European International Journal of Multidisciplinary Research and Management Studies
(ISSN: 2750-8587)

Received: 20 April 2026 | Accepted: 05 May 2026 | Published: 31 May 2026

DOI: 10.55640/eijmrms-06-05-11

RESEARCH ARTICLE

Assessing the Impact of Non-Consideration of Stationarity on Regression Analysis: Evidence from Labor Market Data

Aliqulov Abbos Baxtiyor o'g'li
Senior Lecturer at Qarshi International University, Uzbekistan

VOLUME: Vol.06 Issue05 2026

PAGE: 61-66

Abstract

This study examines the empirical errors that may arise from ignoring stationarity properties when assessing the relationship between labor market indicators using regression analysis. The main objective of the research is to demonstrate, through empirical examples, the impact of the data preprocessing stage on regression results when working with time series and panel data.

The analysis is based on open statistical data on the labor market of Uzbekistan, where wage and employment indicators are considered as panel data across regions for the period 2017–2024. Within the empirical framework, regression results obtained without accounting for stationarity are compared with those estimated using first-differenced variables. The findings indicate that constructing regressions directly on trending and non-stationary series may yield statistically significant results; however, such results can lead to substantively incorrect and unreliable conclusions.

KEYWORDS

Labor market indicators, regression analysis, stationarity, panel data, time series, spurious regression.

INTRODUCTION

Regression analysis is one of the most widely used methods for the empirical investigation of economic and social processes. In particular, regression models based on time-series and panel data are extensively applied in both academic research and practical studies. However, when working with such data, neglecting their statistical properties, especially stationarity, may seriously affect the reliability of regression results. Classical studies by Granger and Newbold empirically demonstrated that regressions estimated using non-stationary series can lead to spurious relationships [1].

Fundamental principles of time-series analysis indicate that estimating regressions directly on trending and non-stationary data may yield high statistical significance and goodness-of-fit measures; nevertheless, such results can lead to substantively misleading conclusions [2]. This issue is particularly evident in the analysis of economic indicators, such as wages and employment, which tend to exhibit persistent changes over time.

Labor market indicators are characterized by considerable dynamics across both time and regions and are frequently used as dependent and explanatory variables in regression analysis. At the same time, the issue of stationarity becomes more complex in panel data settings because both cross-sectional and time dimensions are present simultaneously. As emphasized by Wooldridge, ignoring the time-series properties of panel data may result in misleading interpretations of empirical findings [3].

The primary objective of this study is to demonstrate the empirical consequences of disregarding stationarity in regression analysis using labor market indicators as an example. The study employs publicly available statistical data on the labor market of Uzbekistan and analyzes wage and employment indicators across regions for the period 2017–2024 in the form of panel data. Within the econometric framework, regression results obtained without conducting stationarity tests are compared with those derived from regressions estimated using differenced variables. This approach aims to highlight the empirical importance of the data preprocessing stage in regression analysis and to emphasize the necessity of accounting for stationarity when conducting econometric research.

LITERATURE REVIEW

The issue of stationarity is considered one of the key prerequisites for ensuring the methodological validity and reliability of regression analysis when working with time-series and panel data. The credibility of regression results largely depends on the statistical properties of the data employed, and neglecting these properties may lead to misleading empirical conclusions. In studies based on panel data, the problem of stationarity becomes particularly important because such datasets simultaneously incorporate both temporal dynamics and cross-sectional dimensions. Wooldridge emphasizes that ignoring the time-series properties of panel data can result in incorrect interpretations of empirical findings and argues for the necessity of evaluating the temporal characteristics of data prior to conducting regression analysis [3].

In contemporary econometric literature, differencing techniques and unit root tests are regarded as the primary tools for identifying stationarity. Enders demonstrates that statistical significance in regression models does not necessarily guarantee substantive validity and highlights the importance of testing the stationarity properties of time series in advance to avoid empirical errors [4]. Similarly, the classical econometric approach presented by Gujarati and Porter notes that regressions estimated using non-stationary series may produce spurious relationships, leading researchers to draw misleading conclusions [5]. This issue is especially relevant in the analysis of socioeconomic indicators, where careful interpretation of regression results is required.

