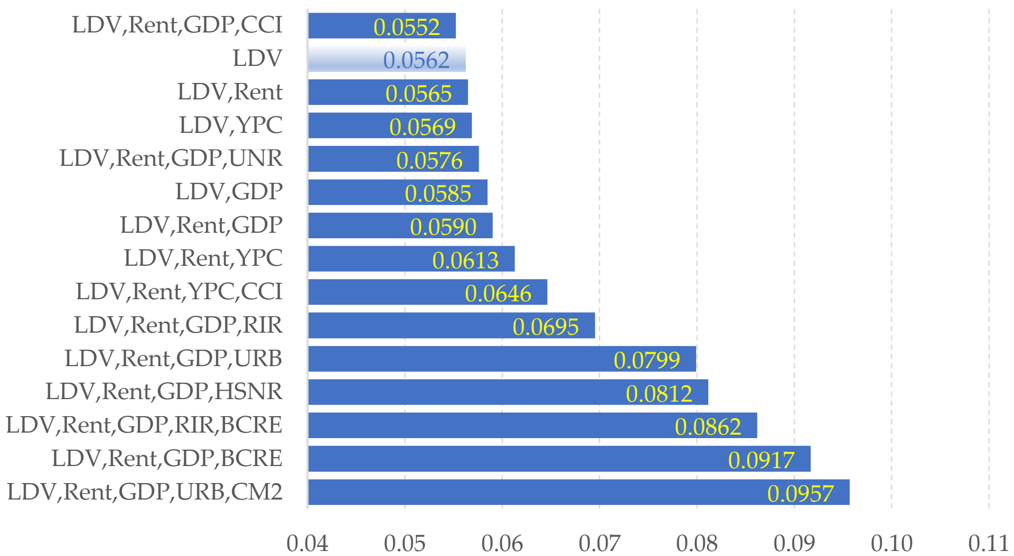
Finally, the link between house prices and the selected informative variables is tested in a pseudo out-of-sample period consisting of the last five years, which are excluded from the network learning process. More precisely, the (adjusted) period under investigation, which ranges from 1999Q1 to 2022Q4 has been divided into the so-called training and the validation period up to 2017Q4,[2] and the forecast evaluation period that runs from 2018Q1 through 2022Q4. The forecast evaluation is based on the smallest estimated forecast errors, as measured by the root mean square error (RMSE). For each of the LSTM specifications we note down its forecast performance for the 1, 4, 8 and 12 quarters ahead. Consequently, the evaluation process for every model is repeated 20 (17, 13 and 9) times for the 1 (4, 8 and 12) quarter(s) ahead forecast horizon, until the last horizon ending in 2022Q4 is reached.

[2] The validation dataset consists of 16 quarters.

