doi: 10.36689/uhk/hed/2026-01-019 HED Hradec Economic Days

Comparing Housing Affordability in the European Union

Michala STEFANOVIĆ* and Jan KAŇKA

Czech University of Life Sciences, Prague, Czech Republic; stefanovicm@pef.czu.cz; kankaj@pef.czu.cz
* Corresponding author: stefanovicm@pef.czu.cz

Abstract: This paper presents an alternative approach to the comparative analysis of housing affordability in European Union countries, utilizing linear regression on logarithmically transformed data. Methodologically, it builds on the model proposed by Kratochvíl and Havlíček (2024), who critically questioned the validity of per capita indicators in international comparisons. Traditional ratio-based indicators often distort reality by systematically overestimating values for small countries and underestimating them for countries with large populations, particularly when no direct proportional relationship exists between the observed variables and population size. The proposed regression approach enables the identification of residuals as indicators of deviations from the expected price trend based on population size. By comparing these residuals with conventional indicators, we provide evidence that regression-based models can serve as a more robust and equitable tool for assessing housing affordability, thereby supporting more informed public policymaking in the housing sector.

Keywords: housing affordability; log–log regression; residual analysis; per capita bias

JEL Classification: C21; R31; O18

1. Introduction

Housing affordability has become a key economic and social issue across the Member States of the European Union in recent years (Kikerec, 2024; Matějková & Tichá, 2025). It is typically assessed using ratio-based indicators, such as the share of housing expenditures relative to income or the housing cost overburden rate, often defined as spending more than 40% of disposable income on housing (European Construction Sector Observatory, 2019; Eurostat, 2022). While these approaches provide a certain degree of practical guidance (Bogdon & Can, 2003; Stone, 2006), they remain limited, as they fail to adequately account for structural differences between countries, such as population size, demographic density, or the level of urbanization (Licchetta, Dias Martins & Vašíček, 2025; Kikerec, 2024).

Kratochvíl and Havlíček (2024) demonstrated that ratio-based (per capita) indicators systematically bias results: small countries tend to be overestimated, while large countries are underestimated, whenever there is no direct proportional relationship between the observed variable and population size. This critique has proven relevant not only for GDP, mortality rates, or CO_2 emissions, but the same principle also applies to housing prices and income ratios.

In assessments of housing affordability, standardized indicators such as the housing cost overburden rate—defined as the share of households spending more than 40% of their disposable income on housing—are frequently used. This indicator is officially employed by institutions such as Eurostat and the OECD (European Construction Sector Observatory, 2019), but its simplicity entails substantial limitations. For instance, it does not account for the uneven distribution of income across population groups and may therefore conceal the actual difficulties faced by low-income households within average values (Kikerec, 2024). Another drawback is that the indicator operates as a simple ratio (costs/income), making it prone to distortions like those affecting other per capita measures. Countries differ in their tax structures, levels of social transfers, market prices of real estate, and cultural models of homeownership, all of which affect data comparability. For this reason, some authors recommend more sophisticated tools, such as multidimensional models or quantile regressions, which better capture the differentiated impacts of housing prices across socioeconomic strata (Fairbrother, 2014; Chung et al., 2019). These methods enable the analysis of affordability not only on average, but also at the level of poorer households, for whom the problem of housing affordability is often most acute.

Therefore, we propose applying linear regression on a log–log scale (Kratochvíl & Havlíček, 2024) to purge price indicators of the effect of population size. The model residuals (i.e., deviations from the expected price trend) then serve as a more robust measure of housing affordability, free from the inherent limitations of ratio-based indices. This regression-based approach enables the identification of countries that face genuinely higher or lower housing cost burdens relative to their economic and demographic profiles.

2. Methodology

The analysis is based on a combination of harmonized macroeconomic and demographic data obtained primarily from Eurostat databases (2024a, 2025) and the Organisation for Economic Co-operation and Development (OECD, 2024), with data from national statistical offices of individual EU Member States used where necessary. For each country, the following variables were collected:

- Total population (as of 1 January of the respective year).
- Median market price of a dwelling (in euros), with imputed values used in cases of missing data based on prices per square meter and a typical dwelling size.
- Median net annual income of a household or an individual, depending on data availability and methodological comparability.

