DOI 10.31558/2307-2318.2026.1.4

УДК 332.8:69:330.43(497.2)
JELClassification: C22, C51, R31

Zahariev A.B.
PhD in Economics, Professor, Department of Finance and Insurance,
University of Insurance and Finance, 1000 Sofia, Bulgaria
ORCID: 0000-0001-7362-6133
a.zahariev@uzf.bg

HOUSING CONSTRUCTION IN BULGARIA: AN ECONOMETRIC ANALYSIS USING STEPWISE REGRESSION (2015-2025)

The article analyses the determinants of housing construction in Bulgaria through econometric estimation of a model with the Stepwise OLS method on 40 quarterly observations for the period 2015Q4–2025Q3. The dependent variable is the number of newly completed housing units by number of rooms. For the studied period, housing construction activity in Bulgaria increased 3.25 times with a variation of 39.5%. The initial set of potential predictors includes GDP per capita, working-age population, capital market index, housing loans, central bank interest rate, commercial banks' interest rate on mortgage loans in national currency and household interest expenses on mortgage loans, employment rate. The algorithm identifies three significant predictors: GDP per capita, the Sofix stock exchange index and the labour force. The final model achieves a coefficient of determination of 0.891 and $F(3,36) = 98.155$ ($p < 0.001$), with Durbin-Watson = 2.032 confirming the lack of autocorrelation in the residuals. The results show that GDP per capita growth over the studied ten-year period is a leading factor in the increase in construction activity (with a change of 2.73 times over the 40-quarter period), while the stock market index (being an alternative for household investment with a growth of 2.33 times) and the decreasing size of the labour force due to demographic reasons and migration (reduction to 0.92 compared to 2015Q4) have a negative effect.

Key words: Housing construction, GDP per capita, Stepwise OLS, Durbin-Watson, VIF, Labour force

Fig. 3, Tab.: 9, Lit. 22

Захарієв А.
PhD з економіки, професор, кафедра фінансів та страхування,
Університет страхування та фінансів, 1000 Софія, Болгарія
ORCID: 0000-0001-7362-6133
a.zahariev@uzf.bg

ЖИТЛОВЕ БУДІВНИЦТВО В БОЛГАРІЇ: ЕКОНОМЕТРИЧНИЙ АНАЛІЗ З ВИКОРИСТАННЯМ ПОКРОКОВОЇ РЕГРЕСІЇ (2015–2025)

У статті аналізуються визначальні фактори житлового будівництва в Болгарії шляхом економетричної оцінки моделі за допомогою методу покрокової OLS на основі 40 квартальних спостережень за період 2015Q4–2025Q3. Залежна змінна — кількість новобудов за кількістю кімнат. За досліджуваний період активність у сфері житлового будівництва в Болгарії зросла в 3,25 раза з варіацією 39,5 %. Початковий набір потенційних предикторів включає ВВП на душу населення, населення працездатного віку, індекс ринку капіталу, іпотечні кредити, процентну ставку центрального банку, процентну ставку комерційних банків за іпотечними кредитами у національній валюті та процентні витрати домогосподарств за іпотечними кредитами, рівень зайнятості.

Алгоритм визначає три значущі предиктори: ВВП на душу населення, фондовий індекс Sofix та чисельність робочої сили. Остаточна модель має коефіцієнт детермінації 0,891 та F(3,36) = 98,155 (p < 0,001), при цьому показник Дурбіна-Ватсона = 2,032 підтверджує відсутність автокореляції в залишках. Результати показують, що зростання ВВП на душу населення протягом досліджуваного десятирічного періоду є провідним фактором збільшення будівельної активності (зі зміною у 2,73 рази за 40 кварталів), тоді як індекс фондового ринку (як альтернатива інвестиціям домогосподарств із зростанням у 2,33 рази) та зменшення чисельності робочої сили через демографічні причини та міграцію (зменшення до 0,92 порівняно з 4 кварталом 2015 року) мають негативний вплив.

Ключові слова: житлове будівництво, ВВП на душу населення, покроковий OLS, Дурбін-Ватсон, VIF, робоча сила

