Novi Ekonomist
Vol 20(2), Year XX, Issue 40, july – december 2026.
ISSN 1840-2313 (Print) 2566-333X (Online)
DOI: 10.69781/NOEK202640006

Submitted: 09.04.2026.
Accepted: 26.05.2026.
ORIGINAL ARTICLE
UDK: 336.74(4-672EU):339.722(4-15)

TESTIN`G THE VALIDITY OF THE PURCHASING POWER PARITY THEORY FOR THE WESTERN BALKAN COUNTRIES

Srđan Kujundžić
Faculty of Economics Pale, University of East Sarajevo, Republic of Srpska, BiH
srdjan.kujundzic@ekofis.ues.rs.ba
ORCID: 0009-0000-4849-8598

Luka Marković
Faculty of Economics Pale, University of East Sarajevo, Republic of Srpska, BiH
luka.markovic@ekofis.ues.rs.ba
ORCID: 0009-0000-4820-7914

Paper presented at the 14th International Scientific Conference EKONBIZ 2026 - The Economy under Conditions of Uncertainty and Innovation, June 12th 2026, Bijeljina.

Abstract: *The aim of this research is to empirically examine the long-term validity of the Purchasing Power Parity (PPP) theory in the Western Balkan countries—Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Albania—in relation to the U.S. dollar. The study applies modern methodological techniques for testing the stationarity of time series, including univariate ADF, PP, and KPSS tests, as well as panel SURADF tests, both in their classical form and with the inclusion of a Fourier component. The univariate tests indicate that real exchange rates are non-stationary in levels but become stationary after first differencing, implying that standard tests do not confirm the validity of the PPP theory. The classical SURADF test also does not provide strong evidence of stationarity, except in the case of North Macedonia. However, the extended SURADF test with a Fourier function consistently rejects the null hypothesis of a unit root for all analyzed countries, indicating the existence of a long-term equilibrium in real exchange rates. To the best of our knowledge, this study represents the first systematic application of the Fourier-augmented panel SURADF unit root test to the empirical testing of the Purchasing Power Parity hypothesis for Western Balkan economies. By explicitly allowing for smooth nonlinear structural changes, the applied methodology significantly enhances the detection of mean-reverting behavior in real exchange rates that conventional linear and panel unit root tests fail to identify. The findings therefore provide robust empirical support for the long-run validity of the PPP hypothesis in the region and offer relevant implications for exchange rate policy formulation and assessments of international competitiveness.*

Key words: *purchasing power parity, real exchange rate, stationarity, SURADF test, Fourier function, Fourier-augmented panel unit root test, Western Balkans.*

JEL classification*: F31, F41, C33, C22*

1. INTRODUCTION

The Purchasing Power Parity (PPP) theory represents one of the central theoretical foundations of international macroeconomics and exchange rate determination. Its validity has direct implications for the assessment of external competitiveness, exchange rate misalignment, and the formulation of monetary and exchange rate policies. In particular, the long-run validity of PPP is crucial for small open economies, where exchange rate dynamics play a key role in shaping inflationary pressures and external balance.

Despite an extensive empirical literature on PPP, existing studies predominantly focus on advanced and emerging economies, while empirical evidence for the Western Balkan countries remains relatively limited and methodologically fragmented. Moreover, many earlier studies rely on conventional linear unit root and panel tests, which are known to suffer from low statistical power and sensitivity to structural breaks—features that are especially pronounced in

transition economies exposed to frequent economic and institutional shocks.

However, although numerous studies find evidence of a unit root in the real exchange rate, some authors warn that such findings are often the result of the low statistical power of conventional tests. Recent research indicates that these tests not only ignore cross-sectional (inter-regional) information—which reduces estimation efficiency—but also have limited ability to distinguish between nearly non-stationary but actually stationary processes. According to Krugman, P., Obstfeld, M., & Melitz, M. (2017), there are several reasons why PPP may not hold in practice: (1) the presence of barriers to international trade (e.g., tariffs, transportation costs), which hinder arbitrage and violate the law of one price; (2) deviations from perfect competition, where monopolies and oligopolies influence price formation; and (3) difficulties in defining a representative consumer basket, since its composition varies among countries, making it difficult to align inflation differentials through exchange rate adjustments.

