DEVELOPING AN EQUILIBRIUM MODEL FOR ASSESSING HOUSE PRICE DEVIATION

Dragomir Stefanov[1],
Yana Stoencheva[2],
Petar Ivanov[3]

Abstract: This article deals with the problem of developing a reliable diagnostic tool that could help in reducing the unpredictability of house price cycles that cause undesirable reverberations in the economy. Such type of cycles is among the main drivers of price volatility and crises not only in the real estate sector itself but also for the financial system and the whole economy too. The purpose of this study is to introduce a step-by-step methodology for assessing the true deviations of house prices from their long-term equilibrium levels. Factors with a significant impact on house price dynamics are also determined. The result is a ready-to-use econometric model that detects possible formation of house price bubbles early enough, so that policymakers and regulators could take adequate actions against market overheating in due time.

Key words: house prices, equilibrium, econometric model, methodology, price cycles

JEL: C32, E32, R31.

DOI: https://doi.org/10.58861/tae.bm.2025.3.04

Introduction

House prices are historically prone to be *cyclical* in nature as their dynamics goes through constantly repeating phases of *booms* and *bursts*.

[1] Assoc. Prof., PhD, Department of Real Estate Property, University of National and World Economy – Sofia, e-mail: d.stefanov@unwe.bg, ORCID: 0000-0002-0600-3873

[2] Assoc. Prof., PhD, Department of Real Estate Property, University of National and World Economy – Sofia, e-mail: ystoencheva@unwe.bg, ORCID: 0009-0009-9381-4217

[3] PhD, Department of Real Estate Property, University of National and World Economy – Sofia, e-mail: petlv@unwe.bg, ORCID: 0000-0003-1223-8131

However, while housing booms typically *build up silently* in time, market downturn could surprise by its *sudden and rapid evolution* (Corradin & Fontana, 2013). Moreover, cycles could differ in length and time. A major plummet of house prices could have a detrimental impact not only on the real estate and construction sector, but it could also be transmitted to the financial sector and the whole economy, too. Such a powerful and *contagious effect* is a typical trait of the house price bubbles.

A notorious example of the above is the US *housing bubble* that took place in the 2000s. It originated from the uncontrolled expansion of *high-risk mortgage bonds* in the banking sector. This initially domestic crisis of overvalued houses eventually led to the onset of the *Global Financial Crisis* (GFC) in 2007-2008 and to a widespread economic recession as an aftermath. Eigner & Umlauft (2015) and others consider GFC to be the *worst financial crisis* since the 1930s Great Depression. This is convincing evidence for the threat posed by a house price bubble.

The main *motives* behind the current research are the challenges before the *timely detection* of house price bubbles. Finding an instrument to deal with these challenges will assist *central banks* and other institutions to prevent, or at least weaken, the negative consequences of a bursting bubble by undertaking proper macroeconomic and credit policy measures. In that way the *goal* of the article is to develop an *econometric model* for detecting house price bubbles through measuring the *deviations* of house prices from their *equilibrium levels*. *Abnormal behaviour* of house prices could be an early signal for a bubble formation and hence monitoring is important to be regularly done with the proper toolkit.

There are several ways for measuring deviations of house prices. Our methodology for that is based on building a multivariate *equilibrium econometric model* that is a kind of either a Vector Autoregression (VAR) or Vector Error Correction Model (VECM) – two techniques widely used in related papers (see the literature review section below). The model itself presents a *contemporaneous equilibrium of dependable variables* with the change of the national house price index as one of them and the others are factors with long-term impact on it. In that way the equilibrium is built as a single *dynamic system* of *mutual long-term interaction*. In the case of a VECM, a part with *short-term equilibrium* is also added, in which external factors participate, too.

1. Literature review

A *bibliographic survey* by Li & Li (2022) gives a good idea of the scope of research on house prices and the drivers of their change. Their review of *academic publications* on the subject for the last six decades (from 1960 to 2020) showed a total number of 4,125 relevant publications in the Web of Science database. Articles studying the behaviour of *house prices* would most often recourse to econometric models, especially of the VAR/VECM type and their derivatives. The main motive here is that such dynamic models would help override the limitations intrinsic in any static methods, where the long-term average will be calculated as a historically average, i.e., it will remain fixed in time.

Kotseva & Yanchev (2017) conducted one of the most *comprehensive studies in Bulgaria* on the current topic. They developed a VECM-type model for measuring the deviation of house prices in Bulgaria from their *long-term trend* by studying the cointegration relation between house prices, the total amount of *hypothetical borrowings* and the direct foreign investments (FDIs) into the real estate sector. *Hypothetical borrowing* of money is a coefficient derived by the authors from the interaction between household disposable income and mortgage interest rates. Another qualitative research on the topic, specifically for Bulgaria, is that of Yovkova (2009). The author examines the correlation between the domestic real estate and mortgage markets.

Econometric research on the *house price drivers* is very thorough for the *developed countries* (especially the USA, UK and EU countries). In the last two decades an increasing number of studies have been conducted for the *emerging markets* too, including China, the UAE and the CEE countries (where Bulgaria belongs). They are attractive for the researchers with the price dynamics, characterized by prolonged phases of a booming market. At the same time, however, researchers of those markets encounter several difficulties related to the *quality* and *scope* of the available data. Also, the CEE region has its own specifics for the inflating house price bubble preceding the Global Financial Crisis of 2007-2008, among which are the *increased demand for dwellings from abroad* and the *credit boom* connected with financial globalization and liberalization of credit standards (see Egert, B. & Mihaljek, D. (2007)).