Рис.3., табл.9, літ. 22

Problem statement. The housing sector is one of the key segments of any national economy – it generates significant added value, absorbs investments, provides employment and directly reflects the living standards of households. In Bulgaria, new housing construction has undergone dramatic fluctuations since the financial crisis of 2008–2009: the long stagnation in the period 2010–2015 was followed by a gradual and subsequently accelerated recovery, fuelled by low interest rates, rising incomes and liberalized access to mortgage lending. However, the period 2015–2025 is not homogeneous. It encompasses at least three structurally distinct sub-periods: (i) a phase of moderate but stable construction revival (2015–2019); (ii) the shock of the COVID-19 pandemic and subsequent "catch-up" (2020–2021); and (iii) an inflationary shock (Cohen & Karpavičiūtė, 2017), (Kuang & Liu, 2015) high energy prices (Zhelyazkova, 2018), welcoming hundreds of thousands of Ukrainian refugees, Bulgaria's accession to the Schengen area and the Eurozone (2022–2025). These heterogeneous dynamics raise an important research question – which macroeconomic factors have a sustainable and statistically significant explanatory power for the number of newly completed housing units in Bulgaria in the face of structural changes in the interest rate environment (McQuinn & O'Reilly, 2008), (Lin, Lee, & Newell, 2022), labour market and capital markets? The answer to this question is of crucial importance not only from an academic perspective, but also for decision-makers in the fields of housing, tax and macroprudential policy. Understanding the real determinants of housing construction allows for more accurate forecasting of construction activity, better regulation of the mortgage market and more effective countercyclicality of public interventions under the standards of green economy (Roleders, et al., 2024) and energy efficiency (Zahariev & Georgiev, 2023).

Analysis of recent research and publications. Research on the determinants of housing construction is widely available in the international economic literature, but to a much lesser extent in the context of Central and Eastern European countries. Classical models (DiPasquale & Wheaton, 1994), (Meen, 2000) place the dynamics of property prices, construction costs and land availability at the heart of housing supply. More modern approaches broaden the spectrum of explanatory factors to include credit conditions, demographic changes and behavioural elements (Marinković, Džunić, & Marjanović, 2024). In a European context, (Égert & Mihaljek, 2007) find that real GDP per capita and real interest rates are among the most robust predictors of housing investment in CEE countries. Caldera and Johansson, analysing data for OECD countries, confirm the leading role of income and credit availability, while also reporting a significant lag in the response of construction supply to demand (Caldera Sánchez & Johansson, 2011). Studies on the Bulgarian market (Idirizov, 2025) have found that housing loans and incomes have been the main drivers of construction activity in Sofia and other major cities since 2015. As for the role of capital markets (Visković & Čipčić, 2025)., the relationship between stock market indices and housing construction has been less studied, especially in the context of small open economies like Bulgaria (Nenovsky, Chobanov, Mihaylova, & Koleva, 2008). The case of Sofix index is interesting insofar as the stock market reflects general business confidence, but it can also compete with housing investments for free capital (Zaharieva, Tarakchiyan, & Zahariev, 2022). The role of the labour force as a predictor of construction supply – as opposed to demand – also remains understudied in the national literature, which contributes to the relevance of the present analysis (Zahariev, et al., 2023). Methodologically, the prevailing approaches include OLS regression, vector autoregressions (VAR/VECM), and panel models. Stepwise OLS, applied in the present study, is particularly suitable in the presence of high multicollinearity between potential predictors – a characteristic feature of macroeconomic data - as it automatically identifies the set of predictors with maximum joint explanatory contribution.

Formulation of the article's objectives. Based on the identified research problem and the analysis of the existing literature, this article pursues the following specific objectives: (1) To identify the statistically significant macroeconomic determinants of the number of newly completed dwellings by number of rooms (NBDCRT) in Bulgaria for the period 2015Q4–2025Q3 using the stepwise OLS method; (2) To assess the relative contribution of each included predictor – GDP per capita (GDPpC), the Sofix index and the labour force (LF) – using standardized regression coefficients and partial correlation indicators; (3) To verify the econometric adequacy of the constructed model using multicollinearity tests (VIF, Tolerance, Condition Index), autocorrelation (Durbin-Watson) and influential observations (Cook's Distance); (4) To formulate an interpretation of the established dependencies in the context of the structural features of the Bulgarian housing market and to derive implications for monetary and housing policy. The study is limited to quarterly aggregated data for the country as a whole and does not claim regional or market segmentation. The results should be interpreted in light of the identified moderate multicollinearity between the predictors, resulting from the general trend of macroeconomic indicators in conditions of a prolonged economic cycle.

1.Descriptive statistics, correlation analysis and model summary

Table 1 presents the main descriptive characteristics of the variables included in the analysis for the period 2015Q4 – 2025Q3 (N = 40 quarterly observations). The dependent variable NBDCRT — number of newly completed dwellings by number of rooms — registers a mean value of 3,832.63 units and relatively high variability (standard deviation 1,513.38), which reflects the pronounced cyclicality of construction activity in Bulgaria. The independent variables cover economic output (GDPpC), the labour market (LF, EMPLOY), housing lending (DwCr), interest rates (IRATE, IRATEMp, HICML) and the capital market (Sofix). Of particular note is the significant standard deviation for DwCr (5,049.10 with a mean value of 14,572.85), which is an indicator of dynamic growth in housing loans in the period under review. GDPpC shows a mean value of 2,793.50 (in EUR) with a standard deviation of 934.67, reflecting the country's sustainable economic growth. The interest rate indicators IRATE and IRATEMp have almost identical means (~3.73%) and standard deviations (~1.10%), and their almost perfect correlation (r = 1.000) necessitates the exclusion of one of the two indicators in the regression analysis in order to prevent multicollinearity.