To increase the statistical power of stationarity testing, many authors have turned to the use of panel data. Levin, A., C. F. Lin, and C. S. Chu (2002), as well as Im, K. S., M. H. Pesaran, and Y. Shin (2003), developed extended ADF tests adapted for panel analyses, which demonstrated significant improvements in power even in small samples. However, these tests share a common shortcoming—dependence among time series within the panel (so-called cross-sectional dependence). To overcome this issue, Zellner, A. (1962) proposed the Seemingly Unrelated Regressions (SUR) method, and O'Connell, P. G. J. (1998) showed that SUR estimates can eliminate bias in testing without a significant loss of power.

Despite methodological advancements, Taylor, M. P. and Sarno, L. (1998) and Breuer, J. B., McNown, M. S., and Wallace, M. (2001) emphasize that the problem of "all-or-nothing" tests still remains—tests that do not allow the identification of which specific series within the panel are stationary, as the null hypothesis of a unit root is tested jointly for all panel members. In this context, Breuer and co-authors (2001) draw a parallel with the F-test in regression analysis: the rejection of the joint null hypothesis does not necessarily mean that all individual coefficients are statistically significant. Analogously, rejecting the null hypothesis of a unit root does not imply that all observed series are stationary.

Unlike such conventional approaches, panel SURADF tests allow for testing individual null hypotheses for each series separately, thereby precisely identifying which and how many series within the panel exhibit stationary behavior. For this reason, the SURADF methodology—with and without the Fourier function—was applied in this paper to test the long-term validity of PPP among the Western Balkan countries.

Chang, T., P. I. Chou, C. H. Lee, and S. C. Wang (2010) applied the SURADF test to monthly real exchange rate data (1980–2008), constructed from CPI indices and nominal exchange rates against the USD, and demonstrated that PPP holds for most G-7 countries (except Canada and Italy). Similarly, Chang, T., Lee, C. H., Chou, P. I., and Wang, S. C. (2012) used the panel SURADF test with a Fourier function for nine transition economies (Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Russia) for the period January 1995–December 2008. Their analysis confirmed the validity of PPP in eight of these countries, with Russia as the exception.

The analysis conducted in this paper relies on the methodological approaches of the aforementioned authors. For the initial examination of the stationarity of real exchange rates by country, univariate tests—ADF, PP, and KPSS—were employed. However, considering their limited power and sensitivity to structural breaks, panel SURADF tests (both with and without the Fourier function) were additionally applied, allowing for a deeper understanding of the time-series dynamics.

Finally, it is important to emphasize that this research represents one of the first systematic empirical investigations of the Purchasing Power Parity theory in the context of Western Balkan countries, thereby contributing to the expansion of the existing literature on this topic.

In this context, the present study aims to empirically examine the long-run validity of the Purchasing Power Parity hypothesis for the Western Balkan countries—Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Albania—using a set of advanced time-series and panel econometric techniques. In addition to standard univariate unit root tests (ADF, PP, and KPSS), the analysis applies the panel SURADF test in both its classical and Fourier-augmented forms. By allowing for smooth nonlinear structural changes, the Fourier-SURADF approach enables a more accurate identification of mean-reverting behavior in real exchange rates. The contribution of this paper lies in providing new empirical evidence for the region based on a methodology that explicitly accounts for the complex and nonlinear dynamics characterizing transition economies.

2. THE REAL EXCHANGE RATE MODEL AND THE THEORY OF PURCHASING POWER PARITY (PPP) – THEORETICAL IMPLICATIONS

The real exchange rate represents the relative price of domestic goods in terms of foreign goods, expressed in the same currency. The general formula is:

$$r_t = e_t \times (p_t^*/p_t)$$

where e is the nominal exchange rate of the domestic currency against the foreign currency, r is the real exchange rate, and p and p^* represent the domestic and foreign price levels, respectively.

The main prediction of the Purchasing Power Parity (PPP) theory is that the real exchange rate remains constant and is therefore used to quantify deviations from PPP as well as to analyze macroeconomic conditions of supply and demand in open economies. It also reflects the competitiveness of domestic goods relative to foreign ones when taking into account the exchange rate and price levels in both countries.