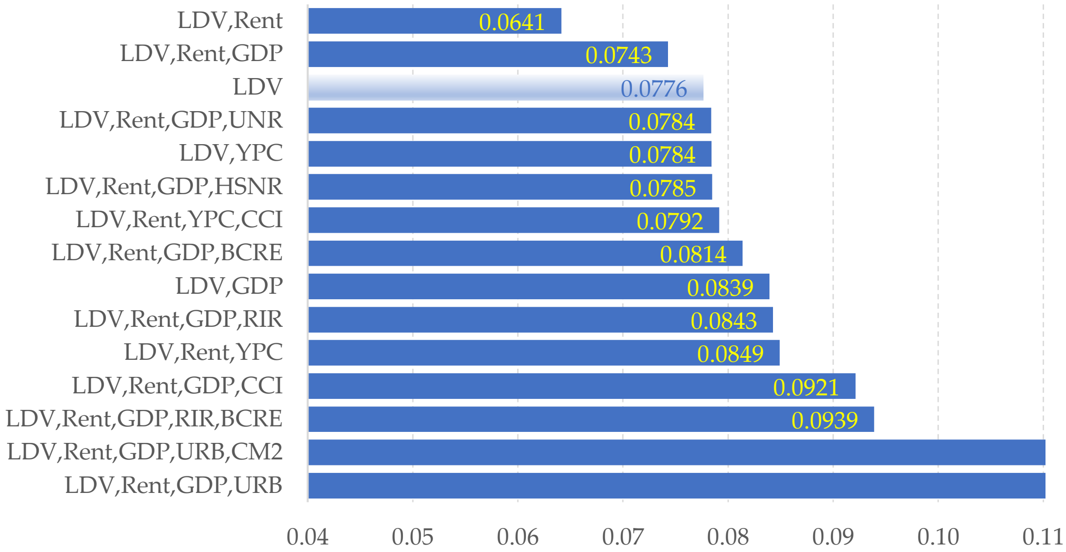
6. Empirical Findings

The power of deep neural networks stems from their freedom to choose a large number of parameters for training that can fit well different types of datasets. However, this power can turn into a weakness, if the model is "saturated" with a lot of parameters that can help it learn too many details in the training dataset but performs very poorly on the unseen or validation period. One way to check for the problem of overfitting is to monitor the evolution of predicted errors as we increase the number of epochs. If error values in the training period decrease significantly but errors in the validation period show substantial increase, overfitting needs to be taken into account. Measuring the model loss for all of our LSTM networks (not shown here) indicates that the size of forecast errors is always (considerably) descending on training data. On the other hand, forecast improvement in the validation dataset is clearly not as much of, and errors often incline upward at higher numbers of epochs. Nevertheless, it was resolved the uprise in errors' size is not large enough to make a case for reconsidering network configuration due to overfitting.

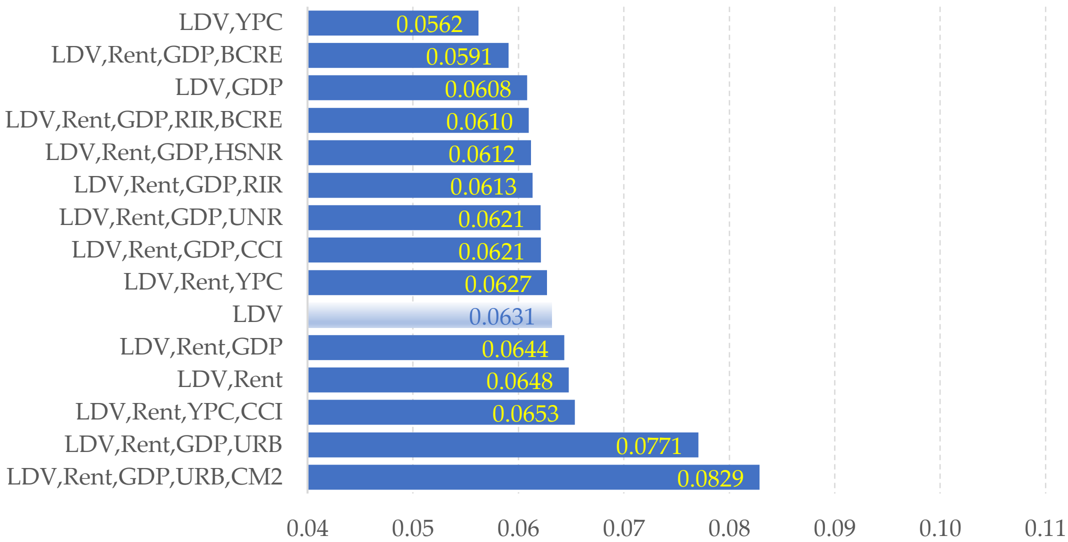
Turning to the focal point of the analysis, we now try to differentiate the economic indicators that can best serve as informative variables for predicting housing prices, particularly when compared with the univariate model performance. Figures 4a-4d compare the ability of all network designs to forecast in the short and the medium run during the whole out-of-sample period spanning from 2018Q1-2022Q4. Stimulatingly, the univariate model, which only relies on its own past values of house price growth, does not show as the best forecasting model in any of the forecast horizons. Its underperformance holds the attention that residential prices are indeed linked to, and have not moved in detached mode from the theory-driven economic indicators neither in the short, nor in the longer-term period. Another salient feature of the results is the rising number of explanatory variables that get involved in the top-ranking models as we aim for the medium-term forecasts. Real rental and GDP developments seem to provide relatively important information for predicting short-term movements in housing prices, as it is shown in the results for the 1 and 4 quarters ahead forecast horizons. Going for the two and three years ahead forecasts the group of relevant indicators gets augmented, particularly with the construction cost index and bank credit for real estate. Interestingly, urban population growth that captures the impact of evolving demographic factors falls in the group of worst forecast performers. In the same vein, area of residential building permits which may be used as a supply-side determinant seems highly ineffectual in the endeavor of beating the univariate network forecasts. The forecast errors (RMSEs) of models that contain these indicators are at least 50 percent higher than those of the univariate model.



