MBH JELZÁLOGBANK

# A methodological guide of the MBH House Price Index

www.mbhindex.hu

# Table of contents




**Published:** The first publication of the MBH Index (originally FHB Index): 1 October 2009.

**Updated:** 2 May 2023.



The methodology was established based on an order placed by MBH Mortgage Bank Co. Plc., in the ELTINGA professional workshop. Dóra Barát, Áron Horváth and Gábor Révész have also taken part in doing statistical calculations.


MBH JELZÁLOGBANK

# 1. The MBH House Price Index

The MBH House Price Index measures the development of Hungarian residential real estate prices. It shows how the value of Hungarian residential properties has changed since 1998.
The index is published quarterly, and starts from first quarter 1998. At the time of quarterly disclosures, the time series ends with the data of the previous quarter. I.e. at the first publication, in October 2009 it contained 46 pieces of data, from which the last one applies to the 2nd quarter of 2009. The value of the index is normalised with the average of year 2000, i.e. the average index value in 2000 is 100.

During the development of the index, we regarded such long-standing, internationally well-known indices as examples as the house price index of Halifax[1] or Hypoport AG2. We made it our aim to prepare the MBH House Price Index on similar standards, and to make it an unavoidable reference in the profession. This methodological guide was written for those who are interested in details of the construction of the index. Here we present the structure and the features of the database used for the calculations. We briefly demonstrate the hedonic method applied when computing the index, with the help of which we developed an indicator of actual price movements from changeable data. Finally, we share some conclusions and interesting points which we experienced during the process. The description is closed with the planned direction of our upcoming work, as certain index parts can be formed from the compiled database as well that could be subject to the interest of the profession.

_____

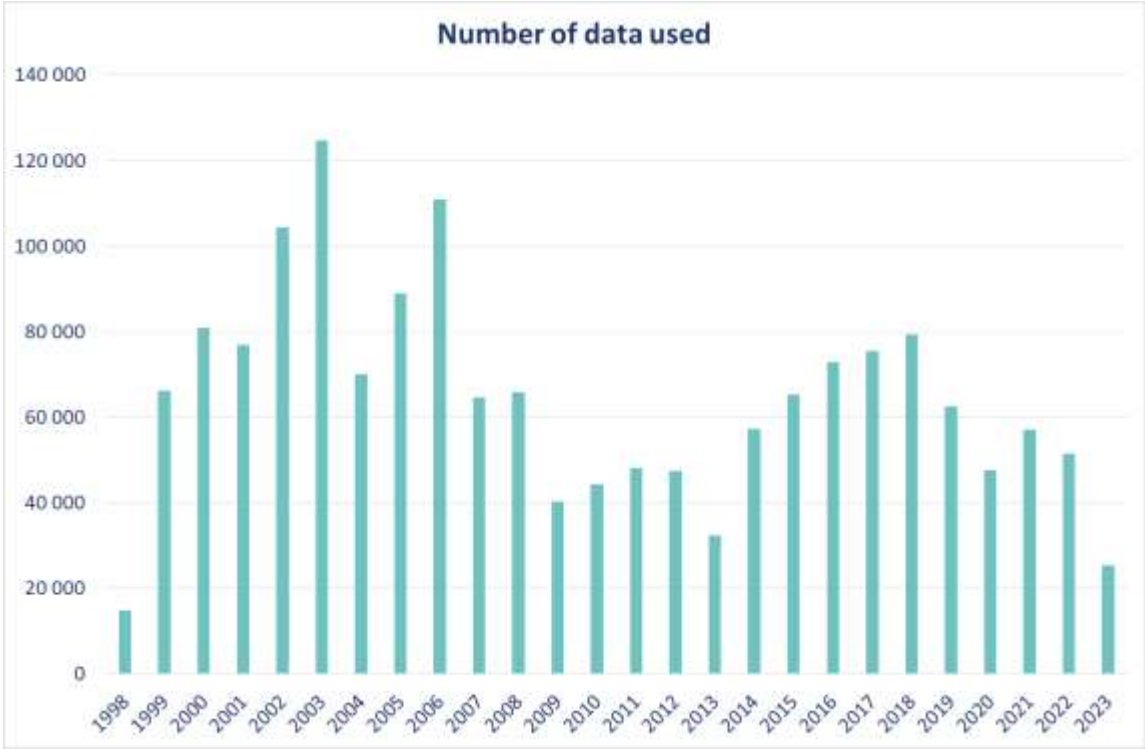
1. https://www.halifax.co.uk/media-centre/house-price-index.html
2. https://report.europace.de/index-epx-mean/

# 2. Data used

In the following, the range of data sources and data used in the production of the MBH House Price Index and the changes in the information set are described.