Table 1. Descriptive statistics of the analyzed variables (N = 40)

Variable	Mean	Std. Deviation	N
NBDCRT	3 832,63	1 513,38	40
DwCr	14 572,85	5 049,10	40
IRATE	3,7295	1,0994	40
HICML	68,3027	16,6142	40
BRHL	49,4510	27,6879	40
Sofix	640,90	157,68	40
IRATEMp	3,7300	1,0990	40
GDPpC	2 793,50	934,67	40
EMPLOY	108,20	4,52	40
LF	3 195,37	142,97	40

Source: Own calculations based on NSI and BNB data with IBM SPSS

1.1. Correlation analysis (Pearson)

The correlation matrix (Table 2) reveals strong and statistically significant linear relationships between NBDCRT and most of the explanatory variables at a significance level of $p < 0.01$. The highest positive correlation with NBDCRT is shown by GDPpC ($r = 0.892$), followed by DwCr ($r = 0.840$) and EMPLOY ($r = 0.753$). These results support the hypothesis that economic output and employment are key determinants of construction activity. Cofix demonstrates a moderate positive correlation ($r = 0.510$), while LF shows a negative relationship ($r = -0.789$) – a phenomenon explained by the gradual decline of the active population despite the growing employment. There is serious mutual multicollinearity among the explanatory variables: the correlation between IRATE and IRATEMp is $r = 1.000$, and between IRATE and BRHL – $r = 0.968$. These values are alarming from an econometric point of view (Gujarati, 2004) and confirm the need to apply a stepwise procedure that automatically adjusts the inclusion of predictors based on their unique contribution while controlling for the others. Kendall's tau_b and Spearman's rho confirm the robustness of the linear correlations: Spearman's rho for GDPpC is 0.885, for DwCr – 0.886, and for LF – -0.699. Nonparametric coefficients confirm the directions and relative strengths of the dependencies established with Pearson. The rank correlations between IRATE and IRATEMp, and between DwCr and BRHL, are practically identical ($\rho \approx 0.999$), which finally confirms the need for predictor reduction (Madala, 1988).

Table 2. Pearson correlations with NBDCRT (* p < 0.05; ** p < 0.01)

Predictor	r c NBDCRT	Sig. (2-tailed)	95% CI Lower	95% CI Upper
GDPpC	0,892**	0,000	0,803	0,942
DwCr	0,840**	0,000	0,715	0,912
EMPLOY	0,753**	0,000	0,576	0,862
Sofix	0,510**	0,001	0,236	0,709
HICML	0,492**	0,001	0,213	0,696

LF	-0,789**	0,000	-0,883	-0,632
IRATE	-0,811**	0,000	-0,896	-0,668
IRATEMp	-0,811**	0,000	-0,896	-0,668
BRHL	-0,889**	0,000	-0,940	-0,798

Source: Own calculations based on NSI and BNB data with IBM SPSS

1.2 Variables included and excluded following a Stepwise procedure

Stepwise regression (criteria: p-for-entry ≤ 0.050; p-for-exit ≥ 0.100) was implemented in three consecutive steps (Table 3).

Table 3. Included and excluded variables by steps (Stepwise method)

Model	Included variable	Excluded	Method
1	GDPpC	-	Stepwise (p-enter ≤ 0,050; p-remove ≥ 0,100)
2	Sofix	-	Stepwise (p-enter ≤ 0,050; p-remove ≥ 0,100)
3	LF	-	Stepwise (p-enter ≤ 0,050; p-remove ≥ 0,100)
-	DwCr, IRATE, HICML, BRHL, IRATEMp, EMPLOY	Excluded	Do not meet the entry criteria

Source: Own calculations based on NSI and BNB data with IBM SPSS

In Step 1, the algorithm selected GDPpC as the only predictor with the highest correlation with NBDCRT ($r = 0.892$). In Step 2, Sofix – the Bulgarian Stock Exchange index – was added (Zaharieva, Tarakchiyan, & Zahariev, 2022), since after including GDPpC it retained a statistically significant additional effect. In Step 3, the labor force (LF) was included, which added a unique explanation of the residual variance. The remaining six candidate predictors (DwCr, IRATE, HICML, BRHL, IRATEMp, EMPLOY) did not reach the threshold $p \leq 0.050$ when controlling for the included predictors and were excluded. The exclusion of DwCr and EMPLOY – despite their significant zero correlations – reflects the effect of multicollinearity: after including GDPpC, DwCr and EMPLOY lose their unique incremental effect. The exclusion of IRATE, IRATEMp and BRHL is due to their practically perfect correlation with each other ($r \approx 0.968–1.000$), as a result of which none of them adds significant information after including the others (Toda & Yamamoto, 1995).