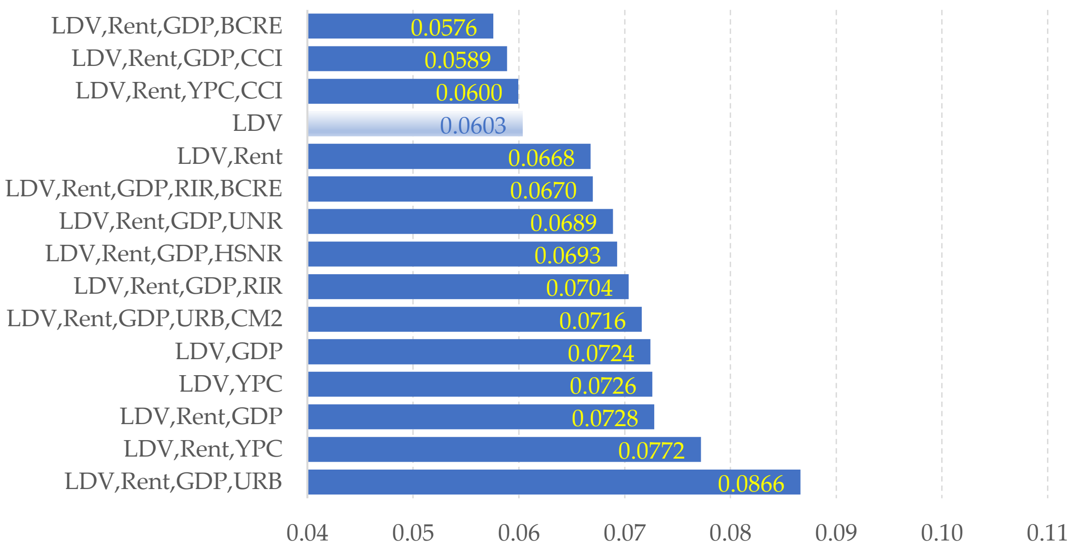
(a) RMSE ranking for the forecast horizon, h = 1.



(b) RMSE ranking for the forecast horizon, h = 4.



(c) RMSE ranking for the forecast horizon, h = 8.



(d) RMSE ranking for the forecast horizon, h = 12.

Figure 4. RMSE Ranking from Lowest to Largest (out-of-sample period, 2018Q1-2022Q4).

The annual growth of real house prices during our forecast evaluation period was not so even. It accelerated from 5.9 percent in 2018 to 13.9 percent during pandemic in 2020. Despite cooling down to 5.8 percent in 2021, it accelerated even faster to 18.2 percent in 2022, therefore resisting to the adverse shocks stemming from struggling global economic and geopolitical issues. For that reason, it is worthwhile to check whether the networks training and validation up to the end of 2017 hold and provide consistent results during the post-Covid lockdown period (2020Q3-2022Q4). Again, the forecast ability of multivariate models is clearly stronger than the univariate model.[3] Their superiority is more noticeable in the short run predictions (for horizons equal to 1 and 4 quarters), where the majority of fundamental indicators appear to have carried useful information for real house prices in the coming year. All the same, urban population and area of building permits fall in the worst ranking models, giving the impression of being unhelpful in our exercise.