In the conducted research, the logarithmic values of the real exchange rate were used. Starting from the general specification of the real exchange rate:

$$r_t = e_t \times (p_t^*/p_t)$$

taking natural logarithms of both sides yields:

$$\ln(r_t) = \ln(e_t) + \ln(p_t^*) - \ln(p_t)$$

which can be written in a simplified log-linear form as:

$$r_t = e_t + p_t^* - p_t$$

where lowercase variables denote the natural logarithms of the corresponding uppercase variables.

From a statistical perspective, testing the validity of the Purchasing Power Parity (PPP) theory reduces to examining the existence of a unit root in the time series of the real exchange rate r. If the presence of a unit root is confirmed, this implies that PPP does not hold in the long run, meaning that there is no stable long-term relationship between the nominal exchange rate and relative price levels. Conversely, the stationarity of the real exchange rate indicates that the nominal exchange rate adjusts to inflation differentials over the long term, thereby confirming the theoretical foundation of PPP.

The non-stationarity of the real exchange rate carries important macroeconomic implications. Dornbusch, R. (1987) emphasized, among other things, that the depreciation of the real exchange rate can increase international competitiveness, potentially leading to a reallocation of employment in favor of the country whose currency is weakening. Therefore, determining the empirical sustainability of PPP is of crucial importance for economic policies that rely on exchange rate stability.

Additionally, one of the key effects of the non-stationarity of the real exchange rate is the possibility of persistent arbitrage profits in goods trade. Parikh, A. and Williams, G. (1998) warn that such exchange rate dynamics may lead to serious macroeconomic imbalances, which often require corrective devaluation to restore external equilibrium.

3. DATA

The conducted empirical analysis covers the following five countries of the region: Albania, Bosnia and Herzegovina, Serbia, North Macedonia, and Montenegro. Monthly data were used, and the analyzed period spans from July 2006 to December 2023 (a total of 210 months). All data, including the Consumer Price Index (CPI) with a 2010 base year and nominal exchange rates of the countries' currencies against the U.S. dollar, were obtained from the International Monetary Fund – International Financial Statistics (IFS).

Before conducting the econometric analysis, all CPI and nominal exchange rate data were transformed into natural logarithms. A summary of the data statistics is presented in Table 1 below. The obtained standard deviations indicate relatively low dispersion across all series, with the highest variability observed in Bosnia and Herzegovina (0.1493) and the lowest in Albania (0.0905). The Jarque–Bera normality test indicates that the nominal exchange rates for all countries are normally distributed.

As a robustness consideration, alternative specifications using the euro as the reference currency may be explored in future research to assess the sensitivity of the results to the choice of numeraire.

All macroeconomic data required for the empirical analysis were obtained from official international statistical sources. Consumer Price Index (CPI) series are expressed with a common base year (2010 = 100), ensuring cross-country comparability, while nominal exchange rates are defined as units of domestic currency per U.S. dollar. The choice of the U.S. dollar as the numeraire follows the standard practice in the empirical PPP literature and facilitates international comparability of the results.

Table 1. – Data Description and Sources

Variable	Description and Source
CPI	Consumer Price Index (2010 = 100), IMF – International Financial Statistics
ER	Nominal exchange rate (domestic currency per USD), IMF – International Financial Statistics
RER	Real exchange rate constructed by authors
Period	Monthly data, July 2006 – December 2023

Source: Author's calculations

Table 2. - Summary Statistics

Country	Skewness	Kurtosis	Std. dev.	Min	Max	JB-stat	Prob
B&H	0,2991	0,0	0,1493	0,2291	0,7713	3,1311	0,209
Serbia	0,002	0,4381	0,1113	4,0514	4,5633	1,6795	0,4318
Montenegro	0,0982	0,0	0,1032	-0,4181	-0,0135	0,3375	0,8447
North Macedonia	0,0941	0,0	0,1134	3,6497	4,0881	0,3099	0,8565
Albania	0,1174	0,2568	0,0905	4,3882	4,8434	1,0594	0,5888

Source: Author's calculations

4. METHODOLOGY

Breuer and co-authors (2001) developed the SURADF tests, which represent an extended version of the ADF tests based on panel estimation using the Seemingly Unrelated Regressions (SUR) method.