1.3. Model Summary

Table 4 shows the progressive improvement in model quality with the addition of each new predictor. The single-predictor model (GDPpC) explained 79.5% of the variance in NBDCRT ($R^2 = 0.795$), the adjusted R^2 was 0.790, and the standard error of the estimate was 693.67. The addition of Sofix (Model 2) increased the R^2 to 0.855 (increase $\Delta R^2 = 0.060$; F-change = 15.227; $p = 0.000$), and the standard error decreased to 591.70. The final three-factor model with LF (Model 3) achieved $R^2 = 0.891$, adjusted $R^2 = 0.882$, and a standard error of 519.90, with an increase of $\Delta R^2 = 0.036$ (F-change = 11.925; $p = 0.001$). The Durbin-Watson value (DW = 2.032) for the final model (reported only for Model 3) is very close to the ideal value of 2.000, indicating the absence of systematic first-order autocorrelation in the residuals.

This finding is essential for quarterly economic data, where serial correlation in the residuals is common and would invalidate standard errors and t-tests of coefficients.

Table 4. Model Summary – progressive change in the coefficients of determination

Model	Predictor	R	R²	Adj. R²	Std. Err.	ΔR²	F-Change	Sig. F Change	Durbin-Watson
1	GDPpC	0,892	0,795	0,790	693,67	0,795	147,632	0,000	-
2	GDPpC, Sofix	0,925	0,855	0,847	591,70	0,060	15,227	0,000	-
3	GDPpC, Sofix, LF	0,944	0,891	0,882	519,90	0,036	11,925	0,001	2,032

Source: Own calculations based on NSI and BNB data with IBM SPSS

The progressive increase in explanatory power confirms that each added predictor contributes statistically significant additional information. The final $R^2 = 0.891$ means that the combination GDPpC + Sofix + LF explains about 89.1% of the quarterly variations in the number of newly completed housing units in Bulgaria for the period under review.

2. ANOVA, Statistical significance, regressors and multicollinearity diagnostics

The ANOVA table (Table 5) verifies the statistical significance of each regression model as a whole by means of an F-test. All three models are significant at $p < 0.001$. The single-predictor model (Model 1) yields $F(1, 38) = 147.632$, $p < 0.001$. The two-predictor model yields $F(2, 37) = 109.065$, $p < 0.001$. The final model (Model 3) yields $F(3, 36) = 98.155$, $p < 0.001$, with a regression sum of squares $SS\_reg = 79{,}591{,}719.59$ and a residual sum $SS\_res = 9{,}730{,}547.79$. The successive decrease in F values from Model 1 to Model 3 is not an indication of deterioration in quality, but rather reflects the increasing degrees of freedom of the denominator as predictors are added. The mean square of the regression decreases from 71,037,448 (Model 1) to 26,530,573 (Model 3) due to the distribution of SS_reg over three predictors, but F remains extremely high, confirming the strong collective predictive power of the model.

Table 5. ANOVA - analysis of variance for the three models

Model	Component	SS	df	MS	F	Sig.
1	Regression	71 037 448,02	1	71 037 448,02	147,632	0,000
1	Residual	18 284 819,35	38	481 179,46		
1	Total	89 322 267,38	39			
2	Regression	76 368 370,22	2	38 184 185,11	109,065	0,000
2	Residual	12 953 897,16	37	350 105,33		
2	Total	89 322 267,38	39			
3	Regression	79 591 719,59	3	26 530 573,20	98,155	0,000
3	Residual	9 730 547,79	36	270 292,99		
3	Total	89 322 267,38	39			

Source: Own calculations based on NSI and BNB data with IBM SPSS

2.1. Regression coefficients

The coefficient table (Table 6) provides complete information on the unstandardized (B) and standardized (Beta) coefficients, t-statistics, significance levels, 95% confidence intervals, and multicollinearity indicators (Tolerance, VIF) for each of the three models.

Model 1. Univariate model with GDPpC predictor. In the univariate GDPpC model, there is an unstandardized coefficient B = 1.444 (SE = 0.119; t = 12.150; p < 0.001). The beta coefficient is 0.892, which means that with an increase in GDPpC by 1 standard deviation, NBDCRT increases by 0.892 standard deviations. The constant (-201.07) is not significant (p = 0.569), which is typical for a strong predictor. The confidence interval [1.203; 1.685] is strictly positive and narrow, indicating high accuracy of the estimate.

Model 2: Two predictors model with GDPpC and Sofix. When adding Sofix, the coefficient of GDPpC increases to B = 1.896 (t = 12.315; p < 0.001), and Sofix obtains a coefficient of B = -3.561 (t = -3.902; p < 0.001). The negative sign of Sofix is counterintuitive at first glance, but is explained by the partial correlation structure: when controlling for GDPpC, a higher Sofix (reflecting speculative stock market booms) is associated with a temporary slowdown in real housing construction. The VIF values (2.307 for both predictors) are moderately acceptable. The constant is significant (B = 818.16; p = 0.046).