The employed regularization method used on normalized data should help to impede overfitting by modifying the cost function in our LSTM neural networks. Yet, underperformance in the validation dataset could come from overly simplistic assumptions (resulting in biased forecast errors) and/or from model's sensitivity to fluctuations in the training set (resulting in high variance of errors). For that reason, we compute the bias and variance proportions of the mean squared forecast errors in our models. The former (latter) proportion indicates how far the mean (variance) of predictions is from that of actual data. Both proportions take values between 0 and 1. The smaller they are, the better the model performs in tracking the mean and variance of real house prices.

Table 4 displays the results for the univariate and multivariate networks in the testing dataset, spanning from 2018Q1 to 2022Q4. Generally speaking, all models seem to predict fairly well the mean of real house prices, particularly at longer horizons of 2-3 years. The bias proportion of the univariate model is 0.31 at h=1 and decreases to merely 0.02 at h=12. By and large, augmenting the network with economic fundamentals improves the mean forecasts at all horizons. Furthermore, all LSTM networks show satisfactory results in foretelling the variability of future house prices. The computed variance proportions appear very low in most cases, for both types of models. In sum, the average of mean (0.16) and variance (0.07) proportions for all models at all horizons only amount to 0.23 (out of 1), suggesting that more than three-fourth of the forecast error may be considered as irreducible error. This implies that the selected LSTM structures may suffer neither from underfitting (high bias), nor from overfitting (high variance). The statistical robustness of our models provides, thus, support to above findings that (certain) fundamental indicators are important in forecasting house prices in Albania.

[3] The results are not shown here, but can be put at the readers' disposal upon request.

Table 4. Measuring Bias and Variance Proportions of Mean Squared Forecast Errors.

	Bias proportion				Variance proportion			
Forecast horizons:	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12
LDV (Univariate)	**0.312**	**0.338**	**0.198**	**0.021**	**0.245**	**0.055**	**0.064**	**0.137**
LDV,GDP	0.308	0.389	0.102	0.016	0.144	0.001	0.094	0.033
LDV,YPC	0.305	0.374	0.164	0.000	0.188	0.012	0.220	0.031
LDV,Rent	0.075	0.086	0.180	0.063	0.179	0.001	0.103	0.047
LDV,Rent,YPC	0.128	0.226	0.011	0.008	0.075	0.000	0.061	0.005
LDV,Rent,GDP	0.118	0.130	0.016	0.030	0.141	0.003	0.009	0.017
LDV,Rent,GDP,CCI	0.008	0.393	0.134	0.009	0.184	0.004	0.107	0.147
LDV,Rent,YPC,CCI	0.129	0.314	0.052	0.010	0.232	0.030	0.048	0.142
LDV,Rent,GDP,BCRE	0.232	0.155	0.076	0.000	0.001	0.024	0.227	0.201
LDV,Rent,GDP,RIR	0.340	0.182	0.173	0.008	0.163	0.001	0.187	0.014
LDV,Rent,GDP,RIR,BCRE	0.244	0.250	0.003	0.000	0.010	0.009	0.070	0.095
LDV,Rent,GDP,HSNR	0.049	0.200	0.130	0.026	0.015	0.020	0.186	0.026
LDV,Rent,GDP,UNR	0.047	0.304	0.091	0.053	0.133	0.018	0.104	0.034
LDV,Rent,GDP,URB	0.396	0.492	0.329	0.065	0.090	0.008	0.012	0.000
LDV,Rent,GDP,URB,CM2	0.446	0.380	0.290	0.021	0.015	0.011	0.000	0.035

Note: LDV = lagged values of the dependent variable (real HPI); GDP = real GDP; YPC = real GDP per capita; RENT = real rental, CPI-deflated; CCI = construction cost index; BCRE = bank credit for real estate; RIR = real interest rate on new loans; HSNR = house sales to non-residents; UNR = unemployment rate; URB = urban population; CM2 = area of construction permits in thousand square meters. Shadowed cells indicate that multivariate models perform not as well as the univariate network in tracking the mean and variance of actual house prices.