A review of the existing empirical literature suggests that researchers often focus primarily on the statistical performance of regression models, such as significance levels and goodness-of-fit measures, while paying insufficient attention to the stationarity properties of the underlying data. This methodological limitation is particularly evident in studies employing regional panel datasets, where the issue of stationarity is frequently overlooked. The present study seeks to address this gap by examining the consequences of disregarding stationarity in regression analysis using labor market panel data and by comparing estimation results obtained before and after applying stationarity adjustments.

METHODOLOGY

This study employs an econometric panel data approach to examine the empirical consequences of ignoring stationarity in regression analysis. The analysis is based on labor market indicators of Uzbekistan, specifically wage and employment data across regional units over the period 2017–2024, structured in the form of a balanced panel dataset.

The empirical investigation is conducted in two main stages. In the first stage, a baseline regression model is estimated without performing stationarity tests on the variables. This specification reflects a typical empirical approach where time-series properties are not explicitly considered. In the second stage, the variables are transformed through differencing to ensure stationarity, and the regression analysis is re-estimated using the transformed data. This two-step procedure allows for a direct comparison between non-stationary and stationary model specifications.

Stationarity is assessed using unit root tests commonly applied in econometric analysis. The regression results are evaluated and compared in terms of statistical significance, coefficient stability, and economic interpretability. All computations are carried out using the EViews software environment, which is widely applied for econometric analysis of panel data.

RESULTS AND DISCUSSION

The panel dataset used in this study covers wage and employment dynamics across regions of Uzbekistan for the period 2017–2024. Descriptive statistics for the main variables are presented in Table 1. The results indicate substantial variation in both wage and employment indicators across regions and over time. The observed range of values suggests that these variables exhibit pronounced dynamic behavior, reflecting structural and temporal heterogeneity in the labor market.

Table 1. Descriptive statistics of wage and employment indicators

Indicators	Wage	Employment
Mean	241 442.0	974 036.6
Median	213 496.0	1 029 900.0
Maximum value	721 771.0	1 563 900.0
Minimum value	99 411.87	332 200.0
Standard deviation	110 107.5	375 162.3
Skewness	1.437384	−0.226705
Kurtosis	6.170358	1.789953
Jarque–Bera test	85.47214	7.792383
Probability	0.000000	0.020319
Sum	2.70E+08	1.09E+08
Sum of squared deviations	1.35E+14	1.56E+13
Number of observations	112	112

Source: Authors' calculations based on open data from the State Statistics Committee of the Republic of Uzbekistan, using EViews software.

The overall time dynamics of the data are presented in Figure 1. The graph is constructed based on average values across regions and clearly demonstrates that both wage and employment indicators exhibit a stable upward trend over the analyzed period. This pattern provides preliminary evidence suggesting the possible non-stationarity of the variables and highlights the necessity of conducting separate stationarity tests prior to performing regression analysis.