As we see, studies on the current topic are predominantly model based with a wide range of applied methods. For example, Cevik & Naik (2022) apply a relatively innovative approach to analyse house price change in 10 emerging housing markets in Europe. They use a method based on a *panel quantile regression*. Applying this method, they perform segmental price analysis, which allows them to compare the changes that occur in the *upper and lower price ranges* of the housing market. In a study covering the impact of monetary policy on house prices, Lee & Park (2022) use a VAR-type model with *a time-varying component*, or the so-called TVP-VAR (time-varying parameter). This model accounts for the variable influence of interest-rate shocks during the periods of implementation of different monetary policies, in contrast to the standard VAR model, where the influence of the factors is assumed constant.

The model of Beraldi & Zhao (2023) proves that the *elasticity of housing demand* depends on the change in the *interest rate* on *mortgage loans*. The authors empirically find out that a 1% reduction in mortgage rates will theoretically result in rise of house prices by about 10%. Aastveit & Anundsen (2022) and several other authors emphasize on the effects of *monetary policy* of central banks on the house price movement. For example, during the COVID-19 pandemic the *monetary policy easing* has its own contribution to house price appreciation that is intensified by the *shift of household preferences* from renting to owning a home.

The European Central Bank (ECB), for example, has developed and implemented a comprehensive *toolkit for monitoring and detecting* the signals of bubble-inflated house prices. For this purpose, the institution monitors a set of four indicators: on the one hand, deviations from the long-term average of the “price-to-income” and “price-to-rent” ratios, and on the other: econometric estimates obtained by applying inverse demand and asset valuation models[4]. In addition to these, the ECB uses a model called “*Residential real estate price-at-risk model*”, which collects information from the above four indicators and others related to them, including *financial conditions*, the *level of cyclical systemic risk*, the current rate of *house price growth* and *consumer attitudes* (Jarmulska et al.,2022).

As we could *summarize* our findings from the above and from other sources, the use of *econometric models* has become *ubiquitous* in almost any quantitative research, especially when they cover macro-financial indicators. In support of the thesis for *qualitative accumulation* in economic theories is the study of Guo et al. (2024). What is crucial to the persistence of a successful econometric model, however, is the choice of *appropriate variables*. Bojinova et al. (2010) presented a methodology for that in a study

[4] The inverse demand function considers price as a function of quantity, rather than the regular scenario in which the quantity demanded is a function of price.

of public companies in Bulgaria. In our article, we also introduce a *procedure* based on quantitative methods for *sorting out the country-specific factors* most closely connected with the change of house prices.

2. Presenting the model

We adopt an approach in which we *group* diverse factors into separate *cluster groups* based on a common feature. Therefore, the study of the impact on house prices will first begin at the cluster group level, and then will proceed *individually* for each of the factors included in it, i.e., we follow the principle from the general to the specific. As Bespalov et al. (2021) note in their study, *cluster analysis methods* are widely used in economics, especially when comparing objects that have common properties.

We present in detail the methodology for building the so-called by us *Equilibrium Econometric Model of House Prices* (EEMHP) in the next section. This methodology's composition encompasses 5 stages with a total of 17 steps within them. The main dependent variable (Y) is the *House Price Index* (HPI) at the national level in Bulgaria, calculated by the National Statistical Institute (NSI) on a regular basis. The HPI is a quarterly indicator and represents the *change in market prices of dwellings* (both newly built and existing) purchased by households.

The desired equilibrium econometric model will represent a *system of interrelated regression equations*. Each of these equations will yield an outcome variable, which is the change in Y or one of the other endogenous variables closely interacting with Y. The behaviour of the output variables will be programmed so that it depends on the combined impact of the other *endogenous* and *exogenous* variables included in the model, whereby they form *subsystems of long-term and short-term equilibrium*.

3. Stages of compiling the EEMHP

In **stage 1**, it is planned to carry out an initial review of the time series dynamics and their statistical properties. We work with quarterly data. Step one of this stage is to look up the indicators which interact with house prices we intend to measure in the relevant statistical sources. We *derive* their time series with all the observations for the reference period. Step two involves performing *transformations* on the time series where needed. There are three possible transformations, namely: 1) *negative-to-positive transformation*, 2) *logarithmic transformation* (all indicators are expressed in their natural logarithm form except for indicators representing interest rates, growth rates and shares) and 3) *differentiation* (we apply to time series that are not *stationary*).

The *stationarity test* is the *third step* of stage 1. It is performed visually and statistically. In the latter, we perform unit root tests, which include the *Augmented Dickey-Fuller (ADF), Phillips and Perron (PP) test*, and the *Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test*. The final step of this stage is the fourth, in which we *examine and compare* the behaviour of each of the indicators, as well as its group as a whole, against that of the HPI.

In **stage 2**, we examine the regression dependencies within each of the groups and identify the factors with the strongest influence on house prices. Step one at this stage is to construct regression equations for each group of indicators, in which they would play the role of independent variables (x) with the HPI being the dependent variable (y). We also include a constant term (C) in them. The regression equations will be *double logarithmic* and will take the following form:

$$log\, y_t = \beta_0 + \beta_1\, log\, x_{1,t} + \beta_2\, log\, x_{2,t} + \beta_3\, log\, x_{3,t} + \varepsilon_t \quad (1)$$

The double log form is suitable for the intended purpose as it allows measuring the elasticity of the dependent variable (*y*) with respect to the behaviour of the independent variables (x_i). Thus, the coefficients before *x* will indicate the percentage change in *y* per unit change in *x*. In addition, the logarithmic transformation of all the variables allows us to work with a linear model even if there is no linear dependence between them.

