BANK OF ALBANIA

Economic Review

2022 H2

Published by: Bank of Albania,
Sheshi "Skënderbej", nr. 1, Tiranë, Shqipëri
Tel.: + 355 4 2419301/2/3; + 355 4 2419409/10/11
Fax: + 355 4 2419408
E-mail: public@bankofalbania.org
www.bankofalbania.org

www.bankofalbania.org

** Views expressed in this review are of the authors and do not necessarily reflect those of the Bank of Albania.*

C O N T E N T

USING MACHINE LEARNING TOOLS TO STUDY THE UNEMPLOYMENT AND OUTPUT RELATIONSHIP IN ALBANIA

Blerina VIKA[1], Ilir VIKA[2], Kozeta SEVRANI[3]

ABSTRACT

The field of economics is nowadays increasingly employing artificial intelligence to complement and further improve the tools for analyzing and making future decisions. One of the most common machine-learning techniques that are used in forecasting economic indicators is the recurrent neural network method, which has often proven to be useful in capturing non-linearities in data series. This article applies the long short-term memory (LSTM) technique to test for the Okun relationship in Albania. Apart from examining movements in unemployment as predicted by developments in aggregate demand, we test whether a disaggregated version of the Okun's law – by decomposing aggregate demand into various expenditure components of GDP – provides better predictions for changes in unemployment. In-sample estimations suggest that the Okun's law may hold in Albania, but the response of unemployment to output performance is found to vary over different time periods. Non-linear model forecast evaluations show that unemployment rate movements in the second half of 2010s could be more related to private investment and government spending, while private consumption and external trade developments seem to predict it less.

Keywords: Machine Learning, LSTM model, unemployment, Okun relationship, Albania.

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 "11th International Conference on Business, Technology and Innovation 2022", organized by the University for Business and Technology - UBT, Pristina, Kosovo, in October 29-30, 2022.

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

1. INTRODUCTION

There is a general consensus today that inequitable economic growth might have adverse effects in the society. For that reason, governments around the globe have increased their awareness on the importance of inclusive growth and are pushing their political agenda to make sure that economic growth could deliver more jobs and promote higher living standards for their citizens. Emerging and developing economies have been growing rapidly for a number of decades. Yet, recent empirical findings reveal that their satisfactory performance has been in the large part not necessarily translated into more employment and equal opportunities.

This paper uses the Okun relationship to examine movements in unemployment as predicted by developments in aggregate demand in Albania. In addition, it test whether a disaggregated version of the Okun's law – by decomposing aggregate demand into various expenditure components of GDP – provides better predictions for changes in unemployment. The hypotheses are tested by using the econometric least square method as well as a rather novel technique in this regard by employing the long short-term memory (LSTM) model. The latter is a machine-learning technique that has often proven to be useful in capturing non-linearities in economic data series.

To preview the results, the OLS estimations appear to be in favor of the Okun relationship in Albania. Estimating different sub-samples suggests that the response of unemployment to output could be time-varying, while faster domestic growth may be needed to keep unemployment rate from rising. On the other hand, the preliminary results in this ongoing project show that adding output growth to the LSTM forecast model does not improve the unemployment predictions that are generated from a univariate model. Nevertheless, the LSTM model predictions with GDP components reveal that unemployment could be more related to private investment and government spending, and perhaps not so much linked with private consumption and external trade.

The remainder of the article is organized as follows. In section 2 we provide a glimpse at the empirical literature on unemployment and growth in advanced and developing countries. Section 3 estimates the Okun coefficient in a simple OLS regression to examine the contemporaneous relation of unemployment with aggregate demand in Albania. In addition, to shed light on the stability of their relationship, the Okun coefficient is estimated for i) the early transition years, ii) the faster growth period, and iii) the post-global crisis decade. Section 4, then, reassesses the relevance of economic growth in the Okun relationship by relying on the information embedded in GDP and its expenditure components to improve upon the benchmark univariate forecasts of the unemployment rate. Moreover, as the in-sample estimations in section 3 evidenced a time-varying Okun coefficient, the uncommon cross-checking analysis in section 4 is based on the nonlinear LSTM model. The concluding remarks are available in section 5.