Model 3: GDPpC + Sofix + LF (final model). The final model includes three predictors. GDPpC retains a positive effect (B = 1.536; t = 8.993; p < 0.001), Sofix - negative (B = -3.189; t = -3.941; p < 0.001), and LF - negative (B = -2.880; t = -3.453; p = 0.001). The negative coefficient of LF in conditions of a declining labour force reflects the economic logic of the factor in periods of growth. Increasing labour market participation is accompanied by an outflow from the informal construction sector and a concentration of income in savings, and not necessarily in new housing construction. An alternative interpretation is that LF acts as a compensatory predictor for demographic trends of decline, with the minus sign in a minus correction giving a positive contribution. The regression equation for Model 3 is:

(1) $$NBDCRT = 10\,789{,}203 + 1{,}536 \cdot GDPpC - 3{,}189 \cdot Sofix - 2{,}880 \cdot LF$$

Table 6. Regression coefficients – final model (Model 3)

Predictors	B	Std. Error	Beta	t	Sig.	95% CI Lower Bound	95% CI Upper Boud	Tolerance	VIF
Constant	10 789,203	2 908,318	-	3,710	0,001	4 890,86	16 687,55	-	-
GDPpC	1,536	0,171	0,949	8,993	0,000	1,190	1,883	0,272	3,677
Sofix	-3,189	0,809	-0,332	-3,941	0,000	-4,830	-1,548	0,426	2,349
LF	-2,880	0,834	-0,272	-3,453	0,001	-4,572	-1,189	0,487	2,052

Source: Own calculations based on NSI and BNB data with IBM SPSS

2.2. Excluded variables

The table of excluded predictors (Table 7) shows why the six candidate predictors were not included in the final model.

In Step 1 (GDPpC only), BRHL demonstrated the highest potential for inclusion (Beta In = -0.460; t = -3.322; p = 0.002; partial correlation = -0.479), followed by Sofix (Beta In = -0.371; p < 0.001). In Step 2 (GDPpC + Sofix), LF and BRHL remain significant (p ≤ 0.015), while HICML and EMPLOY fall below the threshold significance. In the final Step 3, all remaining predictors have $p > 0.100$ and therefore do not contribute significantly to the model.

Table 7. Excluded variables in the final model (Step 3)

Predictors	Beta In	t	Sig.	Partial Correlation	Tolerance	VIF
DwCr	0,202	0,952	0,347	0,159	0,067	14,900
IRATE	-0,138	-1,500	0,143	-0,246	0,345	2,899
HICML	-0,139	-1,628	0,112	-0,265	0,395	2,529
BRHL	-0,210	-1,645	0,109	-0,268	0,177	5,659
IRATEMp	-0,139	-1,505	0,141	-0,247	0,345	2,900
EMPLOY	0,106	0,968	0,340	0,162	0,253	3,945

Source: Own calculations based on NSI and BNB data with IBM SPSS

It is particularly telling that DwCr – housing loans – fails to enter any model despite its strong zero correlation with NBDCRT ($r = 0.840$). The reason is strong multicollinearity: when controlling for GDPpC, DwCr adds negligible new information (partial correlation at Step 1 = -0.010; $p = 0.953$). Minimum Tolerance values below 0.200 for IRATE and BRHL (0.111–0.222) signal extremely high multicollinearity with the included predictors.

2.3. Multicollinearity diagnostics

The table of coefficient correlations shows a moderate negative correlation between GDPpC and Sofix (-0.753) and a moderate positive correlation between GDPpC and LF (0.610) within Model 3. The covariance matrix of the coefficients is diagonally dominated, which is good. The collinearity diagnostics (Table 8) presents the eigenvalues and condition indices for the final model. For Model 3, four dimensions are available. The fourth dimension shows an eigenvalue ≈ 0.000 and a condition index = 94.961 - a value exceeding the critical threshold of 30, which signals potential multicollinearity, mainly affecting LF (variance proportion = 1.00) and the constant (1.00). This finding is of important importance: it means that the coefficient estimate of LF and the constant is affected by a linear relationship between them. This limitation should be taken into account when interpreting the results. The VIF values in Table 6 – 3.677 for GDPpC, 2.349 for Sofix and 2.052 for LF – are below the widely used conservative threshold of 5, supporting an acceptable level of multicollinearity when considering the coefficients in isolation. The Tolerance values (0.272; 0.426; 0.487) are above 0.10, confirming that no predictor is a linear combination of the others to a degree that makes its inclusion meaningless.