The applicable method for calculating the above type of equations (1) is that of *linear ordinary least squares* (OLS). We examine the readings of each factor and check the direction of their coefficients to see whether it corresponds to preconceived theoretical notions. When we find *discrepancies* between empirical results and theoretical assumptions, we should *reconsider the relationship* between house prices and the specific factor that has caused the disturbance. Finally, we highlight the factors that are *statistically significant* for Y and have high coefficients. These will be among the *potential candidates* to play the role of endogenous or exogenous variables in our econometric model.

We extend the analysis by presenting each of the regression equations in *two variants with different scope*: the first variant would be narrower and would include the period up to the end of 2019 (i.e., before the onset of the COVID-19 pandemic), while the second would cover the *entire period* with available data, e.g., until the end of 2023. The reason behind this division is to *measure the impact of the pandemic* on the existing regression relationships and also to find out whether the economic shock from it causes their disruption. Finally, in this step, we draw conclusions about the joint *influence* of the factors of the particular group on the dependent variable Y by comparing the readings of the *coefficient of determination (R^2) and the Durbin-Watson (DW) autocorrelation statistics* for the whole equation.

Step two would involve conducting pairwise *tests of Granger causality* between the HPI (Y) and each of the factors of the specific group. The *F-statistic and probability (ρ)* of each pair in any of the two directions of causality are compared. We note all cases, in which we reject the null hypothesis of no causality. We single out the factors from the respective pair that exhibit causality in *both directions*—i.e., from the corresponding indicator to Y, as well as from Y to it. Factors meeting this condition together with Y would be candidates for inclusion in the role of *endogenous variables* in our projected model. We check if all other factors that exhibit Granger causality only in *one direction* or *none at all* are statistically relevant in the regression equation from the previous step. In case of a positive answer, we direct the corresponding factor to the group of potential candidates for *exogenous variables*.

The last step three of the second stage involves compiling a summary table based on the conclusions drawn. There, we indicate the *role of each of the considered indicators* – whether it will participate in the model as an endogenous variable, as an exogenous variable, or will not participate at all.

In **stage 3**, we *model the long-term dependence* of the HPI on other variables that we have chosen to be endogenous. By long-term equilibrium we mean a system, in which the variables would *move in synchronization*, i.e., exhibit a form of co-integration. In this step, we will define the model specification type. We will limit the choices to *two classic types* of vector models – *Vector Auto Regression* (VAR) and *Vector Error Correction Model* (VECM). The algorithm for choosing the suitable type of specification is presented in the form of a diagram (see *Fig. 1*).

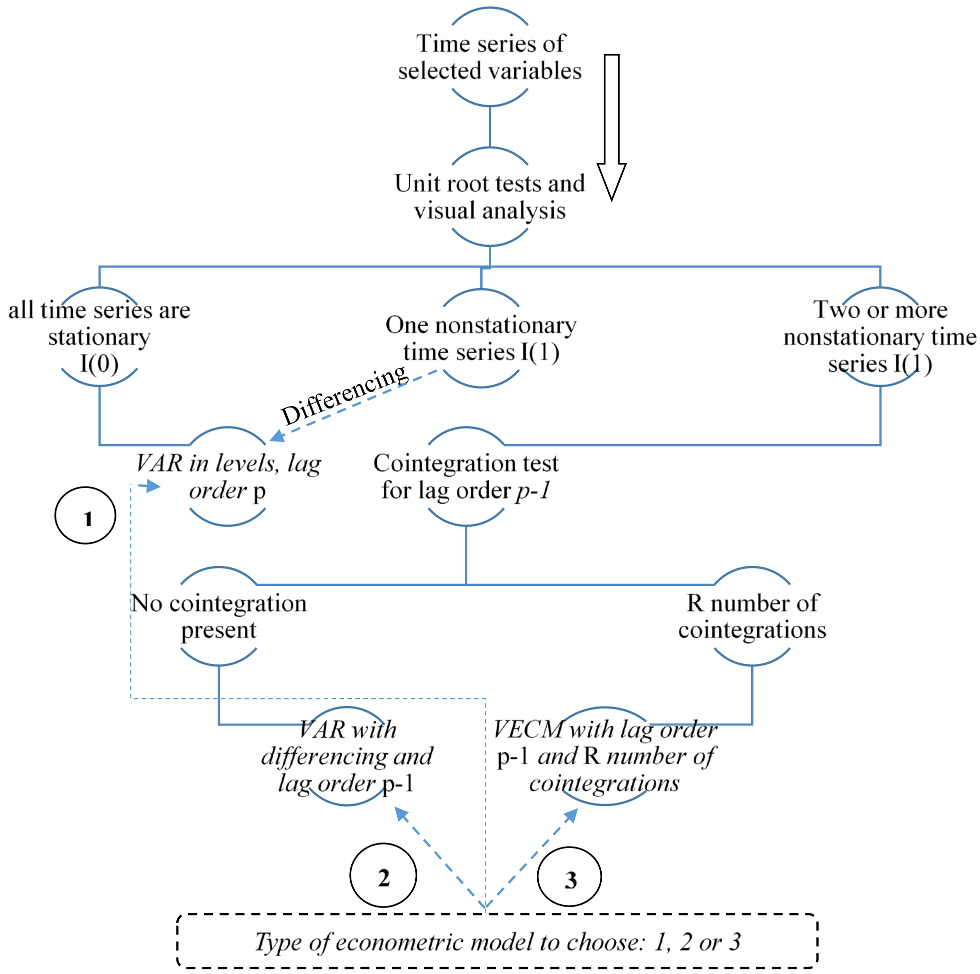


Figure 1. Procedure for choosing an econometric model depending on the statistical properties of the data
Source: own methodology, see Ivanov (2024)