For each panel member, the augmented Dickey–Fuller (ADF) regression estimated within the SUR framework can be expressed as follows:

$$\Delta X_{1,t} = \alpha_1 + \beta_1 X_{1,t-1} + \sum_{j=1}^{k1} \theta_{1,j} \Delta X_{1,t-j} + \varepsilon_{1,t}$$

$$t = 1,2,\dots,T$$

$$\Delta X_{2,t} = \alpha_2 + \beta_2 X_{2,t-1} + \sum_{j=1}^{k2} \theta_{2,j} \Delta X_{2,t-j} + \varepsilon_{2,t}$$

$$t = 1,2,\dots,T$$

$$\Delta X_{N,t} = \alpha_N + \beta_N X_{N,t-1} + \sum_{j=1}^{kN} \theta_{N,j} \Delta X_{N,t-j} + \varepsilon_{N,t}$$

$$t = 1,2,\dots,T$$

Following the approach proposed by Perron (1989), the lag length p_i is selected individually for each country in order to avoid size distortions arising from imposing a common lag structure across panel members. The optimal lag length is determined using standard information criteria, including the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC). The system of ADF equations estimated here allows for individual testing of the null and alternative hypotheses for each panel member, with test statistics obtained from the SUR estimations of the specified system.

The authors emphasized that applying a uniform lag structure for all panel members could lead to biased results; therefore, the lag structure is determined separately for each equation, in accordance with the approach proposed by Perron (1989).

The key difference between the panel SURADF test and other panel unit root tests—such as the multivariate augmented Dickey–Fuller (MADF) test proposed by Sarno, L. and Taylor, M. P. (2002)—lies in the formulation of the null hypothesis. Specifically, while the aforementioned tests jointly examine the existence of a unit root for all panel members, the SURADF test allows for separate testing of the null hypothesis for each individual member. This enables clear identification of which series within the panel exhibit stationary properties and how many satisfy the conditions of stationarity.

Each of the *N* null and alternative hypotheses is tested separately:

$$H_0^1: \beta_1 = 0;\ \ H_A^1: \beta_1 < 0$$
$$H_0^2: \beta_2 = 0;\ \ H_A^2: \beta_2 < 0$$
$$H_0^N: \beta_N = 0;\ \ H_A^N: \beta_N < 0$$

In this study, a system of ADF equations that includes a Fourier function was also estimated.

The introduction of the functions [sin(2πkt/T), cos(2πkt/T)] is justified by the fact that the Fourier approximation allows for a very precise representation of any absolutely integrable function, where *k* denotes frequency, and the coefficients [ai, bj]′ determine the amplitude and phase shift of the frequency component. The presence of at least one frequency component indicates the existence of a structural break in the series.

Numerous studies, including those by Becker et al. (2004), Enders and Lee (2004), Gallant (1981), and Pascalau (2010), show that the Fourier approximation can reliably describe even functions that are not necessarily periodic.

The authors emphasize that, for conducting the test, it is sufficient to appropriately select the frequency within the model. This reduces the number of estimated parameters, thereby improving the size and power properties of the test, regardless of the timing and shape of possible structural breaks.

Since the form of structural changes in the data is not known in advance, a grid-search method is used to identify the optimal frequency, after which the test statistics are calculated based on the SUR estimates of system (1).

Furthermore, Breuer et al. (2001) showed that imposing the same lag structure for all panel members can result in biased tests. Therefore, the number and order of lags are chosen separately for each equation, following the approach proposed by Perron (1989).

The system of ADF equations with the Fourier function estimated in this study is as follows.

To account for potential smooth structural changes, the ADF regression is augmented with a Fourier function and can be written as:

$$\Delta X_{1,t} = \alpha_1 + \beta_1 X_{1,t-1} + \sum_{j=1}^{k1} \theta_{1,j} \Delta X_{1,t-j} + a_1 \sin\left(\frac{2\pi kt}{T}\right) + b_1 \cos\left(\frac{2\pi kt}{T}\right) + \varepsilon_{1,t}$$

$$\Delta X_{2,t} = \alpha_2 + \beta_2 X_{2,t-1} + \sum_{j=1}^{k2} \theta_{2,j} \Delta X_{2,t-j} + a_2 \sin\left(\frac{2\pi kt}{T}\right) + b_2 \cos\left(\frac{2\pi kt}{T}\right) + \varepsilon_{2,t}$$

$$\Delta X_{N,t} = \alpha_N + \beta_N X_{N,t-1} + \sum_{j=1}^{kN} \theta_{N,j} \Delta X_{N,t-j} + a_N \sin\left(\frac{2\pi kt}{T}\right) + b_N \cos\left(\frac{2\pi kt}{T}\right) + \varepsilon_{N,t}$$