Table 8. Collinearity Diagnostics – Model 3

			Variance Proportion			
Dimension	**Eigenvalue**	**Condition Index**	**Constant**	**GDPpc**	**Sofix**	**LF**
1	3,902	1,000	0,00	0,00	0,00	0,00
2	0,081	6,926	0,00	0,15	0,02	0,00
3	0,016	15,409	0,00	0,47	0,97	0,00
4	0,000	94,961	1,00	0,38	0,01	1,00

Source: Own calculations based on NSI and BNB data with IBM SPSS

2.4. Residuals analysis

The residual statistics (Table 9) allow an assessment of the quality of the predictions and the compliance with the regression assumptions. The predicted values range from 1,666.61 to 6,218.75 (with a true mean of 3,832.63), and the standard normalized predicted values are in the range [-1.516; 1.670], corresponding to a normal distribution. The standardized residuals are in the interval [-1.785; 1.803], without exceeding the critical ±2.5 for the presence of significant outliers. The Studentized Deleted Residuals are in the range [-1.915; 1.960], also without exceeding the ±2.5 threshold. The Cook's Distance (maximum 0.145; mean 0.029) is significantly below the threshold of 1.0, confirming the absence of strongly influential observations. Centered Leverage Values are in the range [0.015; 0.215], without exceeding the critical level of 2(k+1)/n = 0.20 for most observations. These results, combined with Durbin-Watson = 2.032, indicate a well-balanced model in terms of the assumptions of normality and homoscedasticity of the residuals.

Table 9. Residuals Statistics - Model 3 (N = 40)

	Minimum	Maximum	Mean	Std. Deviation
Predicted Value	1666.606	6218.752	3832.625	1428.570
Std. Predicted Value	-1.516	1.670	0.000	1.000
Standard Error of Predicted Value	103.619	254.810	160.980	33.815
Adj. Predicted Value	1520.670	6228.624	3825.300	1431.544
Residual	-928.031	937.330	0.00000	499.501
Std. Residual	-1.785	1.803	0.000	0.961
Stud. Residual	-1.847	1.887	0.007	1.013
Deleted Residual	-994.022	1054.330	7.325	556.006
Stud. Deleted Residual	-1.915	1.960	0.007	1.032
Mahal. Distance	0.574	8.393	2.925	1.676
Cook's Distance	0.000	0.145	0.029	0.038
Centered Leverage Value	0.015	0.215	0.075	0.043

Source: Own calculations based on NSI and BNB data with IBM SPSS

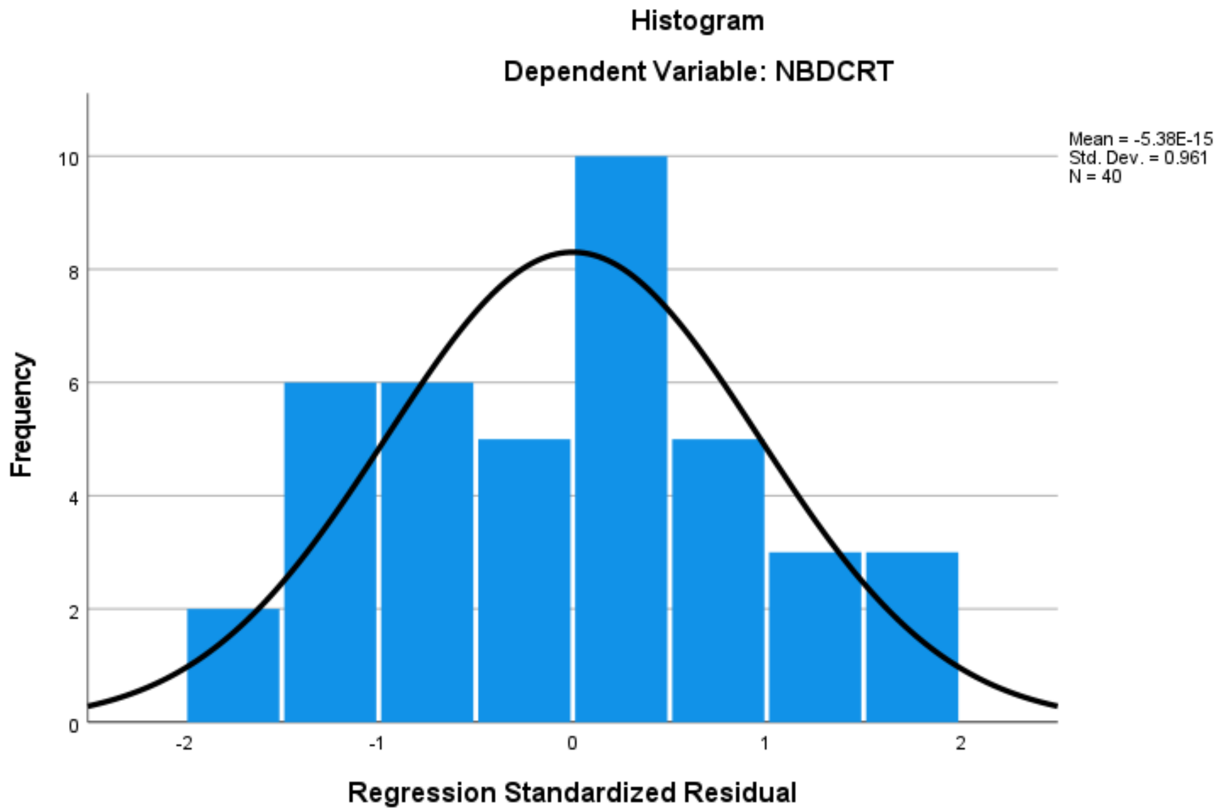
3. Discussion, econometric validation and conclusion