Finding a *stable long-term (co-integration) relationship* between the endogenous variables involved suggests choosing a VECM type model. It would allow to determine the pace with which deviations from the long-run equilibrium are cleared out after a shock. BNB applies a similar type of model to determine the *equilibrium* (long-term) level of house prices in Bulgaria and to *assess the impact* of identified macroeconomic factors on their dynamics (Kotseva & Yanchev, 2017).

Initially, model construction starts with a VAR-type specification, which is expected to subsequently *undergo modifications* depending on the results of the tests conducted. Essentially, a VAR is a system of equations that can be written mathematically in the manner shown in (2). As to the example below, we will assume that we are working with three variables (x, y, z), with each of them participating with 2 lags, i.e., the system would match type VAR (2):

$$x_t = c_1 + \sum_{i=1}^{2} a_{1,i} y_{t-i} + \sum_{i=1}^{2} \beta_{1,i} x_{t-i} + \sum_{i=1}^{2} \gamma_{1,i} z_{t-i} + \epsilon_{x,t}$$
$$y_t = c_2 + \sum_{i=1}^{2} a_{2,i} y_{t-i} + \sum_{i=1}^{2} \beta_{2,i} x_{t-i} + \sum_{i=1}^{2} \gamma_{2,i} z_{t-i} + \epsilon_{y,t}$$
$$z_t = c_3 + \sum_{i=1}^{2} a_{3,i} y_{t-i} + \sum_{i=1}^{2} \beta_{3,i} x_{t-i} + \sum_{i=1}^{2} \gamma_{3,i} z_{t-i} + \epsilon_{z,t} \quad (2)$$

where $c_{1,2,3}$ is a constant (intercept); α, β and γ are coefficients before the variables x, y and z; ϵ is a random error, with t being the current period.

Depending on the results of the *stationarity tests* performed for each time series, the VAR specification may need additional operation for differentiation. The latter would imply differentiating the variables on both sides of the equations, i.e., variables to be represented by their differences (e.g. Δx, Δy, Δz) from each two adjacent periods. This is the reason why the number of lags in this variant shall be reduced by one. If we come across a hypothesis of *non-stationary time series* and *existing co-integration* between these time series, then we should convert the model specification from VAR type to VECM-type with *p-1* number of lags and *r* number of co-integrations. In this case, we have an *error correction component*.

Having considered the theoretical settings, we would move on to list the individual steps needed for the implementation of **stage 3**. The first step thereof will be to determine the *optimal lag* of the model (first task) and possibly to detect *co-integration relationships* between the endogenous variables (second task). *Finding the optimal lag* will require comparing the readings of different types of information criteria. In our case, we choose to work with the following three types: Akaike's criterion, Schwartz's criterion, and Hannan-Quinn's criterion.

The technical choice concerning the ***optimal number of lags*** will be based on the rank, at which the system would have the maximum number of *lowest values* of such information criteria. As long as a smaller number of lags would typically *simplify* the model, we will go for a variant with *fewer lags*, should the loss of information, compared to a model with more lags, is not significant.

Then we would move on to the implementation of the *second task* of the current step: searching for co-integration relations, however, only if the condition is met that at least two of the endogenous variables involved in the model exhibit *non-stationarity of the first integral order* I (1). In this case, we proceed to study the existence of a long-term *co-integration relationship* between them. For this purpose, we conduct *Johansen's co-integration test*, which would either prove or reject the existence of a general trend (i.e., long-term dependence). In the case of a positive result, we find out what the number of *co-integration relationships* is.

At the end of step one of stage 3, what we do is analyse the results of all tests performed and following the diagram in *Fig. 1,* we select a type of model *specification* against one of the three variants proposed. Then we proceed by carrying out a *transformation* of the original VAR specification, if necessary, and in the new version of the model we take into account the proposed number of lags in the previous procedure and possibly the presence of co-integration relationships.

In step two of the third stage, we conduct a series of tests to determine the optimal composition of the model. For this purpose, while we try a *series of combinations* between the endogenous variables involved so far, we monitor their significance and coefficient readings. Moreover, we also compare the coefficients of determination (R^2), DW statistics and information criteria of the different model variants. During this step, some of the selected endogenous variables may *be dropped* if they do not fit optimally into the modified model structure. However, if dropped, they will also be tested whether they would be suitable for *exogenous or not*. In case endogenous drop out, we should again carry out the tests for the detection of an optimal lag and to see whether there is co-integration or not. Where we detect changes, we update the current model specification.

Step three is the last step of the current stage, and here we should make sure that the model constructed so far is both *sustainable and reliable*. For this purpose, we conduct *three types of tests* on the model's *residuals*, namely for the presence of *autocorrelation, heteroscedasticity and normal distribution*. When we are satisfied with their results, we can continue with the model, saving the new settings on it, if we have made any.