The datasets were further adjusted for logarithmic transformation and cross-country harmonization. Following the methodology proposed by Kratochvíl and Havlíček (2024), a log–linear regression model is employed to analyze the relationship between population size and median dwelling prices across EU Member States.

In the literature and official statistics, housing affordability is most often assessed using ratio-based indicators, including in particular:

- PIR – Price-to-Income Ratio
- The ratio between the average property price and the annual median household income. Used, for example, by the World Bank and the OECD. Lower values indicate higher housing affordability (OECD, 2024a; IMF, 2025).
- HCM – Housing Cost Overburden Rate
- The share of the population spending more than 40% of their income on housing-related costs (including energy). A standard Eurostat indicator. A higher share indicates lower housing affordability (Eurostat, 2024a; OECD, 2024b).
- MIR – Mortgage-to-Income Ratio
- The ratio between monthly mortgage payments and monthly household income. Accounts not only for housing prices but also for interest rates and loan maturity (OECD, 2024a).
- Rent-to-Income Ratio
- The ratio between rent and income (monthly or annual). Used primarily in analyses of rental housing markets (OECD, 2024a).

Although these indices are intuitive and easy to compute, they suffer from inherent limitations: as ratio-based measures, they ignore the effects of population size, demographic density, and market structures. Therefore, we propose a regression-based approach inspired by Kratochvíl and Havlíček (2024), which addresses these distortions through a log–log model and residual analysis.

3. Results

The results of the analysis show that the traditional Price-to-Income Ratio (PIR) exhibits high variability and pronounced sensitivity to extreme values, particularly in countries with very low-income levels (e.g., Bulgaria and Romania) and in countries with exceptionally high housing prices (e.g., Luxembourg and the Netherlands). These extremes tend to affect relative cross-country comparisons significantly. In contrast, the residual values derived from the log–linear regression model display considerably greater stability, as the influence of extreme observations is substantially reduced.

Countries in Central and Eastern Europe—specifically Czechia, Slovakia, Hungary, Poland, and Croatia—exhibit PIR values well above the general trend line. This pattern points to the presence of structural factors, most notably rapid housing price growth in recent years, a persistently limited supply of new housing, tightening financing conditions, and a strong preference for owner-occupied housing. The regression residuals, however, suggest that the broader relationship between housing prices and income levels can explain part of these differences.

By contrast, countries such as Denmark, Bulgaria, and Italy record low residual values, indicating that housing prices in these countries are broadly in line with, or slightly below, the levels expected given household income. This finding suggests that high or low PIR values do not necessarily reflect the true extent of housing cost burden. Still, it may instead result from methodological limitations inherent in ratio-based indicators.

Overall, the results confirm that the regression-based approach, which relies on the logarithmic transformation of variables, mitigates distortions arising from extreme values and small-sample effects, thereby providing a more stable and methodologically consistent framework for international comparisons of housing affordability.

The quantitative results indicate substantial differences in dispersion between the PIR index and the regression-based residuals. While PIR values range from 3.093 to 13.305, regression residuals vary only between −0.262 and 0.300, demonstrating markedly lower variability across countries. This narrower interval suggests that the regression approach effectively controls for income differences and reduces cross-country distortions. Overall, the empirical comparison confirms the greater statistical stability of the residual-based indicator.

3.1. Detailed Results of the Analysis

The results of the log–linear regression analysis of the relationship between prices of newly built dwellings and household incomes in European Union countries allow for a systematic comparison with the traditional Price-to-Income Ratio (PIR). The analysis assesses differences in the explanatory power of both approaches in evaluating housing affordability.

In the selected model, the dependent variable was the average price of a newly built 70 m² apartment (in EUR), while the independent variable was the average annual gross household income (in EUR). By applying a log–log linear regression, we examined whether housing prices correspond to income levels across individual countries, that is, whether a proportional relationship exists between these two variables.