2. A BRIEF LITERATURE REVIEW

Despite some theoretical limitations of Okun's law, Prachowny (1993) and Blinder (1997) conclude that Okun relationship is a useful principle of macroeconomics in which 'we should all believe'. In his seminal paper on the relationship between GDP developments and changes in unemployment, Okun (1962) originally treated unemployment as the exogenous variable while output as the dependent variable. However, the rich empirical literature estimating the Okun relationship has been mostly interested in the reverse direction, hence assuming the output growth as the driving force of unemployment.

The link between unemployment and output is generally assessed by transforming the variables in gaps or in differences. The "Gap" version relates the divergence of current unemployment (U_t) fromt its "natural" rate (U^*) with the deviation of actual output (Y_t) from its potential level (Y^*) $[U_t - U^* = \beta(Y_t - Y^*) + \varepsilon_t]$. On the other side, "Difference" version relates changes in unemployment rate to economic growth $[\Delta U_t = \alpha + \beta \Delta Y_t + \varepsilon_t]$, thus assuming that the "natural" rate of unemployment is zero $[\Delta U^*=0]$ while potential GDP growth is constant $[\Delta Y^*=g]$.

The original Okun's paper estimated that β equaled negative 0.3 in the U.S. during the 1947-60 period. However, the recent empirical literature on the Okun's law has found heterogeneous evidence with respect to the size of the Okun coefficient. The difference among countries might have much to do with the very definition of unemployment, as the term does not follow a worldwide standard definition of measuring employment, nor can data availability be comparable between formal and shadow economies. Also, labor market rigidities and government social protection policies on employment have often been emphasized when trying to explain the heterogeneous sensitivity of unemployment to output decline, particularly among advanced economies [see e.g. IMF, 2010; Cazes, Verick, & Hussami, 2013].

A review by Pizzo (2019) shows Okun's coefficients in rich economies as estimated with the difference (gap) version are around -0.29 (-0.39), while in developing and emerging market economies they are -0.18 (-0.20). Other studies have similarly concluded that unemployment is less sensitive to production in developing than in advanced economies [see e.g., Ball et al., 2019; An et al., 2017; Bartolucci et al., 2018]. Using a sample of 176 countries for the period 1993 until 2015, Farole et al. (2017) evidence that Okun's coefficient is not constant across countries. For every percentage point increase in GDP, unemployment rates would decline on average for the high income, upper-middle income, lower middle-income, lower income countries by 0.21, 0.08, 0.03, and 0.005 pp, respectively.

Furthermore, reassessments after the Global Crisis period indicate that Okun coefficients all over the world have declined when comparing the pre-crisis (1992-2007) with the post-crisis (2010-2017) period (Lee et al., 2020). Yet, Farole et al. (2017) find that in transition economies like Ukraine and Croatia unemployment seems to react asymmetrically over the business cycle:

the size of coefficients show relatively strong reaction during downturns, but much weaker upside response during periods of upward growth.

Using only the "difference" version and running pooled OLS for ILO data sets on unemployment, Lee et al. (2020) find that during the 1992-2017 period the Okun's coefficients are -0.22 in twenty nine NSW European countries and -0.15 in ten Eastern European economies. Anyhow, Farole et al. (2017) find a large difference among similar income countries in CE Europe. For example, the Okun coefficient is estimated 0.85 in Poland, but just 0.03 in Hungary. In Albania, Garo (2020) finds that "for a 1 p.p. drop in the output growth, the annual change in the unemployment rate increases about 0.184 p.p. over 4 consecutive quarters".

3. DOES OKUN'S LAW HOLD IN ALBANIA? AN OLS APPROACH WITH AGGREGATE DEMAND

The link between unemployment and output is typically examined by assessing their contemporaneous relationship:

$$\Delta U_t = \alpha + \beta_y \Delta Y_t + \varepsilon_t \qquad (1)$$

where ΔU denotes annual change in unemployment rate; ΔY is annual growth rate of output; α is an intercept coefficient that captures the trend growth in unemployment rate ($\alpha = -\beta \Delta Y^*$); β is the so-called "Okun coefficient", denoting the sensitivity of unemployment rate to GDP growth, which is expected to be negative, since a higher GDP growth rate should lead to lower unemployment rate; ε is an error term.