The optimal Fourier frequency *k* is selected using a grid-search procedure that minimizes the sum of squared residuals, allowing the model to flexibly capture smooth structural changes of unknown form and timing. The selected frequency values for each country are reported in the Appendix.

5. EMPIRICAL RESULTS

In the first phase of the analysis, the stationarity of the real exchange rate was tested for five Western Balkan countries—Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Albania—using standard univariate tests: the Augmented Dickey–Fuller (ADF), Phillips–Perron (PP), and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. The results presented in Table 1 indicate that the null hypothesis of a unit root (non-stationarity) cannot be rejected in levels but is strongly rejected after first differencing for all examined countries. Specifically, all ADF and PP tests show highly negative test statistic values after differencing (at the 1% significance level), confirming that all series are integrated of order I(1). For example, the ADF statistic for Bosnia and Herzegovina in differenced form is −11.271 with zero lags, which is well below the critical value at the 1% significance level. Similar findings apply to the other countries—Serbia (−9.679), Montenegro (−11.334), North Macedonia (−11.250), and Albania (−5.601)—indicating a very strong presence of stationarity after differencing. On the other hand, the KPSS tests—which have the reverse null hypothesis (the null assumes stationarity)—further confirm non-stationarity in levels. Specifically, all level series exhibit statistic values well above the corresponding critical thresholds. For instance, the KPSS statistic for Montenegro is 1.4714, while for Bosnia and Herzegovina it reaches as high as 1.8886, clearly indicating the rejection of the null hypothesis of stationarity in levels.

Table 3. – Univariate Unit Root Tests (ADF, PP, KPSS)

Country	ADF Level	ADF Diff	KPSS Level	KPSS Diff	PP Level	PP Diff
B&H	-1.121 (1)	-11.271 (0)***	1.8886*** (1)	0.0899 (1)	78.969 (1)	3.075 (0)
Serbia	-2.549 (1)	-9.679 (0)***	1.1131*** (1)	0.0591 (0)	38.940 (1)	5.259 (0)
Montenegro	-1.871 (1)	-11.334 (0)***	1.4714*** (1)	0.1107 (0)	54.617 (1)	3.413 (0)
North Macedonia	-1.715 (1)	-11.250 (0)***	1.6172*** (1)	0.0768 (0)	57.039 (1)	3.419 (0)
Albania	-2.489 (4)	-5.601 (3)***	0.8297*** (4)	0.1159 (3)	41.361 (4)	3.245 (3)

Source: *Author's calculation*

The failure of conventional univariate unit root tests to confirm the validity of PPP may be partly attributed to the presence of major global and regional shocks during the sample period. In particular, the global financial crisis of 2008–2009 and the COVID-19 pandemic introduced pronounced and persistent deviations in real exchange rates, driven by abrupt changes in capital flows, trade patterns, and domestic price dynamics. Such events are likely to generate nonlinear adjustment paths that standard linear tests are unable to capture adequately.

Based on these findings, it can be concluded that the real exchange rates for all observed countries are non-stationary in levels but become stationary after first differencing.

These results are consistent with previous empirical research suggesting that univariate tests have low power, particularly for time series with high autocorrelation (Taylor, M. P. and Sarno, L., 1998; Im et al., 2003). Accordingly, these results imply that the Purchasing Power Parity (PPP) hypothesis cannot be confirmed when traditional univariate testing methodologies are used.

In the next phase of the empirical analysis, the SURADF test without the inclusion of the Fourier component was applied to further examine the stationarity of real exchange rates in the analyzed countries.

Unlike classical univariate tests, the SURADF methodology allows for the simultaneous estimation of interdependent time series while retaining the individual characteristics of each country (Breuer et al., 2001). The results are presented in Table 4.