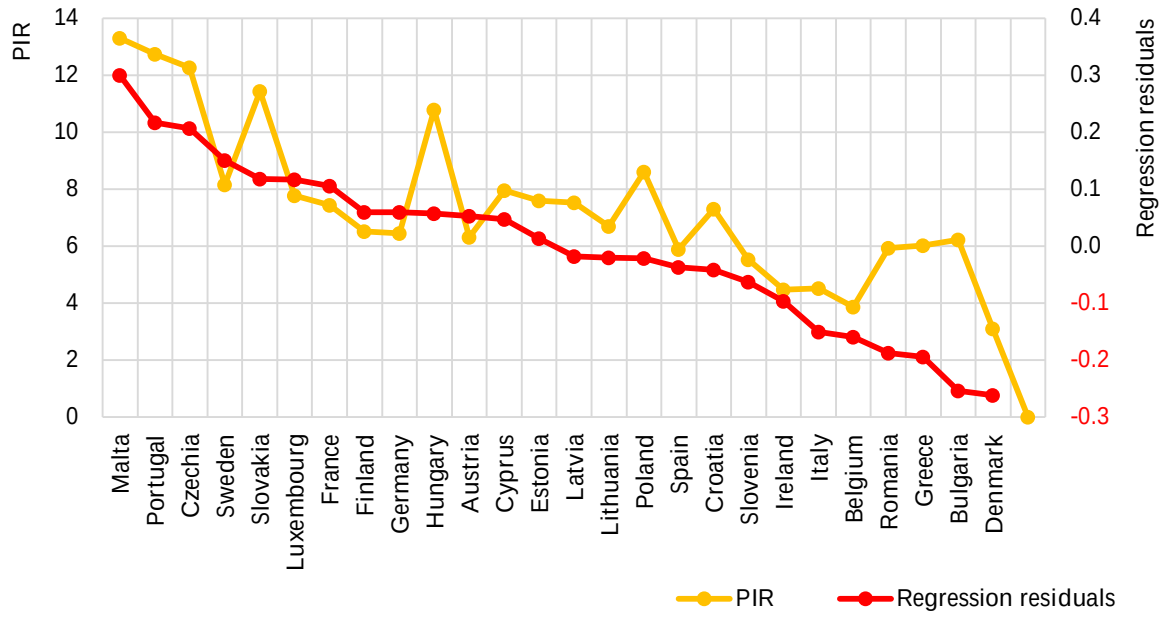

Figure 1. Graphical comparison of the two examined metrics

The figure contrasts two alternative metrics used to evaluate housing affordability across EU Member States.

- The yellow line (Price-to-Income Ratio, PIR) represents a conventional indicator that measures the number of average annual incomes required to purchase a dwelling of a given size.
- The red line (representing residuals from the log–linear regression model) indicates the deviation of observed dwelling prices from values implied by national income levels, showing whether prices exceed or fall below expectations.

The analysis demonstrates that the PIR is characterized by substantial dispersion and marked sensitivity to extreme observations. This is particularly evident in countries with very low-income levels, such as Bulgaria and Romania, as well as in countries with exceptionally high housing prices, including Luxembourg and the Netherlands, where relative comparisons are disproportionately affected. In contrast, residuals obtained from the log–linear regression framework exhibit considerably lower volatility, as the impact of extreme values is attenuated mainly, resulting in a smoother and more stable pattern.

Several Central and Eastern European countries—namely Czechia, Slovakia, Hungary, Poland, and Croatia—record PIR values that significantly exceed the prevailing trend. This outcome points to a combination of structural factors, including rapid increases in housing prices, persistently constrained housing supply, tightening credit conditions, and a strong cultural preference for owner-occupied housing. At the same time, the regression residuals indicate that part of these deviations can be attributed to the underlying income–price relationship rather than to affordability distortions alone.

Conversely, Denmark, Bulgaria, and Italy are associated with relatively low residual values, suggesting that housing prices in these countries broadly correspond to, or remain slightly below, levels predicted by household income. This finding highlight that elevated or depressed PIR values do not necessarily reflect the actual housing cost burden. It may instead be influenced by methodological shortcomings inherent in ratio-based measures.

Taken together, the results indicate that the regression-based approach, employing logarithmic transformations, reduces biases associated with extreme values and limited sample sizes. Consequently, it offers a more robust and internally consistent framework for cross-national comparisons of housing affordability.