Then in **stage 4**, we finalise the model by constructing the *short-run equilibrium* subsystem related to the endogenous variables included. This step is feasible in cases where we are working with a VECM type model (this is expected to be the default case). The *short-run equilibrium* is constructed by adding *exogenous (external) variables* to the system of equations that have a significant influence on the endogenous ones, but without experiencing a counter-effect on themselves. By adding these, the model takes into account the influence of a *wider range* of relevant factors, adjusting better to short-term fluctuations. The result is that the explanatory power of the model increases. The *group of potential exogenous factors* should already be compiled at the end of step three of stage 2, in which, having made a review and analysis, we have determined the *role* of each factor that was proposed at the beginning of the study. To this group, we add the endogenous ones, which during the numerous modifications of the model were *dropped* from it in the previous stages.

In step one of stage 4, we check *whether there is a structural break* in the relationship between the endogenous variables and Y. What we understand under this term is a *sudden change* in the behaviour of a time series from a certain point onwards, which is expressed in a change in the average or some other parameter in the data generating process. Should such a phenomenon not be taken into account, the constructed econometric model would run the risk of not being sustainable and would be likely to produce large deviations in the forecasts. Based on the results of this step, we will determine whether we need to introduce an *additional structural break* variable into the model.

We check for structural breaks by visually inspecting the dynamics of the *logarithmic ratios* between each endogenous variable and Y. To confirm the visual observations, we proceed with conducting a specific *test to check whether there are any multiple structural breaks or not*. Essentially, the test consists of multiple *consecutive* single tests performed according to Bai and Perron's methodology. The detection of a structural break requires the inclusion in the model of *dummy variables* that take only *two possible values*, which are either *0 и 1*, with the border between them being the date of the break. We *add* such *dummy variables* if we have identified the need for them to the *list* of potential exogenous factors to be tested for significance.

In step two of stage 4, we search for the *optimal combination of endogenous and exogenous variables*. Having completed this step, we would arrive at the *final version* of the model, provided that subsequent control tests confirm its strength, stability and reliability.

In case we are using a VECM-type specification, we should perform some *additional settings*. The *first* of them is related to the presence of a *co-integration equation* in the model, the calculation of whose parameters would require that *the first endogenous variable coefficient* (i.e., the Housing Price Index) *be fixed at 1, i.e.,* $b_1 = 1$ in the above example. This is how we will be able to *measure the influence* that each of the remaining endogenous variables exerts on Y in order to maintain the system in long-term equilibrium.

The *second setting* in the VECM specification of the model would be related to the imposition, if necessary, of a weak exogeneity condition in the error correction term of one or more of the model's endogenous ones. While the mentioned *second setting* made to a VECM-type model would not always apply, it will depend on whether the correction coefficients α_m, α_p and α_y are significant, cf. (3).

$$\Delta m_t = \alpha_m (1 * m_{t-1} + b_2 p_{t-1} + b_3 y_{t-1} + b_4) + \lambda_{MM} \Delta m_{t-1} + \lambda_{MP} \Delta p_{t-1} + \lambda_{MY} \Delta y_{t-1} + v_t$$
$$\Delta p_t = \alpha_P (1 * m_{t-1} + b_2 p_{t-1} + b_3 y_{t-1} + b_4) + \lambda_{PM} \Delta m_{t-1} + \lambda_{PP} \Delta p_{t-1} + \lambda_{PY} \Delta y_{t-1} + u_t$$
$$\Delta y_t = \alpha_Y (1 * m_{t-1} + b_2 p_{t-1} + b_3 y_{t-1} + b_4) + \lambda_{YM} \Delta m_{t-1} + \lambda_{YP} \Delta p_{t-1} + \lambda_{YY} \Delta y_{t-1} + \eta_t \quad (3)$$

In this case, by *m, p* and *y* we denote endogenous variables. α_m in particular (the coefficient in front of house prices) is initially important to be statistically significant (*so that its t-statistic is below -1.96*), otherwise the VECM-type model will not be applicable.

This would mean that if we prove that one of the coefficients in the error term is statistically insignificant, then its variable exhibits a property of *weak exogeneity*, i.e., we can no longer treat it as a fully-fledged endogenous variable. It will continue to influence the long-run equilibrium, however it *will not participate with an error correcting coefficient in the short-term settings* of the model, as it responds more slowly to the influence than the other endogenous variables. Therefore, if any of the first differences of the endogenous variables show signs of weak exogeneity, we introduce a *constraint* on its error correction coefficient and fix it to *zero*.

The *short-term equilibrium subsystem* of the model resembles a standard VAR system of equations. When constructing this short-term equilibrium, we *do not need to impose any constraints* as opposed to setting the long-term equilibrium and the error correction term. In particular, *there would be no need to make a weak exogeneity check* as long as what we examine is the exogenous influence of all variables. In the VECM structure, we add the exogenous variables to the second part of the equations of (3) as follows (see underlined):

$$\Delta m_t = [\ldots] + \lambda_{MM} \Delta m_{t-1} + \lambda_{MP} \Delta p_{t-1} + \lambda_{MY} \Delta y_{t-1} + \underline{\lambda_{MEx1} \Delta Ex_{1,t} + \lambda_{MEx2} \Delta Ex_{2,t} + [\ldots]} + v_t$$
$$\Delta p_t = [\ldots] + \lambda_{PM} \Delta m_{t-1} + \lambda_{PP} \Delta p_{t-1} + \lambda_{PY} \Delta y_{t-1} + \underline{\lambda_{PEx1} \Delta Ex_{1,t} + \lambda_{PEx2} \Delta Ex_{2,t} + [\ldots]} + u_t$$
$$\Delta y_t = [\ldots] + \lambda_{YM} \Delta m_{t-1} + \lambda_{YP} \Delta p_{t-1} + \lambda_{YY} \Delta y_{t-1} + \underline{\lambda_{YEx1} \Delta Ex_{1,t} + \lambda_{YEx2} \Delta Ex_{2,t} + [\ldots]} + \eta_t \quad (4)$$

where the content of the second square brackets follows the same pattern with the exogenous variables denoted before it, that is to add the product between an exogenous variable (Ex) and its corresponding coefficient (λ) until all involved exogenous variables are listed.