It is observed that the t-statistics for most countries (Bosnia and Herzegovina, Serbia, Montenegro, and Albania) exceed the critical value at the 10% significance level, implying that the null hypothesis of a unit root cannot be rejected. This means that the real exchange rate in these countries remains non-stationary even when cross-sectional dependencies among panel members are taken into account, indicating that PPP does not hold in the long run.

The only exception is North Macedonia, for which the t-statistic is −2.315, significant at the 5% level (marked with **). Since this value is lower than the corresponding critical value (−1.497), it can be concluded that the null hypothesis is rejected and the alternative hypothesis of stationarity of the real exchange rate is accepted for North Macedonia, thereby partially confirming the validity of the PPP theory.

The estimated β parameter values are negative for all countries, indicating the presence of a medium-term mean-reverting force toward equilibrium. However, their low absolute magnitudes result in high estimated half-lives of adjustment—from 27 months for North Macedonia to as high as 133 months for Albania—implying a very slow convergence process of exchange rates toward long-term equilibrium.

These findings are consistent with those of Breuer et al. (2001) and O'Connell, P. G. J. (1998), who pointed out that standard SURADF tests without correction for structural changes may have limited power in detecting stationarity.

By contrast, the Fourier-augmented SURADF test explicitly allows for smooth structural changes in the data-generating process, making it particularly suitable for economies characterized by gradual institutional reforms and repeated external shocks.

The rejection of the unit root null hypothesis for all countries under the Fourier-SURADF framework suggests that real exchange rates exhibit nonlinear mean-reverting behavior once such structural features are properly accounted for

Table 4. – SURADF Unit Root Test (without Fourier Components)

Country	B	t-statistics	Half-Life (months)	Critical value 1%	Critical value 5% and 10%
B&H	-0.010	-1.586	69.315	-2.687	-1.689 -1.465
Serbia	-0.017	-1.985	39.576	-2.610	-1.555 -1.377
Montenegro	-0.008	-1.234	86.356	-2.571	-1.545 -1.338
North Macedonia	-0.023	-2.315**	27.198	-2.498	-1.497 -1.297
Albania	-0.005	-1.141	132.876	-2.486	-1.488 -1.298

Source: *Author's calculation*

The estimated half-life values indicate a relatively slow speed of adjustment toward long-run equilibrium, ranging from just over two years in the case of North Macedonia to more than ten years for Albania. These prolonged adjustment periods are consistent with the structural characteristics of Western Balkan economies, including limited market integration, exchange rate management practices, and persistent nominal rigidities. Similar magnitudes of half-life estimates have been reported for other transition and emerging economies in Central and Eastern Europe.

Due to the aforementioned limitations and the low statistical power of conventional tests in the presence of nonlinear shocks or structural breaks, the extended SURADF test with a Fourier component is applied in the subsequent part of the study. This approach allows for more flexible modeling of potential structural changes in the time series (Becker et al., 2004; Enders, W. and Lee, J., 2004).

Unlike the previous tests without the Fourier component, all observed countries—Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Albania—now exhibit statistically significant negative t-values, all of which are lower than the critical value at the 1% significance level. For example, Serbia's t-statistic is −4.354, compared to the critical value of −2.336, while Bosnia and Herzegovina's t-statistic is −2.874, with a corresponding critical value of −2.694. Similar results are obtained for the other countries as well.

These findings clearly indicate the rejection of the null hypothesis of a unit root and confirm the stationarity of real exchange rates in all analyzed countries once potential smooth structural breaks are taken into account. In other words, the Purchasing Power Parity (PPP) theory finds empirical support in the long run for all Western Balkan countries included in this analysis.

The estimated β coefficients are all negative and statistically significant, indicating the presence of a medium-term mean-reverting force toward equilibrium. Moreover, the estimated half-lives of adjustment are considerably shorter than in the tests without the Fourier component—ranging between 6.31 months (Serbia) and 10.9 months (Bosnia and Herzegovina)—implying a faster convergence of exchange rates toward long-term equilibrium. This result aligns with claims in the literature that incorporating Fourier approximation enhances test sensitivity and improves the detection of stationarity in the presence of nonlinearities and unaccounted regime shifts (Gallant, R., 1981; Pascalau, R., 2010).