As a result, these countries exhibit a significant divergence between PIR values and regression-based predictions since the simple price-to-income ratio fails to capture all relevant structural factors. In contrast, the regression model partially stabilizes these effects through its underlying trend function.

The traditional PIR index is limited by its sensitivity to absolute income and price levels and may overstate affordability problems in wealthier countries while understating them in poorer ones. Residual analysis based on the log–log model mitigates this shortcoming, enabling a more equitable assessment of housing affordability across countries.

The map illustrates the residual values from the log–linear regression between the average price of a newly built dwelling and average annual income. The blue color scale represents deviations from the trend-based prediction:

- Light blue: prices below the predicted level.
- Medium blue: prices in line with the trend.
- Dark blue: prices above the expected level.

This visualization facilitates a rapid spatial overview of differences in housing affordability across the European Union.

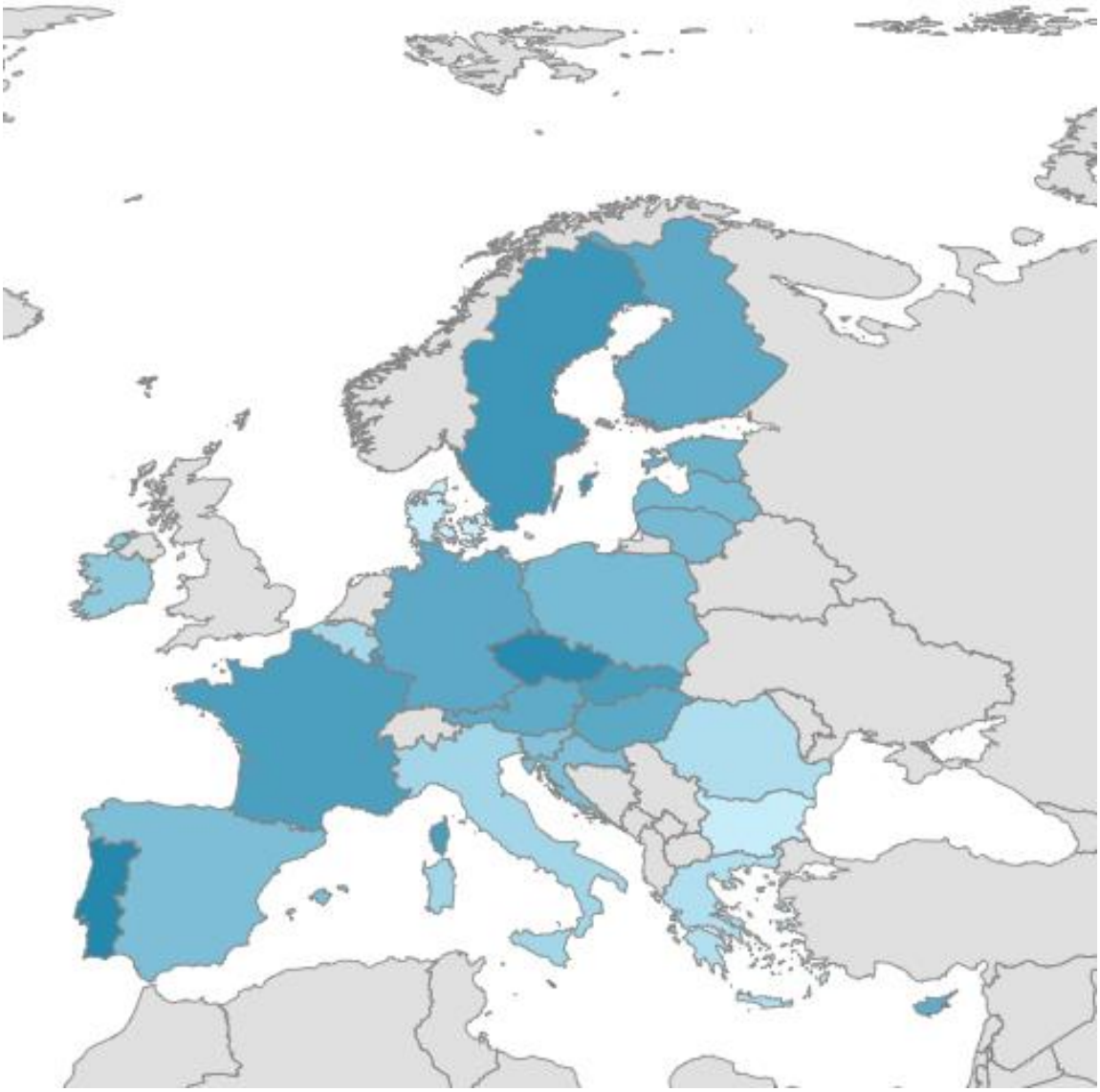

Figure 2. Geographical distribution of housing price residuals relative to income levels