As we described at the beginning of the step, we are looking for the composition with the best combination of variables both in terms of their *individual contribution* to the variation of Y and in terms of the overall representation of the system of equations. Specifically, we examine the *t-statistic* of each of the outliers with respect to Y in search of values equal to, below (≤) *the cutoff of* -1.96 or above (≥) *the cutoff of 1.96* (equivalent to $p < 0.05$). We remove from the model factors that fail achieving this criterion (which t-statistic is between -1.96 and + 1.96) because they are insignificant for the system.

Moreover, we also monitor the general performance of the system of equations, taking into account the readings of its main indicators: the coefficient of determination, DW statistics and information criteria among others. We also check the coefficients in front of the corresponding variable, whether they correspond to the initially set theoretical assumptions about the *direction of the relationship* (positive or negative correlation) with respect to HPI. In case of drastic discrepancies, we switch to another modification of the model. After all these operations, we single out an optimal variant of a model.

Starting from here through the end of stage 4, we run a series of control tests to validate the model and make final adjustments to it, as appropriate. The first of these is a check for the presence of *multicollinearity* between the exogenous variables. We conduct it as part of the third step of stage 4. We perform the testing by constructing a *correlation matrix* of the exogenous factors and checking the pairs to see whether there is a *strong correlation* or not: that means values either *above 0.7* or *below -0.7* with a possible range of -1 to 1. Based on the results, we should decide *whether to remove* any of the variables in case it causes a strong correlation with another and its exclusion does not lead to significant information loss.

In the last step four of stage 4, we again run the tests for *autocorrelation, heteroscedasticity and normal distribution* on the last modification of the model (see the procedure in step three of stage 3). If the results are satisfactory, then we accept the tested configuration as final and proceed to the *interpretation* of the results obtained. Otherwise, we consider whether to make any new *structural changes* provided the potential to find a more optimal configuration that is still existent. It is possible that the efforts for the latter might go beyond the scope of the present research, so, in this case, we should settle for the last best configuration and proceed with analysing the results thereof.

In the last stage of the current methodology, **stage five**, we *review and interpret* the findings for the *observed interrelations* between the HPI and the other variables in the final variant of the model.

4. Empirical results from applying the model

The methodology was executed with an initial data set of *35 indicators*, which were supposed to influence in a direct or indirect way the nationwide *House Price Index* in Bulgaria. These 35 factors, distributed into *9 cluster groups*, were potential participants in the model, but only a few were chosen after applying the sorting out procedure. Thus, we succeeded to identify the exact *endogenous* and *exogenous factors* to be used together with the HPI in the construction of the EEMHP (Ivanov, 2024). The selected factors are displayed in Table 1.

Table 1.
Endogenous and exogenous variables included in the EEMHP

ENDOGENOUS		EXOGENOUS	
Y	House price index	F02	Consumer price index
F10	Average salary	F03	Index of housing rents
F15	Current account of the balance of payments	F09	Unemployment rate
F21	Volume of new housing loans	F22	Number of registered mortgages in the country

After identifying the variables, we can build the model itself. The focus of the research is that single equation from the dynamic system, in which the *dependable variable* is the *change of house prices* ($\Delta\gamma_t$). We express it with the following notation:

$$\Delta\gamma_t = \varphi_\gamma(\lambda_1\gamma_{t-1} + \lambda_2 F10_{t-1} + \lambda_3 F15_{t-1} + \lambda_4 F21_{t-1} + \lambda_5) + \qquad (5)$$
$$+\beta_{\gamma,1}\Delta\gamma_{t-1} + \beta_{\gamma,2}\Delta F10_{t-1} + \beta_{\gamma,3}\Delta F15_{t-1} + \beta_{\gamma,4}\Delta F21_{t-1} + c_\gamma +$$
$$+\beta_{\gamma,5}\Delta F02 + \beta_{\gamma,6}\Delta F03 + \beta_{\gamma,7}\Delta F09 + \beta_{\gamma,8}\Delta F22$$

Running the model with the real values of the above listed variables for the period 2005-2022 brings out the result in Figure 2. What we see on the graph is the *modified equilibrium HPI* (dashed line) suggested by the EEMHP (Y'), *the actual HPI* (solid line, Y) and their *difference* (columns) that is a projection of the house price deviation from the equilibrium. As we see on the graph, a clearly pronounced *real estate bubble* was observed in the period 2007-2008. After it burst out, there was a price plunge, followed by a prolonged period of *stagnation*. Eventually, it turns out that the model *precisely identified* the boom and the burst phase of the first contemporary house price bubble in Bulgaria. This outcome confirms that the model is reliable.