We test for the Okun relationship using quarterly data from 1996q1 to 2019q4. As Albanian economy has experienced significant fluctuations during this period, it may be possible that the correspondence between variables be both nonlinear and time-varying. Thus, we test for changes in the Okun relationship during the i) early transition period of 1996q1-2003q4; ii) fast economic growth from 2003q1-2010q4; and the post-global financial crisis decade from 2011q1-2019q4.

Table 1. Okun coefficients for Albania: aggregate relationship

	(1) Full Sample 1996q1:2019q4		(2) Early Transition 1996q1:2003q4		(3) Fast Growth 2003q1:2010q4		(4) Post-Global Crisis 2011q1:2019q4	
Dependent Var. ΔU	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.
ΔY (βy)	-0.090***	0.00	-0.075*	0.10	-0.212***	0.00	-0.581***	0.00
Constant (α)	0.236	0.23	0.436	0.28	0.820***	0.00	1.182***	0.01
Adj. R-squared	0.07		0.06		0.54		0.32	
Included observations	96		32		32		36	
Ratio of $-\alpha/\beta$	2.6		5.8		3.9		2.0	

Table 1 displays the estimated Okun coefficient for the aggregate relationship in Albania. The estimated response of unemployment to output growth is statistically significant and has a negative sign as expected, hence suggesting a well-functioning labor market. The Okun coefficient, β, is, however, not constant and appears to have considerably increased in magnitude over time. The relationship is estimated to be modest in the full sample (-0.09). The result is in line with the findings from Farole et al. (2017) for middle-income countries, which are often characterized by insufficient safety nets and relatively large share of self-employed workers. However, the relatively low reaction seems to be influenced by the early transition period. Later on, the size of β coefficient increases significantly and implies a relatively significant job creation in the post-GFC period (-0.58). But, so has the trend growth in unemployment rate as captured by the intercept parameter, α; in the last period in column (4) is about 5 times higher than for the full sample, hence well above the assumption of zero changes in the "natural" rate of unemployment. Finally, the time-varying parameters are reflected in a declining ratio of model coefficients ($-\alpha/\beta$) – found to be statistically significant since 2003 – which may imply that the Albanian economy needs to grow faster in order to leave the unemployment rate unchanged.

4. DECOMPOSING THE OKUN'S LAW: DO EXPENDITURE COMPONENTS ENHANCE PREDICTION ABILITY?

METHODOLOGY

The vast empirical literature on the Okun relationship evidences that economists widely believe on the existence of a long-run relationship between GDP performance and unemployment. Yet, the short-run dynamics may disguise the long-run bond between the variables. The previous section revealed a rather time-varying Okun relationship in Albania, suggesting a lack of stability in the coefficients over the past two and a half decades.

To shed more light on the suitability of Okun's law as a "rule of thumb" for predicting unemployment rate, certain studies have resorted to analyses that decompose the Okun relationship into expenditure or production components. We follow them by decomposing here the aggregate demand into private consumption (con), public expenditure (gov), private investment (inv), exports (exp) and imports (imp).

Furthermore, the mainstream studies have commonly relied on traditional econometric models to estimate and test the stability of the Okun coefficient. In contrast, we follow a novel approach to test for the Okun relationship by relying on unemployment predictions generated by nonlinear nonparametric Machine Learning (ML) methods. These innovative techniques could be suitable in resolving possible issues of nonlinearity and time-varying relations in our variables, due to changing domestic and external economic conditions. Among the wide class of these techniques, we have selected to apply in our

analysis the long short-term memory (LSTM) model. This method belongs to the artificial recurrent neural network family, which have gained popularity in the area of artificial intelligence and deep learning. Moreover, machine learning methods are nowadays being seen as a new exciting 'car' on the street by many central bank economists, who are increasingly experimenting with ML techniques as additional supporting tools to help improve economic forecasts and monetary policy decisions. Unlike the feedforward neural network that we have experimented before (e.g., Vika & Vika, 2021), the recurrent neural network LSTM has feedback connections and is capable of learning long-term dependencies. Once we train and test the LSTM network structure, we use the out-of-sample forecasts computed by it[4] to assess the relevance of economic growth in the Okun relationship by relying on the information embedded in GDP and its expenditure components to improve upon the benchmark univariate forecasts of the unemployment rate.