Table 5. – SURADF Fourier Unit Root Test

Country	B	t-Statistics	Half-Life (months)	Critical value 1%	Critical value 5% and 10%
B&H	-0.0616	-2.874***	10.9	-2.694	-2.024 -1.700
Serbia	-0.1040	-4.354***	6.31	-2.336	-1.660 -1.326
Montenegro	-0.0974	-3.665***	6.76	-2.223	-1.835 -1.492
North Macedonia	-0.1009	-3.890***	6.52	-2.003	-1.450 -1.165
Albania	-0.0929	-3.732***	7.11	-2.373	-1.766 -1.249

Source: *Author's calculation*

Compared to the test without the Fourier component, it is clear that this approach provides more robust and informative conclusions, further supported by the findings of Chang et al. (2012), who noted that the Fourier approach significantly improves the test's size and power properties, especially for time series characterized by latent structural breaks.

6. EMPIRICAL SIZE AND POWER OF THE SURADF TEST – FOURIER VS. CLASSICAL APPROACH

To assess the statistical properties of the panel SURADF tests, a simulation analysis was conducted to examine the empirical size and power of the test with and without the Fourier component. The results are presented in Table 6.

Empirical size measures the probability that the test incorrectly rejects a true null hypothesis (Type I error), while empirical power indicates the test's ability to correctly reject a false null hypothesis (Type II error). Ideally, a test should exhibit a low size and high power.

In terms of test size, the SURADF test with the Fourier component shows a slightly higher, yet still acceptable, rate of false rejection of the null hypothesis. For instance, North Macedonia records a size of 0.15, which is slightly above the conventional tolerance level, while for most countries (Bosnia and Herzegovina: 0.06; Serbia: 0.05) the size remains within acceptable limits (up to 0.10). In contrast, the classical SURADF test without the Fourier component displays an almost ideal size (close to zero), but this may make the test overly conservative, thereby reducing its sensitivity.

Regarding power, the results show a clear advantage of the Fourier approach: the test with the Fourier component achieves a power of 1.00 (or 100%) for nearly all countries, including Serbia, Montenegro, North Macedonia, and Albania, while for Bosnia and Herzegovina the power is slightly lower but still very high (0.99). By comparison, the test without the Fourier component demonstrates somewhat weaker performance (e.g., 0.91 for Bosnia and Herzegovina; 0.99 for Montenegro), though still at a satisfactory level.

These findings confirm the claims in the literature (Lau, C. K., 2009; Enders, W. and Lee, J., 2004) that the inclusion of the Fourier transformation significantly improves the statistical properties of the test, particularly in the presence of nonlinear trends or unknown structural changes in time series.

CONCLUSION

The conducted research aimed to examine the long-term validity of the Purchasing Power Parity (PPP) theory using the example of five Western Balkan countries: Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Albania. By applying modern time-series and panel analysis methods—including univariate ADF, PP, and KPSS tests, as well as panel SURADF tests with and without the Fourier component—the study explored whether real exchange rates in these countries revert to equilibrium levels in the long run. The univariate tests showed that all real exchange rates are integrated of order I(1), that is, they become stationary only after first differencing. This suggests that standard tests do not provide sufficiently strong evidence to confirm the PPP theory, consistent with previous findings in the literature (Taylor, M. P. and Sarno, L., 1998; Im et al., 2003). However, the panel SURADF test without the Fourier component provided additional insights. The results showed that the null hypothesis of a unit root could not be rejected for most countries, except for North Macedonia, indicating partial support for the PPP theory. Nonetheless, the high estimated half-lives of adjustment suggest a slow return to equilibrium, further highlighting the limitations of standard panel approaches in the presence of potential structural changes. The application of the extended SURADF test with the Fourier component represents a crucial enhancement of the analysis. The results clearly indicate the rejection of the null hypothesis for all observed countries at a high level of significance, confirming the stationarity of real exchange rates. These findings suggest that the PPP theory holds in the long run for all countries in the sample, but only when flexible nonlinearities and potential smooth structural breaks—which standard tests fail to detect—are incorporated into the model. Moreover, the Fourier-SURADF approach exhibits significantly shorter half-lives of adjustment compared to the classical version of the test, implying faster convergence of exchange rates toward equilibrium and a stronger medium-term mean-reverting force. This confirms the results of previous studies (Becker et al., 2004; Chang et al., 2012), which found that incorporating the Fourier function improves the statistical performance of the test, particularly in the presence of nonlinear or unspecified structural changes. Overall, it can be concluded that the Purchasing Power Parity theory retains empirical validity for the Western Balkan countries in the long run. However, its confirmation depends on the choice of methodology and the model's ability to capture the complex temporal dynamics of real exchange rate series. This has important policy implications, as it