7. Concluding Remarks

After a period of relative calm in the mid-2010s, the housing prices in Albania have exhibited a return to a steep rise for at least four consecutive years, including during the pandemic. The current phase has raised concerns about a possible change in direction as post-pandemic housing price surges face increasing risks of reversal amidst slower economic growth and higher interest rates. Although housing prices appear to be generally in line with its "sustainable" levels as indicated by the deviation of the ratio of house price to per capita GDP from its estimated long-term levels (Vika [29]), fear about a potential price correction necessitates a rigorous monitoring and understanding of the key factors influencing the risks of housing price decline. The prediction-based analysis in this article demonstrates an important role for fundamental indicators in predicting house prices in Albania. Their information content improves upon univariate model forecasts in the short-, as well as the longer-term horizons, albeit the combination of variables in addition to real rental and GDP growth could change over time. Although the Albanian financial sector at present appears relatively strong and resilient, government institutions should be attentive to the housing market trends and its related economic fundamentals in order to ensure the stability of both housing and finance.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, B.V. and I.V.; methodology, B.V.; software, B.V.; validation, B.B. and I.V.; formal analysis, B.V. and I.V.; investigation, B.V. and I.V.; resources, I.V.; data curation, I.V.; writing—original draft preparation, B.V. and I.V.; writing—review and editing, B.V. and I.V.; visualization, I.V.. All authors have read and agreed to the published version of the manuscript."

Funding: "This research received no external funding".

Acknowledgments: The authors are grateful to comments and suggestions received from participants during the presentation at the "Istanbul Annual International Multidisciplinary Conference on Economics, Business, Technology and Social Sciences - 2023", organized by the European Institute for Research and Development (EIRD), Istanbul, Türkiye, in May 13-14, 2023. Special thanks are for Mr Altin Tanku, Head of Research Department at Bank of Albania, for helpful discussions on house price determinants in Albania.

Conflicts of Interest: "The authors declare no conflict of interest. The institutions where the authors worked at the time of conducting this research article had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. Therefore, any views or errors that remain are the sole responsibility of authors."