DATA AND FORECASTING PROCEDURE

All variables are seasonally adjusted by using the X-13 ARIMA-SEATS procedure in EViews software. The whole sample period covers quarterly data from 1996Q1 to 2019Q4. Each model includes three lags as inputs of each variable, as suggested by Akaike and HQ info criteria in the estimated least squared regressions. Forecast evaluation is undertaken for out-of-sample forecasts, using the RMSE measure for a forecast horizon of 4 quarters. In our analysis, the 1996Q1:2019Q4 period has been divided into the so-called training period 1996Q1:2014Q4 (76 quarters) and the forecast evaluation period stretching over the 2015Q1:2019Q4 sample (20 quarters).

Consequently, the model starts for the period 1996Q1:2014Q4. For each specification the forecast and its corresponding RMSE is noted down for the 4 quarters ahead. The training and testing period is recursively extended by one quarter (1996Q1:2015Q1), and similarly calculating and retaining the RMSE of forecasts for the desired one-year ahead horizon. The evaluation process is repeated 16 times for the 4 quarters horizon until we predict the last quarter of 2019.

FORECAST EVALUATION RESULTS

Table 2 displays the forecast gains/losses in using GDP growth and its expenditure components to predict changes in the unemployment rate. A ratio above 1 indicates underperformance of the bivariate LSTM networks. It turns out that including GDP in the unemployment LSTM model does not outperform the latter's forecasts that are based on its own past values. The ratio of RMSE is just above 1, casting doubt on the predictive content of output growth on unemployment rate. However, the evaluation of loss differentials for individual expenditure components reveals some interesting results on the effectiveness

[4] *The model forecasts are generated by using the Python software.*

of this approach with the aggregate relationship. The RMSE ratios for government spending (0.84) and private investment (0.81) suggest us that these expenditure components could be the key factors in delivering jobs and relieving unemployment in Albania. On the other hand, private consumption (1.15) and external trade (1.07) seem to have no predictive power on the rate of unemployment.

Table 2. The loss differential between forecasts, 2015q1:2019q4, h=4 quarters

Bivariate Models:	Δu-Δy	Δu-Δcon	Δu-Δgov	Δu-Δinv	Δu-Δexp	ΔU-Δimp
Ratio of average RMSE: (Bivariate/Univariate f'casts)	1.01	1.15	0.84	0.81	1.07	1.07

5. CONCLUDING REMARKS

The results from OLS estimations seem to be in favor of the Okun relationship in Albania. The coefficient is found to be rather low for the past two and a half decades, in line with other studies for low and middle income countries. Estimating different sub-samples shows considerable lack of constancy in the coefficients that measure the sensitivity of unemployment to output as well as the trend growth of unemployment. Yet, faster domestic growth may be needed to keep unemployment rate from rising. Moreover, LSTM forecasts indicate that unemployment could be more related to private investment and government spending, while private consumption and external trade developments do not seem to improve upon the univariate forecast accuracy.

While social demographic changes, informality and emigration behavior may limit the ability to increase the precision of measuring unemployment, inadequate safety nets and the relatively high level of self-employed persons in the country might contribute to the underestimation of output-unemployment relationship.

Future research may need to focus on alternative approaches to assess the Okun relationship with various model specifications and lag length selections, along with variable transformation such as other ratios of labor market indicators or the "gaps" version. The latter, for instance, could be invaluable in taking into account possible changes in the "natural" rate of unemployment and potential output growth, as the Albanian economy strives to catch-up with other higher per capita income countries.

REFERENCES

An, Z., Ghazi, T. & Prieto, N.G. (2017) "Growth and jobs in developing economies: trends and cycles" IMF Working Paper WP/17/257.