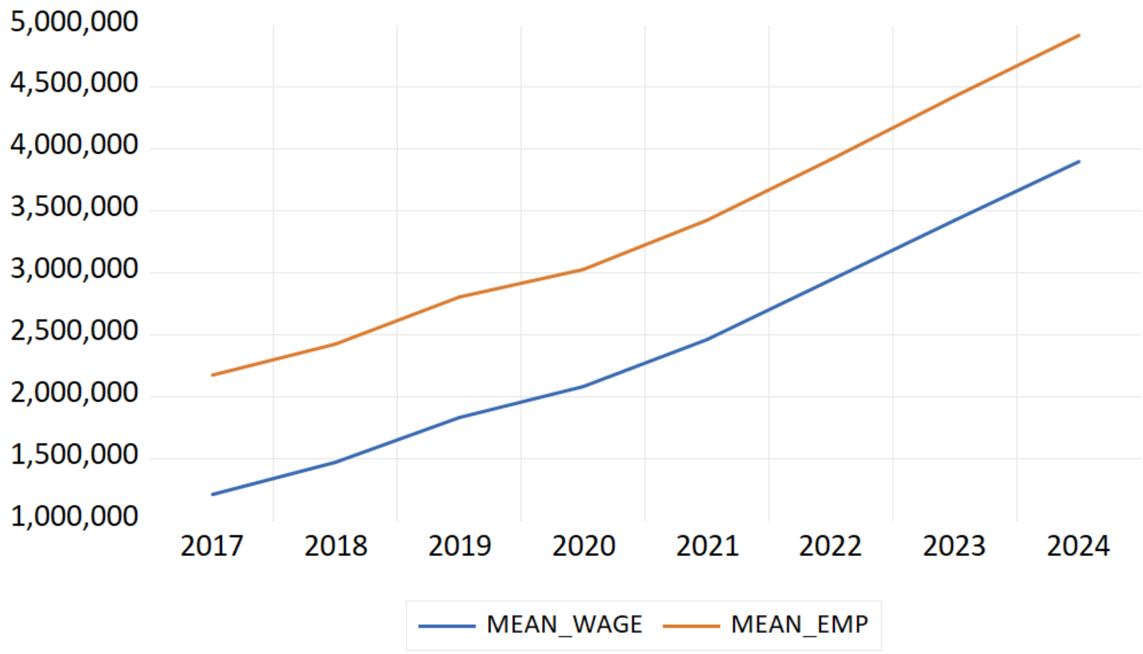

Figure 1: Average time dynamics of wage and employment indicators (2017–2024)

In the next stage, wage and employment variables were taken in level form, and an initial regression model was estimated without conducting stationarity tests. The results of this regression analysis are presented in Table 2.

The results in Table 2 indicate that the regression coefficients are statistically significant and that the model exhibits a relatively high coefficient of determination. However, since the model is estimated using non-stationary level variables with a pronounced trend, these results do not exclude the possibility of spurious regression. Therefore, the findings should be interpreted only as preliminary estimates, and in the subsequent stage, the stationarity properties of the variables were examined separately.

Table 2: Regression results without stationarity testing

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	634 299.8	19 025.62	33.33924	0.0000
WAGE	0.032447	0.007686	4.221604	0.0056
Indicator			**Value**	
Number of observations			8	
Coefficient of determination (R^2)			0.7481	
AdjustedR^2			0.7062	
F-statistic			17.8219	
Prob(F-statistic)			0.0056	
Durbin–Watson statistic			1.6378	

Source: Authors' calculations based on open data from the State Statistics Committee of the Republic of Uzbekistan using EViews software.

To assess the stationarity properties of wage and employment indicators, panel unit root tests were conducted. The test results for both level and first-differenced forms are presented in Table 3.

Table 3: Panel unit root test results for wage and employment indicators

Variable	Specification	Test Method	Statistic Value	p-value	Conclusion
Wage	Level	Levin–Lin–Chu (LLC)	9.8503	1.000	Non-stationary
		Im–Pesaran–Shin (IPS)	9.1108	1.000	Non-stationary
ΔWage	First difference	Levin–Lin–Chu (LLC)	−4.8123	0.000	Stationary
Employment	Level	Levin–Lin–Chu (LLC)	−8.5959	0.000	Stationary
		Im–Pesaran–Shin (IPS)	−1.6497	0.049	Stationary

Source: Authors' calculations based on open data from the State Statistics Committee of the Republic of Uzbekistan using EViews software.

The results in Table 3 indicate that the wage variable is non-stationary in level form according to all major panel unit root tests. However, after first-order differencing, it becomes stationary at least according to the Levin–Lin–Chu (LLC) test.

The employment variable, on the other hand, demonstrates stationarity in level form according to some tests (LLC and IPS). However, due to inconsistencies across different test results, the variable was also differenced as a precaution and included in the subsequent regression analysis in its stationary form.

This situation highlights that in panel data analysis, stationarity should not be assessed based on a single test only, but rather through a combination of multiple test results.

Based on the stationarity test outcomes, both wage and employment variables were transformed into first differences to achieve stationarity, and a new panel regression analysis was conducted. The results of this regression are presented in Table 4.

Table 4: Panel regression results based on differenced variables

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	−17 676.14	6 536.87	−2.7041	0.0081
ΔWAGE	0.0656	0.0154	4.2623	0.0000
Indicator		**Value**		
R^2		0.1591		
Adjusted R^2		0.1504		
F-statistic		18.1672		
Prob(F-statistic)		0.000047		
Durbin–Watson		1.9746		
Number of observations		98		
Number of cross-sections		14		
Number of periods		7		

Source: Authors' calculations.