suggests that exchange rates in the region align with relative price movements in the long term and may serve as an indicator of equilibrium relationships in international trade. From a policy perspective, the obtained results provide relevant insights for exchange rate regime choices and macroeconomic coordination in the context of European integration. The long-run validity of PPP implies that real exchange rates in Western Balkan countries tend to adjust toward equilibrium levels, supporting the use of PPP-based indicators in assessing exchange rate misalignment and external competitiveness. This is particularly important for highly euroized economies and countries operating under managed or fixed exchange rate arrangements, where deviations from equilibrium may accumulate in the form of competitiveness losses or external imbalances. Furthermore, the confirmation of PPP when nonlinear adjustment dynamics are taken into account has implications for euro adoption readiness and the broader EU accession macroeconomic framework. Stable long-run price convergence and mean-reverting real exchange rates are consistent with the requirements of nominal and real convergence embedded in the European Union's accession criteria. In this regard, the results suggest that maintaining macroeconomic stability, fiscal discipline, and credible monetary frameworks is essential for facilitating convergence toward equilibrium exchange rates and supporting sustainable integration into the EU single market.

REFERENCES

[1] Becker, R., Enders, W., & Lee, J. (2004). A general test for time dependence in parameters. *Journal of Applied Econometrics, 19*(7), 899–906.

[2] Breuer, J. B., McNown, R., & Wallace, M. S. (2001). Misleading inferences from panel unit root tests with an illustration from purchasing power parity. *Review of International Economics, 9*(3), 482–493.

[3] Chang, T., Lee, C.-H., Chou, P.-I., & Wang, S.-C. (2012). Purchasing power parity for transition countries. *Eastern European Economics, 50*(4), 42–59.

[4] Chang, T., Liu, W.-C., Tzeng, H.-W., & Yu, C.-P. (2010). Purchasing power parity for G-7 countries: Panel SURADF tests. *Applied Economics Letters, 17*(12), 1223–1228.

[5] Dornbusch, R. (1987). Exchange rates and prices. *American Economic Review, 77*(1), 93–106.

[6] Enders, W., & Lee, J. (2004). Testing for a unit root with a nonlinear Fourier function. *Journal of Time Series Analysis, 25*(3), 369–386.

[7] Gallant, R. A. (1981). On the basis in flexible functional form and an essentially unbiased form: The flexible Fourier form. *Journal of Econometrics, 15*, 211–353.

[8] Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics, 115*(1), 53–74.

[9] International Monetary Fund. (2025). *International Financial Statistics*. https://legacydata.imf.org

[10] Krugman, P. R., Obstfeld, M., & Melitz, M. J. (2018). *International economics: Theory and policy* (11th ed.). Pearson.

[11] Lau, C. K. (2009). A more powerful panel unit root test with an application to purchasing power parity. *Applied Economics Letters, 16*(1), 75–80.

[12] Levin, A., Lin, C. F., & Chu, C. S. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics, 108*(1), 1–24.

[13] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics, 44*(1), 1–20.

[14] Parikh, A., & Williams, G. (1998). Modeling real exchange rate behavior: A cross-country study. *Applied Financial Economics, 8*(6), 577–587.

[15] Pascalau, R. (2010). Unit root tests with smooth breaks: An application to the Nelson–Plosser data set. *Applied Economics Letters, 17*(6), 565–570.

[16] Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361–1401.

[17] Sarno, L., & Taylor, M. P. (2002). Purchasing power parity and the real exchange rate. *IMF Staff Papers, 49*, 65–105.

[18] Taylor, M. P., & Sarno, L. (1998). The behavior of real exchange rates during the post–Bretton Woods period. *Journal of International Economics, 46*(2), 281–312.

[19] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*, 348–368.