Ball, L., Furceri, D., Leigh, D., & Loungani, P. (2019) "Does one law fit all? Cross-country evidence on Okun's law" Open Economies Review 30, 841–874 (2019).

Bartolucci, F., Choudhry, M.T., Marelli, E. & Signorelli, M. (2018) "GDP dynamics and unemployment changes in developed and developing countries" Applied Economics, 50 (31), 3338-3356.

Blinder, Alan S. (1997) "Is there a core of practical macroeconomics that we should all believe?" The American Economic Review, vol. 87, no. 2, 1997, pp. 240-243.

Cazes, S., Verick, S. & Hussami, F.A. (2013) "Why did unemployment respond so differently to the global financial crisis across countries? Insights from Okun's law" IZA Journal of Labor Policy, Springer; Forschungsinstitut zur Zukunft der Arbeit GmbH, vol. 2 (1), pp 1-18, December.

Farole, T., Ferro, E. & Gutierrez, V.M. (2017) "Job creation in the private sector: an exploratory assessment of patterns and determinants at the macro, sector, and firm levels" Jobs Working Paper, no. 5. World Bank, Washington D.C.

Garo, O. (2020) "Inquiring into the relationship between the unemployment rate and output growth", Paper presented at Bank of Albania, 14th South-Eastern European Economic Research Workshop, Tirana, 10-11 December 2020.

IMF (2010) "World Economic Outlook. Chapter 3"

Lee, S., Schmidt-Klau, D., Weiss, J., Chacaltana, J. (2020) "Does economic growth deliver jobs? Revisiting Okun's Law" ILO Working Paper 17, November / 2020.

Okun, A. (1962) "Potential GDP: its measurement and significance" Proceedings of the Business and Economic Statistics Section of the American Statistical Association.

Pizzo, A. (2019) "Literature review of empirical studies on Okun's law in Latin America and the Caribbean" ILO Employment Policy Department Working Paper No. 252, 2019.

Prachowny, M.J.F. (1993) "Okun's law: Theoretical foundations and revised estimates" The Review of Economics and Statistics, vol. 75, no. 2, 1993, pp. 331-336.

Vika, B., Vika, I. (2021) "Forecasting Albanian time series with linear and nonlinear univariate models" Academic Journal of Interdisciplinary Studies, Richtmann Publishing, vol. 10, no. 5, September 2021.

REAL TIME OBTAINED REAL ESTATE PRICES: THE WEB SCRAPPING METHODOLOGY

Orion Garo, Research Department, Bank of Albania

INTODUCTION AND BACKGROUND OVERVIEW

The real estate market is one of the cornerstone markets of the economy and residential property prices (hereafter, house prices) are a key indicator of the domestic consumer dynamics and propensities. Yet, periodic data on this indicator is quite scarce in Albania. The only official source providing a house price index is Bank of Albania's semiannual survey "The Real Estate Market". In this case the data are obtained using questionnaires filled by construction and real estate trade companies making up a representative sample of this sector's population, and are later aggregated to compute the semiannual Fischer price indices for Tirana and the whole country. With the aforementioned exception, neither the Institute of Statistics nor any other governmental body has been involved in gathering, compiling and aggregating this data for official statistics.

Under the conditions of a necessity for alternate channels of housing market data accessibility, online platforms advertising residential properties for sale may be very useful to acquire this type of information, along with its periodicity. They enable users (house owners, or the real estate trade companies representing the latter) to list their properties for sale on the Web, along with the corresponding details such as price, surface area in meters squared, architectural setup, etc. User generated content (UGC) attributed to the housing market constitutes a small portion of the immense volume of Internet-found information better known as "big data", but it leads to tremendous opportunities in the field of research. This is because UGC's data entry frequency, granularity and detailed descriptions can help to compile datasets that allow for sound structural econometric analyses and real time price dynamics monitoring.

The set objective of this research article is to identify the per-meter-squared house price average standings and dynamics during the investigated time period, for three geographical classifications: [I] Tirana; [II] coastal, and; [III] inland areas of the country. This is accomplished through the acquisition of a representative sample of for sale houses price listings countrywide, based on Python's web scrapping technique, and data congregation and organization through MS Excel and MS Office's Power toolset.