The results presented in Table 4 show a substantial difference compared to the initial regression where stationarity was not taken into account. In particular, the reduction in the coefficient of determination (R^2), as well as changes in both the magnitude and statistical significance of the regression coefficients, indicate that the previously observed high model fit was largely driven by common trend effects rather than a true economic relationship between the variables.

At the same time, since this regression is estimated using stationary data, it is considered methodologically appropriate and statistically reliable from an econometric perspective.

In the next stage, the regression results obtained under non-stationary and stationary specifications are compared. The results of this comparison are presented in Table 5.

Table 5: Comparison of regression results with and without stationarity adjustment

Indicator	Level Regression (Table 2)	Differenced Regression (Table 4)
Model specification	EMP = f(WAGE)	ΔEMP = f(ΔWAGE)
Independent variable	WAGE	ΔWAGE
Coefficient	0.0324	0.0656
p-value	0.0056	0.0000
R^2	0.7481	0.1591
Prob(F-statistic)	0.0056	0.000047
Number of observations	8	98
Stationarity status	Not accounted for	Accounted for
Methodological assessment	Possible spurious regression	Methodologically valid

Source: Authors' calculations.

The comparison presented in Table 5 indicates that, although the regression estimated without stationarity testing exhibits a relatively high coefficient of determination and seemingly strong statistical performance, it may be methodologically incorrect. In contrast, the regression based on stationary variables produces less pronounced goodness-of-fit results; however, it fully satisfies econometric theory requirements and is therefore considered more reliable.

This finding empirically confirms the critical methodological importance of accounting for stationarity in regression analysis.

CONCLUSION

This study analyzed the relationship between wage and employment indicators using panel data. During the analysis, the time-series properties of the variables were examined, and it was found that both wage and employment series are non-stationary in levels but become stationary after first-order differencing. This result highlights the importance of the data preprocessing stage prior to model estimation in regression analysis.

Although the initial regression model estimated without considering stationarity produced a high coefficient of determination, further analysis showed that this fit was largely driven by a common trend effect. In contrast, the regression model estimated using stationary (differenced) data exhibited a lower goodness-of-fit but is methodologically correct and statistically reliable for inference. This confirms that model quality in data analysis should not be evaluated solely based on a high R^2 value, but also on whether the statistical properties of the data are properly addressed.

Based on the results obtained, the following recommendations are proposed:

First, stationarity testing should be considered a mandatory step prior to regression analysis in both time-series and panel data applications.

Second, in practice, preference should be given to models that satisfy statistical assumptions and are theoretically consistent, rather than models that only produce visually appealing results.

Third, future research in this area could be extended by incorporating automated testing procedures, dynamic panel data models, and machine learning-based econometric approaches.

REFERENCES

1. Spurious regressions in econometrics (1974). Granger C.W.J., Newbold P. Journal of Econometrics.
2. Applied Econometric Time Series (2015). Enders W. Wiley.
3. Econometric Analysis of Cross Section and Panel Data (2010). Wooldridge J.M. MIT Press.
4. Basic Econometrics (2009). Gujarati D.N., Porter D.C. McGraw-Hill.
5. Time Series Analysis (1994). Hamilton J.D. Princeton University Press.
6. Nonstationary Panels, Panel Cointegration, and Dynamic Panels (2000). Baltagi B.H. Advances in Econometrics.
7. Testing for unit roots in heterogeneous panels (2003). Im K.S., Pesaran M.H., Shin Y. Journal of Econometrics.
8. Panel unit root tests in the presence of cross-section dependence (2007). Pesaran M.H. Journal of Applied Econometrics.
9. Forecasting, structural time series models and the Kalman filter (2012). Harvey A.C. Cambridge University Press.
10. Introduction to Time Series and Forecasting (2016). Brockwell P.J., Davis R.A. Springer.
11. Applied Predictive Modeling (2018). Kuhn M., Johnson K. Springer. (data preprocessing va regressiya uchun)
12. Machine Learning Mastery with Time Series (2020). Brownlee J. Machine Learning Mastery.
13. Iqtisodiy-statistik tahlil asoslari (2020). O‘zbekiston Respublikasi Davlat statistika qo‘mitasi.
14. O‘zbekiston Respublikasi Davlat statistika qo‘mitasining rasmiy ochiq ma’lumotlari (2017–2024). https://stat.uz