References

1. Duca, J. V.; Muellbauer, J.; Murphy, A. (2021). What drives house prices cycles? International experience and policy issues. *Journal of Economic Literature* 59 (3): 773-864.
2. Tsatsaronis, K.; Zhu, H. (2004). What drives housing price dynamics: cross-country evidence. *BIS Quarterly Review*, March.
3. Egert, B.; Mihaljek, D. (2007). Determinants of house prices in Central and Eastern Europe. *Comparative Economic Studies*, Vol. 49, pp. 367-388.
4. Agnello, L.; Schuknecht, L. (2011). Booms and busts in housing markets: determinants and implications. *Journal of Housing Economics*, Vol. 20, pp. 171-190.
5. Cerruti, E.; Dagher, J.; Dell'Ariccia, G. (2015). Housing finance and real-estate booms: a cross-country perspective. *IMF Discussion Note* No. 15/12.
6. Yzeiraj, Endrit (2016). House prices in Albania: deviations from the equilibrium? *Bank of Albania working paper* 20 (59) 2016.
7. Suljoti, Erjona (2017). The implications of financial conditions in the housing market – a cointegrated analysis. *Bank of Albania Working Paper* 27 (66) 2017.
8. Kraja, I.; Sejdini, M.; Sejdini, A. (2018). Tregu i banesave në Shqipëri: regresioni hedonik. *INSTAT Revista e Statistikave dhe Analizave Social Ekonomike* Nr. 1/2018. Available online: https://www.instat.gov.al/media/4268/4-the-housing-market-in-albania.pdf (accessed on 30 April 2023).
9. Marku, E.; Lleshaj, Ll.; Lleshaj, A. (2020). VECM analysis of house price index. Case of Tirana. *European Journal of Marketing and Economics*, Vol. 3, Issue 1, January-June 2020.
10. Lleshaj, Ll.; Korbi, A. (2019). VAR model for short-run endogenous factors in real estate market: case of Albania. *American Research Journal Of Humanities Social Science*, Vol. 02, Issue 10, pp. 29-36, October 2019.
11. Koprencka, L. K.; Bundo, Sh.; Qarri, A. (2014). Factorial analysis of Albanian housing market. Acta Universitatis Danubius, *Economica*, Vol 10, No 4 (2014).
12. Halili, F. (2022). The Impact of Macroeconomic Factors on the Change of Residential Prices – The Case Study of Albania. *International Journal of Applied Research in Management and Economics*, 5(4): 29-44.
13. Mora-Garcia, R.-T.; Cespedes-Lopez, M.-F.; Perez-Sanchez, V.R. (2022). Housing Price Prediction Using Machine Learning Algorithms in COVID-19 Times. *Land* 2022, 11, 2100.
14. Wang, D.; Li, V.J. (2019). Mass appraisal models of real estate in the 21st century: a systematic literature review. *Sustainability* 2019, 11 7006.
15. Park, B.; Bae, J.K. (2015). Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data. *Expert Syst. Appl.* 2015, 42, 2928-2934.
16. Banerjee, D.; Dutta, S. (2017). Predicting the housing price direction using machine learning techniques. Proceedings of the 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 21–22 September 2017; pp. 2998–3000.
17. Kok, N.; Koponen, E.-L.; Martínez-Barbosa, C.A. (2017). Big Data in Real Estate? From Manual Appraisal to Automated Valuation. *J. Portf. Manag.* 2017, 43, 202–211.
18. Ceh, M.; Kilibarda, M.; Lisec, A.; Bajat, B. (2018). Estimating the Performance of Random Forest versus Multiple Regression for Predicting Prices of the Apartments. *ISPRS Int. J. Geo-Inf.* 2018, 7, 168.

19. Fan, C.; Cui, Z.; Zhong, X. (2018). House Prices Prediction with Machine Learning Algorithms. Proceedings of the 2018 10th International Conference on Machine Learning and Computing, Macau, China, 26–28 February 2018; pp. 6–10.
20. Ho, W.K.O.; Tang, B.-S.; Wong, S.W. (2021). Predicting property prices with machine learning algorithms. *J. Prop. Res.* 2021, 38, 48–70.
21. Chatzidis, I.V. (2019). Prediction of Housing Prices based on Spatial & Social Parameters using Regression & Deep Learning Methods. Master's Thesis, University of Thessaloniki, Thessaloniki, Greece, 2019.
22. Alfaro-Navarro, J.-L.; Cano, E.L.; Alfaro-Cortés, E.; García, N.; Gámez, M.; Larraz, B. (2020). A Fully Automated Adjustment of Ensemble Methods in Machine Learning for Modeling Complex Real Estate Systems. *Complexity* 2020, 2020, 5287263.
23. Hong, J. (2020). An Application of XGBoost, LightGBM, CatBoost Algorithms on House Price Appraisal System. *Hous. Financ. Res.* 2020, 4, 33–64.
24. Hacıevliyagil, N.; Drachal, K.; Eksi, I.H. (2022). Predicting House Prices Using DMA Method: Evidence from Turkey. *Economies* 10: 64. Available online: https://doi.org/10.3390/economies10030064 .
25. Gedeshi, I. (2002). The role of remittances from Albanian emigrants and their influence in the country's economy. In Kotios and Petrakos (eds.) Restructuring and Development in Southeastern Europe, University of Thessaly Press, SEED, 2002.
26. Ghysels, E.; Marcellino, M. (2018). Applied Economic Forecasting Using Time Series Methods. Book published by Oxford University Press, 2018.
27. Nair, V.; Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807-814.
28. Ruder, S. (2017). An overview of gradient descent optimization algorithms. arXiv:1609.04747v2 [cs.LG] 15 Jun 2017.
29. Vika, I. (2023). Whither house prices in Albania. *Bank of Albania*, Unpublished research document, May, 2023.