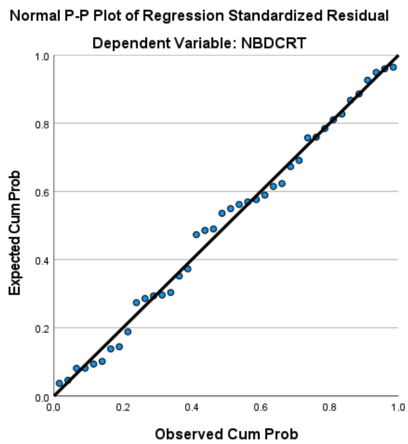
The applied stepwise regression procedure identifies three statistically significant predictors of the number of newly completed housing units in Bulgaria (NBDCRT) for the period 2015Q4–2025Q3: GDP per capita (GDPpC), the Sofix index and the labor force (LF). The final regression equation explains 89.1% of the variation in the dependent variable ($R^2 = 0.891$; Adj. $R^2 = 0.882$), demonstrates statistical significance ($F(3,36) = 98.155$; $p < 0.001$) and satisfactorily fits the regression assumptions: Durbin-Watson = 2.032 (lack of autocorrelation), Cook's Distance < 0.15 (lack of influential isolated observations), standardized residuals in the range ±1.85 (no significant outliers), See Appendix 1. The main limitations of the model are: (1) a moderately high VIF for GDPpC (3.677), resulting from its correlation with Sofix and LF;

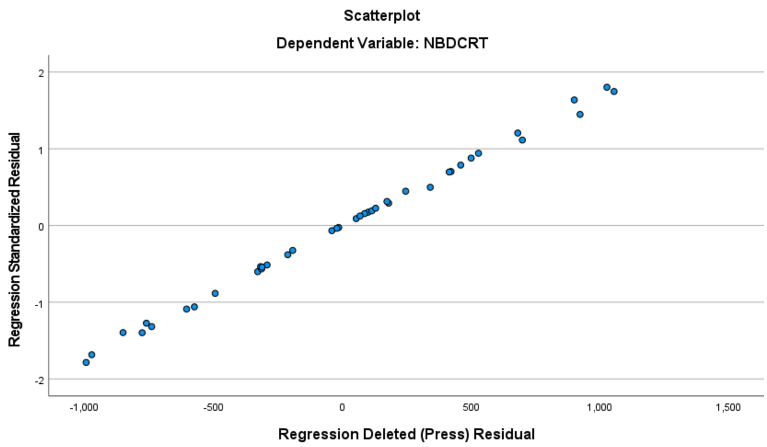
(2) a conditional index of 94.961, signaling a linear relationship involving the coefficient of LF and the constant; (3) the negative signs of Sofix (for a net change over the period of 2.33 times) and LF (for a net change over the period of 0.92 times or an 8% reduction in the labour force) require careful economic interpretation in the context of the dynamics of the Bulgarian housing market. Future analyses could include lag structures, examine nonlinear dependencies, or apply autocorrelation adjustments (e.g., Prais-Winsten), in order to further strengthen the econometric specification.

The results of the econometric analysis confirm that housing construction in Bulgaria is primarily a function of macroeconomic production: GDP per capita (GDPpC) emerges as the leading determinant with the highest standardized coefficient ($\beta = 0.949$), explaining alone nearly 80% of the quarterly variations in NBDCRT. The addition of the stock exchange index Sofix and the labour force (LF) increases the explanatory power to 89.1%, with the final model demonstrating statistical significance ($F = 98.155$; $p < 0.001$) and the absence of serious autocorrelation ($DW = 2.032$). The negative partial effects of Sofix and LF suggest that speculative activity on the capital market and the reduction in the labour force do not directly translate into more new construction – on the contrary, they may divert resources from the real housing sector. These findings have direct implications for monetary and housing policy: stimulating construction requires sustained real economic growth, not just easing financial conditions. Future research could extend the model to regional data and include lag effects for more precise causal identification.