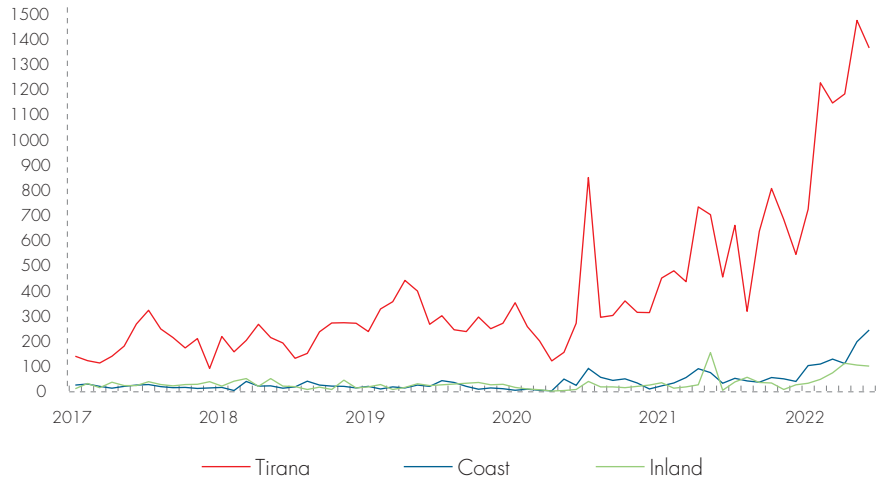
The remainder of the article is structured as follows. The second section covers the research data, the methodology, and describes the tools used to achieve data congregation and organization for the purposes of this research. The fourth sections details the research results, and the last section presents its concluding remarks. The bibliography is given at the end of the paper.

DATA AND METHODOLOGICAL FRAMEWORK

This research article's data is sourced from njoftime.com which, to the best of our knowledge, to date, is Albania's largest UGC transactions information website. The sample size consists of 30,467 observations (house sale ad entries) published over a time span of more than 5 years: January 2017 to June 2022. To refine the sample's representation attributes, we have only included apartments of residential buildings and single to five story houses and villas located in rural and urban areas of the country. The study omits listings of premises used to conduct activities unrelated to residential accommodation, such as business, social, industrial and agricultural activities. Listings of residential properties sold in conjunction with sizable areas of land or other natural assets are also omitted.

Referring to the methodology used in Bank of Albania's "Real Estate Market" semiannual questionnaire, conducted by its Financial Stability Department, the available data was divided into 3 groups: [I] "Tirana" for all entries of advertised sales within a radius of 5 km from the center of the capital;[1] [II] "coast" for all entries of advertised sales in Albania's coastal[2] cities, towns and rural areas, and; [III] "inland" for all entries of advertised sales in Albania's inner[3] cities, towns and rural areas. The sample's proportionality is such that 86% of the ads are for property sales in Tirana, 9% are for property sales in coast areas, and only 5% are for sales in inland areas. The analysis' time frequency is set to be monthly since there is observational abundance to support this frequency; there are approximately 557 observations per month. Figure 1 shows a chart of the sample's number of observations, distributed in time:

Figura 1. Smaple's proportionality of observations by geographical groups



[1] *This limit is set because at a distance of a little over 5 km from the city center, semiurban and rural dwelling localities of Tirana district – such as Babrru, Paskuqan, Kashar etc – are found, and the houses for sale in these localities belong to the inland areas group.*

[2] *The "coast" group includes districts: Durrës, Vlorë, Sarandë, Himarë; towns: Shëngjin, Golem, Velipojë and Divjakë; as well as all pertaining coastal villages.*