Despite its statistical foundation, the proposed regression model is characterized by a surprising degree of simplicity and intuitive clarity. Its core consists of a linear relationship between the logarithmically transformed average housing price and the logarithm of average annual income, which enables the capture of the general trend in the relationship between the two variables. The advantages of this approach lie not only in its low computational complexity but, above all, in the greater stability of the results and their robustness to extreme values when compared with traditional ratio-based indicators (e.g., price-to-income ratios or per capita measures), which are often strongly affected by small bases in low-population countries.

Another significant strength of the regression framework is its methodological flexibility. The model can be readily extended to include additional variables that have been empirically identified as relevant determinants of housing affordability. These factors include:

- Level of urbanization: higher population concentration in urban agglomerations is typically associated with stronger demand and rising prices (see, e.g., OECD, 2021).
- Population density: as an indicator of pressure on space and infrastructure that influences housing markets.
- Demographic structure: the proportion of young households entering the housing market, representing a key demand group.
- Interest rates: a factor affecting practical housing affordability even when nominal housing prices remain unchanged.
- Share of rental housing: different ownership and rental arrangements influence mobility, price dynamics, and the vulnerability of specific population groups.

By incorporating these variables, the predictive power of the model can be strengthened, while simultaneously creating a multidimensional tool suitable for both academic comparison and practical policymaking in the housing field.

3.2. Formatting of Mathematical Components

Following the methodology proposed by Kratochvíl and Havlíček (2024), a log–linear regression model is employed to analyze the relationship between population size and median dwelling prices across EU Member States. Given the assumed nonlinearity of the relationship and the presence of scaling effects between variables, the model is specified as:

$$log\,(Y_i) = a \cdot log(X_i) + b + \varepsilon_i, \quad (1)$$

where:

Y_i denotes the median dwelling price in each country.

X_i represents the population size.

a is the scaling coefficient describing the elasticity of prices with respect to population.

b is a constant capturing the baseline price level.

ε_i is the residual, i.e., the deviation of the observed value from the value predicted by the model.

The model is estimated using the ordinary least squares method after applying a logarithmic transformation to both variables, which removes nonlinearities and stabilizes variance.

Interpretation of residuals:

$\varepsilon_i = 0$ corresponds to the trend.

$\varepsilon_i > 0$ is higher than the expected trend.

$\varepsilon_i < 0$ is lower than the predicted trend.

The residual component thus enables the identification of whether housing price levels in each country are above or below average relative to population parameters, providing an alternative to traditional per capita indicators.

3.3. Interpretation of the Relationship between Absolute and Relative Indicators in the Functional Space

The analysis of functional relationships between absolute and relative indicators identifies the conditions under which ratio-based measures are analytically meaningful. A functional perspective clarifies how growth patterns between the variable and population size affect the stability of per capita indicators.