**a) Source of data, data**

The index is based on purchase-and-sell transaction data of residential real estate, and was prepared processing purchase-and-sell data of more than 1.2 million residential properties from some 3100 settlements from 1998 onwards. Both the own observations of MBH Group and a database purchased from the National Tax and Customs Administration served as sources of data. The following table shows the yearly distribution of the data used.[3] When providing loans for the purchase of real estate, MBH Group devotes special attention to the registration of features of the properties.




[3] We must note that the number of data used unfortunately does not show the number of Hungarian housing transactions from the given year clearly, because the data supply from the National Tax and Customs Administration, in our opinion, was not full in several quarters.

# 2. Data used

During valuation for the assessment of loan-to-value, several attributes of a property are investigated (among others):

- address, location of the real property,
- area (gross, net, useful) of the real property,
- area of accompanying parcel,
- year of construction,
- distance from public transportation (train, local- and distance bus,
- condition,
- technical features of the built construction (walls, roof structure, doors and windows, heating),
- access to public utilities.

These data allow for the application of the hedonic method (to be discussed later).

Another part of the database comes from the National Tax and Customs Administration. Every single Hungarian residential property transaction must find its way to this database.

The data recorded include the following:

- date of the purchase,
- address (settlement, district if in Budapest),
- purchase price documented through transaction,
- sum of valuation by a stamp duty office,
- type of building (detached house, terraced house, condominium or block flat),
- area.

However the incoming data are rather incomplete. The area, which is indispensable for the calculations, is missing in more than a quarter of the database. An even more serious problem is that the quality of available data is poor, errors can be identified in numerous cases: often an impossible sum is stated to be the purchase price.

# 2. Data used

That's why the dataset is first analysed by experts of MBH Group one by one, and obvious errors are corrected and complemented based on formerly accumulated data (for instance, earlier information is often available about the area of the property). Afterwards, data that are useless from a statistical point of view are filtered out from the still deficient database using automated filtering conditions, and remaining data are included in the calculations.

The filtering conditions concern the following:

- the transaction should be between private persons,
- the whole ownership should be transferred,
- an existing date should be connected to the purchase price,
- an area should be belong to the purchase price,
- the price should be reasonable,
- the price of one area unit should be reasonable.

In the 4$^{th}$ quarter of 2007 and in the beginning of 2008, we faced extreme deficiencies in the database,

Probably because the transfer of data from the stamp duty offices to the National Tax and Customs Administration (Tax and Financial Control Administration back then) took place at that time. The price movements of this period were finalised relying on the expertise of the Hungarian Central Statistical Office.

# 2. Data used

**b) Change of the information set**

The two sources of the database are available with a significant shift in time. MBH Group registers its data with special care, so based on them virtually up-to-date data are provided. However, a large part of data coming from the NTCA finds its way into the database at least half a year after the purchase transaction, and not an insignificant number of data have even arrived years after the transactions. On the other hand, we put special emphasis on the index reflecting even the newest state of the housing market, thus we do not wait for all the purchase-and-sell data of the latest period when preparing the index.

Based on our research, we concluded that even if there are fewer observations, it carries substantial information, so we opted for the earliest publication possible. Using the data arriving continuously from the National Tax and Customs Administration, we recalculate the values of the index that are for the latest period, and as such, these may be subject to modification. Our uncertainty due to an incomplete information set is indicated by dashed lines at the end of the graph.

# 3. A The hedonic calculation method

For computing the index, we applied the process based on the hedonic method, which is usual in the profession, and the classic reference of which is the study of Kain and Quigley[4]. The method serves to filter out the composition effect appearing in observations of simple indicators (mean price, median price). The composition effect itself arises, because not every single property is sold in every quarter, thus our observations – even if the transaction data are complete – are accordingly a restricted sample of the real estate stock, this way the sample is not representative. The hedonic method is the most widespread statistical method to measure aggregate property price change, and is based on the theoretical consideration that the value of houses/flats can be divided into the values of their attributes.

So if we assume that the price of real estate depends on its attributes, we get the general change of the price level by filtering out the effect of the change of features in the sample from the change of average price. A simple example demonstrates the point of the method.

Let's suppose that in the country of Residency, flats of unique design are more expensive than flats of the same type. Altogether six flats were built in Residency. The table below shows the distribution of the two types and the value of the flats.

| Of same type | Of unique design |
|---|---|
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |

[4] Kain, John F. – John M. Quigley [1970]: Measuring the Value of Housing Quality. Journal of the American Statistical Association 65 (330), Jun 1970, pp. 532-548.

# 3. The hedonic calculation method

The value of the flats is stable, but it is random, which ones are purchased in the different periods. In the first period, two flats alike and one unique flat changes owners:

| Of same type | Of unique design |
|---|---|
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |

In this case, the average price of the observed transactions is: (10 + 10 + 20) / 3 = 13.3 thousand dollars.