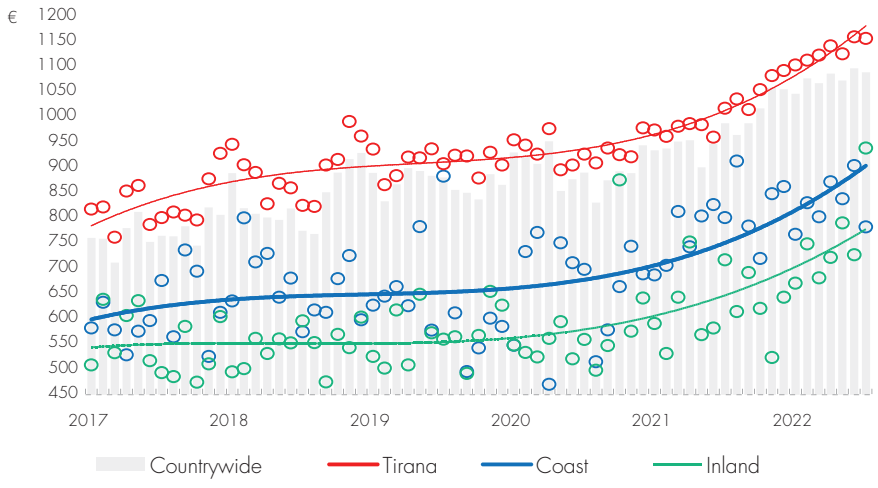
[3] *The main districts and towns pertaining to the "inland" group are: Berat, Cërrik, Delvinë, Devoll, Dibër, Elbasan, Fier, Fushë-Krujë, Gjirokastër, Gramsh, Kamëz, Korçë, Krujë, Kuçovë, Kukës, Laç, Lezhë, Librazhd, Lushnjë, M. e Madhe, Mirditë, Pogradec, Përmet, Shkodër and Tepelenë.*

The tools used for the electronic data gathering procedure include the Python framework and programing language in conjunction with its sub-component Scrapy – a website data scraping engine. The data gathering procedure consists in accessing the HTML structure of the website njoftime.com through the Python framework and Python code, then employing Scrapy in order to identify, collect and save locally all the needed bits of information for each sale ad entry in the residential property sales category of the website. The code instructs the program to collect – for each price listing – the date, the listing entity, the header and the general details of the ad entry, and store it as raw information in a data interchange file. This step is then repeated in a loop fashion for all entries of all specified pages of the website, as determined by the researcher, ending up with a dataset consisting of 4 columns of raw information where each dataset row corresponds to each ad entry. Once raw data were collected and saved, we have cleaned, appended, and organized them into specified variables, with no outliers, no missing values. The variables: city/town/village, geographical group, surface area in meters squared and price in Euros, were determined by identifying the related bits of information from the raw data.

RESULTS

The distribution in time of the monthly averages of the per-meter-squared prices in Euros is given by the chart in Figure 2. The red line is a polynomial regression showing the monthly averages dynamics of for sale house prices in Tirana; the blue line shows the same dynamics for house prices on the coast; the green line shows the dynamics of the house prices in inland areas of the country. The grey bar element of the chart shows the monthly averages of the house prices countrywide, based on all available price listings.

Figure 2. The dynamics of monthly averages of per-meter-squared prices by geographical group



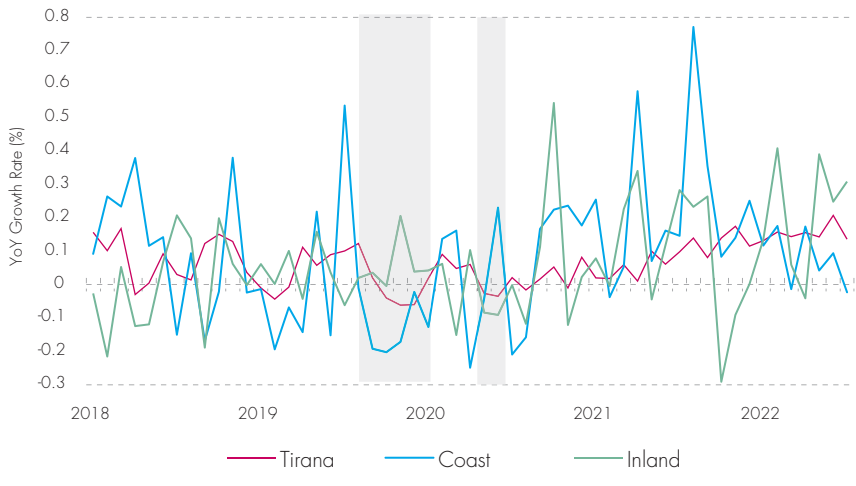
As evident, for the investigated time period, house prices show positive growth inclinations countrywide. The growth appears to accelerate noticeably in