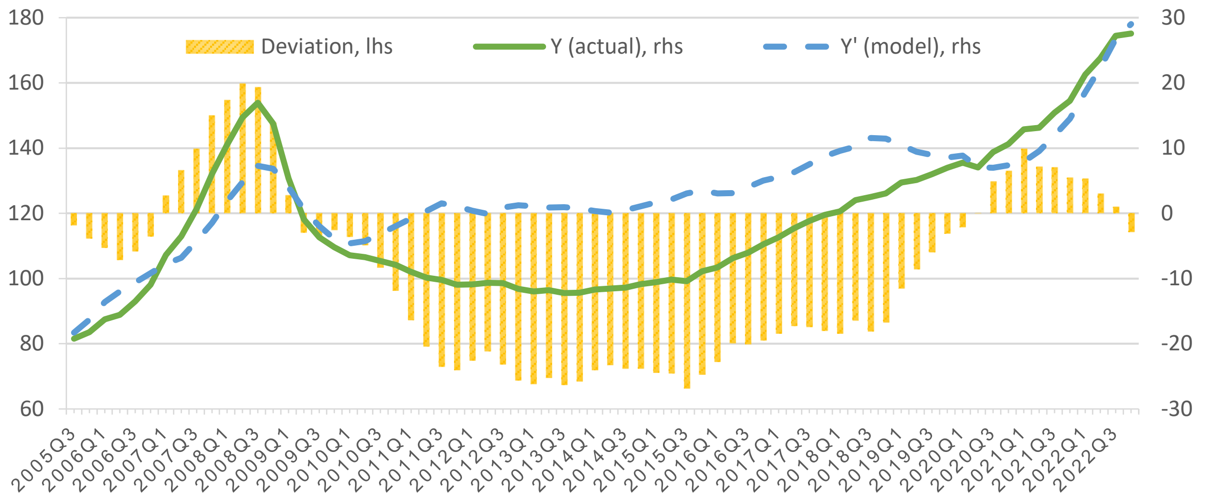


Figure 2. Deviation of house prices from their equilibrium levels according to the EEMHP

Legend: lhs – left hand side axis, rhs – right hand side axis;

The application of the model refutes the prevailing *public speculations* about an *existing housing bubble* that has been inflating since the start of the COVID-19 pandemic. No clear evidence for that was found indeed. Moreover, the model suggests that after some anomalies in house prices in 2020, they were once again returning to their equilibrium. Even at the end of 2022, they *fell* just below the equilibrium level despite the fact they were in an upward trend. The possible explanation of this is that the *surge* of house prices *is not big enough* to catch up the *increase of the other factors* in the system, among which: the growing *inflation* and *salaries*. The model even tells us that some dwellings were *a little underpriced* at the end of the period. To conclude, *no price bubble* seems to be of an immediate threat, at least not for the nationwide HPI. The actual situation, however, could be different for separate cities, especially the capital.

Based on theoretical and empirical evidence, we might well make a link between the probability of house price bubbles formation and the occurrence of *anomalies* in the house price cycle.

Conclusion

House price bubbles occur as a result of *hyperactive behaviour* from the demand side, i.e. the *demand function* approaches a state of *hyper-inelasticity*. Such a phenomenon is typical during the market's booming phase. Its negative consequences, which would often feature a wider economic and financial impact, are associated with the sharp onset of a *severe* and *prolonged phase* of *price correction*.

The procedure for building an *equilibrium econometric model,* presented in this article, reflects the complex behaviour of house prices under the aggregate and dynamic influence of heterogeneous groups of factors. The model has the potential to detect early signals of *house price bubbles*, so that the negative economic and financial effects from its spread could be mitigated with proper actions.

In fact, house prices are often out of *balance*. As the results from the application of the EEMHP point out, *achieving an equilibrium* for house prices is only an *attractive anchor,* but because of their volatile nature, equilibrium is hardly reached and not stable over time. Among the reasons for this are the complex *mechanism of real estate pricing* and the complex *transmission* of effects between *the real and financial sectors*, due to the intertwined interaction of multiple factors. This tends to generate difficulties and challenges in terms of constructing an appropriate econometric model to account for such centrifugal forces.

The model proved its capacity in finding out price abnormalities as it exactly detected the house price bubble in Bulgaria in 2007-2008. Based on a solid methodology, it is capable of bringing out well-argued outcomes. Such an instrument would nicely complement the toolkit of the central banks when determining their policies on credit standards and monetary actions. Nevertheless, almost every econometric model the EEMHP has its *structural drawbacks* and *limitations* for use. That is why, its results should not be taken as absolute. They are just approximations that should be considered in *combination with other methods* and above all to be assessed with an *expert judgement*.

With all that said, the current study accomplished its goal. We successfully constructed, applied and tested the EEMHP in *assessing the deviations of house prices* and for *detecting house price bubbles*. We also presented the empirical results from the application of the EEMHP for the country specific case of Bulgaria. Additional research on the topic could further elaborate the model and supplement it with additional diagnostic tools.

This work was financially supported by the UNWE Research Programme (Research Grant No. 20/2023).

References

Aastveit, K. & Anundsen, A. (2022). Asymmetric Effects of Monetary Policy in Regional Housing Markets. *American Economic Journal: Macroeconomics, 14(4*), pp. 499-529. DOI:10.1257/mac.20190011

Beraldi, F. & Zhao, Y. (2023). The Pricing-Out Phenomenon in the U.S. Housing Market. *International Monetary Fund: IMF Working Paper WP/23/1*. DOI:10.5089/9798400229039.001

Bozhinova, M., Iliychovski, S., & Filipova, T. (2010). Prilozhni aspekti na koncepciyata za upravlenie na baza stoynostta po primera na publichnite druzhestva v Bulgariya. *Almanah Nauchni izsledvaniya, 13, pp.34-85.* [Bozhinova, M., Iliychovski, S., & Filipova, T. (2010). Practical aspects of the concept for value-based management of Bulgarian public companies. (In Bulgarian) *Scientific Research Almanac, 13:pp.34-85*].