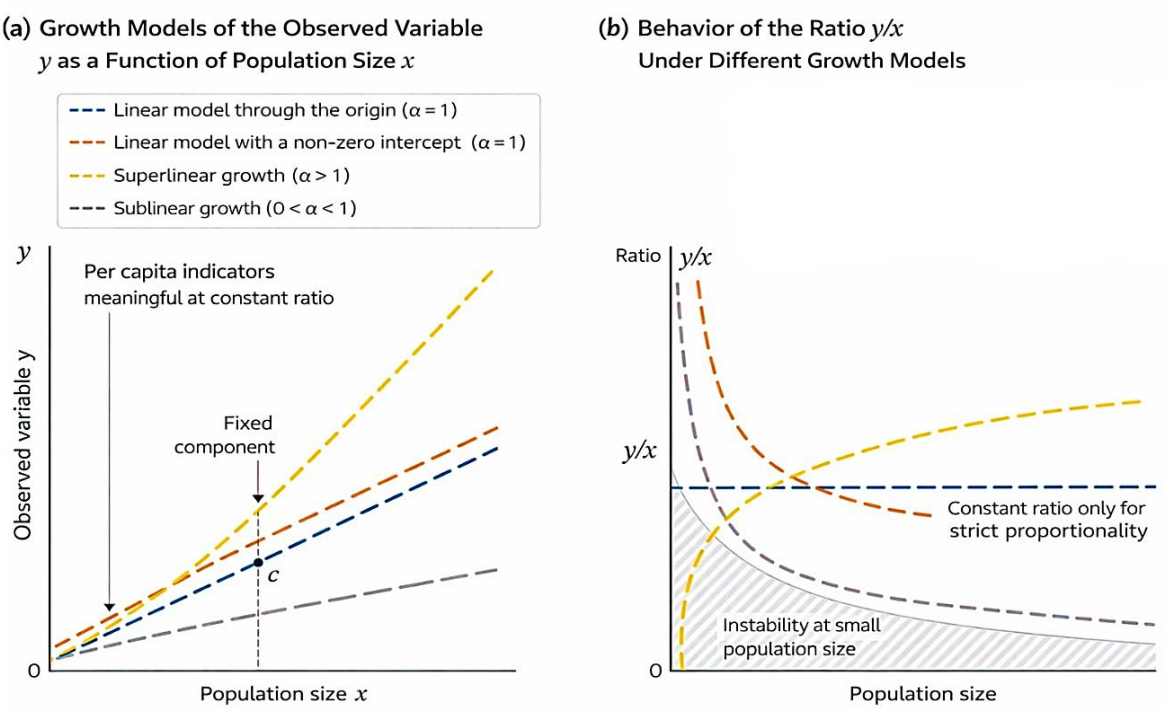

Figure 3. Interpretation of the Relationship between Absolute and Relative Indicators (Kratochvíl & Havlíček, 2024)

Figure (*a*) illustrates different growth models of the observed variable y (e.g., GDP or housing prices) as a function of population size *x*.

The blue curve represents a linear relationship passing through the origin—an ideal case in which the observed variable increases in direct proportion to population size. If the population doubles, the value of y also doubles. Only in this case is it meaningful to use ratio-based "per capita" indicators (i.e., *y/x*), as the ratio remains constant.

The orange curve depicts a linear relationship that does not originate at the origin, corresponding to a situation in which each unit (e.g., a country) has a certain fixed component independent of population size. In this case, ratio-based indicators primarily distort results for small countries.

The yellow curve represents exponential growth with an exponent greater than one, meaning that larger countries generate more than proportionally higher values of *y*, which is often the case for aggregated variables such as GDP. In this scenario, per capita indicators systematically overestimate smaller countries.

The grey curve corresponds to sublinear growth ($0<\alpha<1$), where increments in y decrease as population size increases.

Figure (*b*) shows how the ratio *y/x* changes with population size. It is evident that only for the blue curve does the ratio remain constant, which justifies the use of per capita indicators exclusively in cases of direct proportionality.

For the other types of relationships, however, the ratio exhibits substantial instability, particularly at low values of x: the ratio y/x may take extremely high or low values as a result of the limiting behavior of the functions as $x \to 0$. This phenomenon has important implications for international comparisons, as in countries with small populations even minor absolute changes can significantly affect relative indicators, leading to unreliability and interpretative bias.

It therefore follows that the use of "per capita" indicators is justified only when it can be statistically demonstrated that a direct proportional relationship exists between the variable *y* and population size *x*.

4. Conclusions

The application of the log–linear regression model proposed by Kratochvíl and Havlíček (2024) to the assessment of housing affordability offers an alternative to traditional per capita indicators, which often exhibit systematic biases, particularly in relation to differences in population size across countries. This approach enables the elimination of such distortions, thereby allowing for a more accurate capture of the actual economic affordability of housing in individual countries.