2020, right after the COVID-19 economic lockdown. Data classification by geographical area portrays similar growth trends. These trends are almost identical for Tirana and coastal areas. Meanwhile price averages in the country's inland areas appear static and equable up to mid 2020, displaying a noticeable growth trend thereafter.

The chart reveals differences in price averages as well. There currently is a gap of about 270 Euros between per-meter-squared prices of houses in Tirana and those on the coast, which appears to be present since 2019. While the current gap between prices of houses on the coast and those in inland areas seems to be about 120 Euros. If we refer to the average standing of house prices in Tirana – where the overwhelming share of house sales occurs – which was 1175 Euros in July 2022, the average of house prices for the coast stands at about 80%, while the average of house prices for inland areas stands at about 68%. As of the first half of 2022, the average of per-meter-squared prices: in Tirana was 1128 Euros, on the coast was 837 Euros, and in inland areas was 724 Euros. Compared to the first half of 2021, the YoY growth of the above-noted price averages: for Tirana was 15.6%, for the coast was 9.6%, and for inland areas was 18.3%.

The chart in Figure 3 helps us understand the YoY price growth dynamics by geographical group. In regard to Tirana and the coast, growth rates appear to be positive for most of the investigated time period. The time interval in which prices appear to have a negative growth spur is during the two natural disasters, in the second half of 2019. The chart displays clearly that the coast undergoes a sharper decline of prices; the earthquakes affected these areas the most.

Figure 3. YoY growth of monthly averages of per-meter-squared prices, by geographical group



During the COVID-19 economic lockdowns in the 2nd quarter of 2020, a slight decline of price averages is noted for Tirana, but not for the coast. As for the inland areas, the growth rate fluctuates around zero for about 3 years (2018-2020). Since the start of 2021 onward, all three geographical groups display positive growth of their respective price trends.

CONCLUDING REMARKS

To recap, this research was conducted using microdata gathered electronically by means of big data congregation tools such as the Python framework and a set of web scrapping tools and procedures. The objective from the outset was to estimate average standings of prices and to track their dynamics in time, for three geographical groups: Tirana, the coast, and the country's inland areas. The data was sourced from njoftime.com, Albania's largest UGC transactions information website, and contained 30,467 observations (price listings) spanning from Jan. 2017 to Jul. 2022. The raw data collection phase was followed by a laborious phase of identifying, cleaning, sorting, organizing, and calculating the collected information into tabular form, with specified variables, no outliers and no missing data. This was accomplished through MS Excel and its subcomponent Power Query. The last phase was that of using MS Office's other subcomponent Power Pivot to estimate the monthly averages of prices. The main findings are detailed below:

- The current gap of price averages between Tirana and coastal areas is about 270 Euros, while that between coastal areas and inland areas is about 120 Euros.
- Estimations for data divided by geographical area show that, in contrast to coastal areas and Tirana, where house prices have been on a steady rising trend, prices in inland areas have been stationary for most of the investigated period. Only after the COVID-19 lockdowns prices in inland areas start to follow a rising trend.
- The 2019 natural disasters impacted the house price dynamics negatively, predominately in coastal areas, but in Tirana also.
- YoY gowth of the average standing of house prices during the first half of 2022 ended up at: about 15.6% for Tirana; about 9.6% for coastal areas, and; about 18.3% for inland areas.

BIBLIOGRAPHY

Hajba, G. L. (2018). Website Scraping with Python. Sopron, Hungary: Apress.

Kouzis-Loukas, D. (2016). Learning Scrapy. Birmingham: Packt Publishing Ltd.

Ministry of Finances, & Ministry of Justice. (2018, July 30). Joint directive no. 29 regarding the tax on the transfer of the real estate ownership right. Tirana, Albania.