Bespalov, S., Mitiushkina, K., & Chernichenko, H. (2021). Ecological and economic development dimensions of the countries with the use of cluster analysis tools. *D. A. Tsenov Academy of Economics – Svishtov: Business Management, 3/2021*.

Cevik, S. & Naik, S. (2022). Don't Look Up: House Prices in Emerging Europe. *International Monetary Fund*. DOI:10.5089/9798400227042.001

Corradin, S. & Fontana, A. (2013). House price cycles in Europe. *European Central Bank, Working Paper Series, No.1613*.

Egert, B. & Mihaljek, D. (2007). Determinants of house prices in Central and Eastern Europe. *BIS Working Paper 236, pp. 1-26*. DOI: 10.1057/palgrave.ces.8100221

Eigner, P. & Umlauft, T. S. (2015). The Great Depression(s) of 1929–1933 and 2007–2009? Parallels, Differences and Policy Lessons. *MTA-ELTE Crisis History Working Paper No.2, pp. 1-53*. DOI: 10.2139/ssrn.2612243

Guo, X., Shishmanova, P., Orlova, A., Nesenenko, P., & Mankuta, Y. (2024). Technological theories in economics and management: evolution and applied aspects. *D. A. Tsenov Academy of Economics – Svishtov: Business Management, 1/2024*. DOI: 10.58861/tae.bm.2024.1.0 1

Ivanov, P. (2024). *Izsledvane na zhilishtnite tseni i obuslavyashtite tyahniti izmenenie faktori* (Doktorska disertaciya, Universitet za natsionalno i svetovno stopanstvo), DOI: 10.13140/RG.2.2.15906.85444. [Ivanov, P. (2024). *Research on house prices and the factors determining their* *change* (Doctoral dissertation, University of National and World Economy), DOI (in Bulgarian): DOI: 10.13140/RG.2.2.15906.85444].

Iliychovski, Sv., Filipova, T., & Petrova, M. (2024). Application of Mathematical Models and Hierarchical Relationships Method of Residential Properties Valuation in Bulgaria. *4th International Conference on Sustainable, Circular Management and Environmental Engineering (ISCMEE 2024), Izmir, Turkey. E3S Web of Conf*. Volume 558, 2024, 01019, p.12, DOI: 10.1051/e3sconf/202455801019

Iliychovski, Sv., Filipova, T., & Petrova, M. (2025). *Investigating the influence of determinants on market valuation multiples of non-public companies in Bulgaria*. In Emilia Alaverdov, Muhammad Waseem Bari (Eds.). Cultural Heritage Protection and Restoration in Conflict and Post-Conflict Zones. Chapter 15, ISBN: 9798337325408. DOI: 10.4018/979-8-3373-2540-8.ch015

Jarmulska, B. et. al. (2022). The analytical toolkit for the assessment of residential real estate vulnerabilities. *European Central Bank*: *Macroprudential Bulletin*, *vol. 19*.

Kotseva, P. & Yanchev, M. (2017). Analysis of the Housing Market Developments and the Underlying Macroeconomic Fundamentals in Bulgaria. *Bulgarian National Bank, Discussion Papers, No. 103*.

Lee, C. & Park, J. (2022). The Time-Varying Effect of Interest Rates on Housing Prices. *Land, 11(12)*. DOI: 10.3390/land11122296

Li, N. & Li, R. (2022). A bibliometric analysis of six decades of academic research on housing prices. *International Journal of Housing Markets and Analysis*, 17(4). DOI:10.1108/IJHMA-05-2022-0080

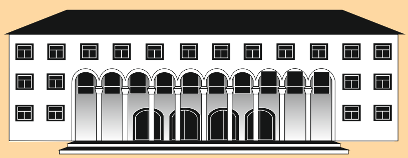
Nikolaev, D., & Petrova, M. (2021). Application of Simple Convolutional Neural Networks in Equity Price Estimation. *2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, 2021, pp. 147-150, DOI: 10.1109/PICST54195.2021.9772160.

Petrova, M., Popova, P., Popov, V., Shishmanov, K. & Marinova., K. (2022). Potential of Big Data Analytics for Managing Value Creation. *2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES),* 2022, pp. 1-6, DOI: 10.1109/CIEES55704.2022.9990882

Yovkova, Y. (2009). Vliyanie na ipotechnoto kreditirane varhu ikonomikata. *Pazar i upravlenie na nedvizhimata sobstvenost v savremennite uslovia*. Sofia: Avangard Prima, str. 9-15. [Yovkova, Y. (2009). Impact of Mortgage Lending on the Economy. *Real Estate Market and Management in Modern Conditions*. Sofia: Avangard Prima, pp. 9-15].

ISSN 0861 - 6604
ISSN 2534 - 8396

BUSINESS management



PUBLISHED BY
D. A. TSENOV ACADEMY
OF ECONOMICS - SVISHTOV

3/2025

BUSINESS management 3/2025

The printing of the issue 3-2025 is funded with a grand from the Scientific Research Fund, Contract KP-06-NP6/29/04.12.2024 by the competition "Bulgarian Scientific Periodicals - 2025".

Submitted for publishinq on 16.09.2025, published on 18.09.2025,
format 70x100/16, total print 80

BUSINESS management

D. A. Tsenov Academy of Economics, Svishtov

Year XXXV * Book 3, 2025

CONTENTS