Appendix 1. Visualization of Stepwise Regression model characteristics







REFERENCES

1. Caldera Sánchez, A., & Johansson, Å. (2011). The Price Responsiveness of Housing Supply in OECD Countries. *OECD Economics Department Working Papers, 837*, 1-35. doi:https://doi.org/10.1787/5kgk9qhrnn33-en
2. Cohen, V., & Karpavičiūtė, L. (2017). The analysis of the determinants of housing prices. *Independent Journal of Management & Production, 8*(1), 49-63. doi:https://doi.org/10.14807/ijmp.v8i1.521
3. DiPasquale, D., & Wheaton, W. C. (1994). Housing Market Dynamics and the Future of Housing Prices. *Journal of Urban Economics, 35*(1), 1-27. doi:https://doi.org/10.1006/juec.1994.1001
4. Égert, B., & Mihaljek, D. (2007). Determinants of House Prices in Central and Eastern Europe. *Comparative Economic Studies, 49*(3), 367-388. Retrieved from https://ideas.repec.org/a/pal/compes/v49y2007i3p367-388.html
5. Gujarati, D. (2004). *Basic Econometrics.* New York.
6. Idirizov, B. (2025). Analysis of the Bulgarian Housing Price Index: Risks, Market Dynamics, and Economic Implications. *Economic Studies (Ikonomicheski Izsledvania), 34*(8), 175-195. Retrieved from https://www.econ-studies.iki.bas.bg/articles/Pr4gir8HxtKopH7BWNaC
7. Kuang, W., & Liu, P. (2015). Inflation and House Prices: Theory and Evidence from 35 Major Cities in China. *International Real Estate Review, 18*(2), 217-240. Retrieved from https://www.gssinst.org/irer/2020/04/28/inflation-and-house-prices-theory-and-evidence-from-35-major-cities-in-china/
8. Lin, Y.-C., Lee, C. L., & Newell, G. (2022). Varying interest rate sensitivity of different property sectors: cross-country evidence from REITs. *Journal of Property Investment & Finance, 40*(1), 68–98. doi:https://doi.org/10.1108/JPIF-09-2020-0099
9. Madala, G. (1988). *Introduction to econometrics.* New York: Macmilian Publishing Company.
10. Marinković, S., Džunić, M., & Marjanović, I. (2024). Determinants of housing prices: Serbian Cities' perspective. *Journal of Housing and the Built Environment*, 1601–1626. doi:https://doi.org/10.1007/s10901-024-10134-5
11. McQuinn, K., & O'Reilly, G. (2008). Assessing the role of income and interest rates in determining house prices. *Economic Modelling, 25*(3), 377-390. doi:https://doi.org/10.1016/j.econmod.2007.06.010
12. Meen, G. (2000). Housing cycles and efficiency. *Scotish Journal of Political Economy, 47*(2), 114-140. doi:https://doi.org/10.1111/1467-9485.00156

13. Nenovsky, N., Chobanov, P., Mihaylova, G., & Koleva, D. (2008). Efficiency of the Bulgarian Banking System: Traditional Approach and Data Envelopment Analysis. (I. -I. Research, Ed.) *ICER Working Papers 22-2008*. Retrieved from https://ideas.repec.org/p/icr/wpicer/22-2008.html

14. Roleders, V., Oriekhova, T., Zaharieva, G., Sysoieva, I., Dobizha, V., Pidhaiets, S., & Kucher, L. (2024). Economic justification of recycling in the processing industry. *Cleaner and Responsible Consumption, 13*, 100195. doi:https://doi.org/10.1016/j.clrc.2024.100195

15. Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics, 66*(1-2), 225-250.

16. Visković, J., & Čipčić, D. (2025). Effects of eurozone and schengen area accession on real estate prices. *Financial Internet Quarterly, 21*(4), 84-93. doi:https://doi.org/10.2478/fiqf-2025-0029

17. Zahariev, A., & Georgiev, T. (2023). The day-ahead energy market: correlation analysis and trends for selected CEE countries. *100th International Scientific Conference on Economic and Social Development – "Economics, Management, Entrepreneurship and Innovations". 100*, pp. 178-192. Svishtov: Varazdin Development and Entrepreneurship Agency, Varazdin, Croatia. Retrieved from https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdSvishtov2023_Online.pdf

18. Zahariev, A., Ivanova, P., Zaharieva, G., Slaveva, K., Mihaylova, M., & Todorova, T. (2023). Interplay between CSR and the Digitalisation of Bulgarian Financial Enterprises: HRM Approach and Pandemic Evidence. *Journal of Risk and Financial Management, 16*(9), 385. doi:https://doi.org/10.3390/jrfm16090385

19. Zaharieva, G., Tarakchiyan, O., & Zahariev, A. (2022). Market capitalization factors of the Bulgarian pharmaceutical sector in pandemic environment. *Business Management, XXXII*(4), 35-51. Retrieved from https://bm.uni-svishtov.bg/title.asp?title=2784

20. Zhelyazkova, V. (2018). The Road to Circular Economy: What Can Europe Learn from the Experience of Germany and Japan? *Economic Studies journal*(6), 167-177. Retrieved from https://ideas.repec.org/a/bas/econst/y2018i6p167-177.html

Стаття надійшла до редакції 05.01.2026
Стаття прийнята до друку після рецензування 20.01.2026