Log-linear regression offers a more robust alternative to ratio-based indicators for comparing housing affordability. The residuals reveal the actual deviations of housing prices from the expected price level implied by the average income of a given country's population. The method is computationally simple, extendable to additional factors (e.g., interest rate, unemployment, mobility), and suitable for the design of practical policies and official statistics.

The findings of this study provide an opportunity for a broader revision of the evaluation tools currently used in the field of housing affordability. Owing to its novel computational framework, this method can significantly contribute to revising existing assessment standards. It may serve as a starting point for the development of integrated, internationally comparable indices. It is therefore desirable that this approach be further tested, refined, and implemented within official statistics and EU strategies. The results of this analysis have direct applicability for formulating more effective, data-driven housing policies.

Conflict of interest: none

References

Bogdon, A.S., & Can, A. (2003). Indicators of Local Housing Affordability: Comparative and Spatial Approaches. *Real Estate Economics, 25*(1), 43–80. https://doi.org/10.1111/1540-6229.00707

Deloitte. (2024). *Deloitte Property Index 2024 - Overview of European Residential Markets*. https://www.deloitte.com/nl/en/Industries/real-estate/perspectives/property-index-deloitte.html

European Construction Sector Observatory. (2019). *Housing affordability and sustainability in the EU: Analytical Report*. EU Single Market. https://single-market-economy.ec.europa.eu/publications/housing-affordability-and-sustainability-eu_en

Eurostat. (2022). Housing cost overburden rate. In *Housing in Europe – 2023 edition* (Interactive Publication). Eurostat. https://ec.europa.eu/eurostat/web/interactive-publications/housing-2023

Eurostat. (2024a). *Housing cost overburden rate*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Housing_cost_overburden_rate

Eurostat. (2024b). *Average full time adjusted salary per employee*. https://ec.europa.eu/eurostat/databrowser/view/nama_10_fte__custom_13597179/bookmark/table?lang=en&bookmarkId=c29eed24-377e-4763-aaf0-0419906d2ecd

Fairbrother, M. (2014). Multilevel modelling for housing affordability. *Journal of Social Policy, 1*(2), 239–262. https://doi.org/10.1017/psrm.2013.20.

Chung, Y.-N., Chung, K.-K., Gordon, D., Mak, J.K.-L., Zhang, L.-F., Chan, D., Lai, F.T.T., Wong, H., Wong, S.Y.-S. (2019). Housing affordability effects on physical and mental health: household survey in a population with the world's greatest housing affordability stress. *Journal of Epidemiology & Community Health, 74*(2), 164–172. https://doi.org/10.1136/jech-2019-212286

International Monetary Fund. (2025). *Global Housing Watch: House price-to-income ratio indicators.* https://www.imf.org/external/research/housing/index.htm

Kikerec, L. (2024). Trends and drivers of housing affordability in the EU: Insights from panel data analysis. *Croatian Economic Review of Business and Social Sciences*, 2(4), 49–62. https://doi.org/10.62366/crebss.2024.2.004

Kratochvíl, L., & Havlíček, J. (2024). The fallacy of global comparisons based on per capita measures. *Royal Society Open Science*, *11*(3), 230832. https://doi.org/10.1098/rsos.230832

Licchetta, M., Dias Martins, V., & Vašíček, B. (2025). Selected macroeconomic and social aspects of housing affordability. *Intereconomics, 60*(2), 101–106.

Matějková, J., & Tichá, A. (2025). Housing Market Trends and Affordability in Central Europe: Insights from the Czech Republic, Slovakia, Austria, and Poland. *Buildings, 15*(10), 1729. https://doi.org/10.3390/buildings15101729

OECD. (2024a). *Analytical House Price Indicators: Price-to-income ratio. OECD Affordable Housing Database*. https://www.oecd.org/en/data/datasets/oecd-affordable-housing-database.html

OECD. (2024b). *Housing costs over income (HC1.2). OECD Affordable Housing Database.* https://www.oecd.org/els/family/hc1-2-housing-costs-over-income.pdf

Stone, M.E. (2006). What is housing affordability? The case for the residual income approach. *Housing Policy Debate, 17*(1), 151–184. https://doi.org/10.1080/10511482.2006.9521564