In the second period, the market gets momentum, and all the flats change owners.

| Of same type | Of unique design |
|---|---|
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |
| | 20 thousand dollars |
| 10 thousand dollars | 20 thousand dollars |

Here, the average price of the observed transactions is: (2 * 10 + 4 * 20) / 6 = 16.6 thousand dollars. So the observed mean price has risen, because more flats of unique design were in circulation in the second time period. I.e. the composition of traded flats changed, and thus the composition effect increased the average price despite the nationwide valuation of the flats remaining unchanged.

# 3. The hedonic calculation method

The hedonic method deals with this problem by trying to identify the value of attributes that define the value of the flats. In the above example, the flats have one feature, namely their type. According to the hedonic method, we observe that two flats of the same type cost 10 thousand on the average in the first period, while the only one of unique design costs 20 thousand. In the second period, again the quality of traded flats may be observed: the average price of flats of the same type is again 10 thousand, and that of unique flats is 20 thousand.

The conclusion of the method now is that the price of the flats did not change, because neither the price of alike nor the price of unique design “attribute” changed from the first period to the second.

The larger the sample available, and the more properties of the flats that are observed, the greater the chance of detecting variations due to changes in composition. The method can therefore also be used to detect the composition effect that the number of transactions of cheaper housing may increase at the time of the introduction of a housing subsidy scheme linked to the value threshold, which reduces the observed average price ceteris paribus.

When applying the hedonic method, we need to run a regression. In the regressions, the prices of properties (or because of the skewness of the price distribution, the logarithms of prices, log pj) are explained by their observable features.

$$log\ p_j = b_0 + b_1 q_{1j} + b_2 q_{2j} + \ldots + b_p q_{pj} + \sum_{i=2}^{t} m_i x_{ij} + u_j$$

The variables indicated with q are the attributes of the flats (size, type or location), and coefficients b are coefficients of the attributes drawn in as explanatory variables. The coefficients show how one more unit from the attributes (e.g. one more bathroom) raises the price of the property on the average. That’s why these coefficients are referred to as shadow prices of the features.

# 3. The hedonic calculation method

The above specification is a so-called restricted hedonic regression, because b coefficients have no time index, i.e. shadow prices are unchanged over time. Finally, xij variables in the summa part are the dummy variables showing the period of observation, i.e. xij = 1 if the $j^{th}$ property was sold in the $i^{th}$ period. Accordingly, the mi estimated coefficients explain the "surplus value" of the given year, and the price index will be the result of raising the coefficients to the exponential power. I.e. the coefficients got by the model calculus, connected to the year-dummies, express the price change compared to the reference year.

The hedonic method gives us a more accurate measurement than the more simple indicators (averages) if we manage to involve variables into the analysis that help tackle the distortion of the composition effect. As it is difficult to describe several qualitative parameters with a quantitative indicator, or there is no indicator available, it is common to include proxy variables among the explanatory variables.

(For instance, the address is not a hedonic qualitative feature, still it contains important information. It does make a difference whether a flat or house is in the city centre or in the suburbs of Budapest.) The reason these variables are important is, because the composition effect mostly arises from the change of the settlement distribution of purchase transactions. Thus with the help of variables drawn into the regression, it can be filtered out if in some periods there were more transactions from Budapest than at other times. The set of variables was formed so that it also tackles the geographical, settlement size-related, property type-related and flat size-related instability of the composition. Where data are available, besides these it is possible to correct the composition effect resulting from the instability of the year of construction, construction technology, and the location within the settlement.

# 3. The hedonic calculation method

The values of MBH House Price Index were computed by creating sample parts that comprise 5 consecutive quarters from the complete database containing 100 quarters. That means consecutive sample parts overlap for 4 quarters. The above regression was estimated using these sample parts with 5 quarters, and the value of the index was always calculated based on the estimation, in the sample of which the given quarter is the last piece of data. This method ensures that after one and a half years, we can seal the past, as from now on the new data do not modify the previously estimated index values any more.

# 4. Conclusions, remarks

In the following, we are going to share some observations that we deem instructive, or may answer some questions that may come up.

First, through two figures we illustrate why we chose the hedonic method for the index. Figure 1 features the changes of the hedonic index and the changes of mean and median prices of the sample. The trends of simpler methods and MBH House Price Index calculated with the hedonic method are similar, they share the same conclusions in the long run. The figure also demonstrates that in the short run, simpler indicators have a bigger volatility, while the index produced with the regression method has a smoother run, i.e. changes in the composition of sold properties, which is corrected by the hedonic method, may have a role in short-term volatilities.

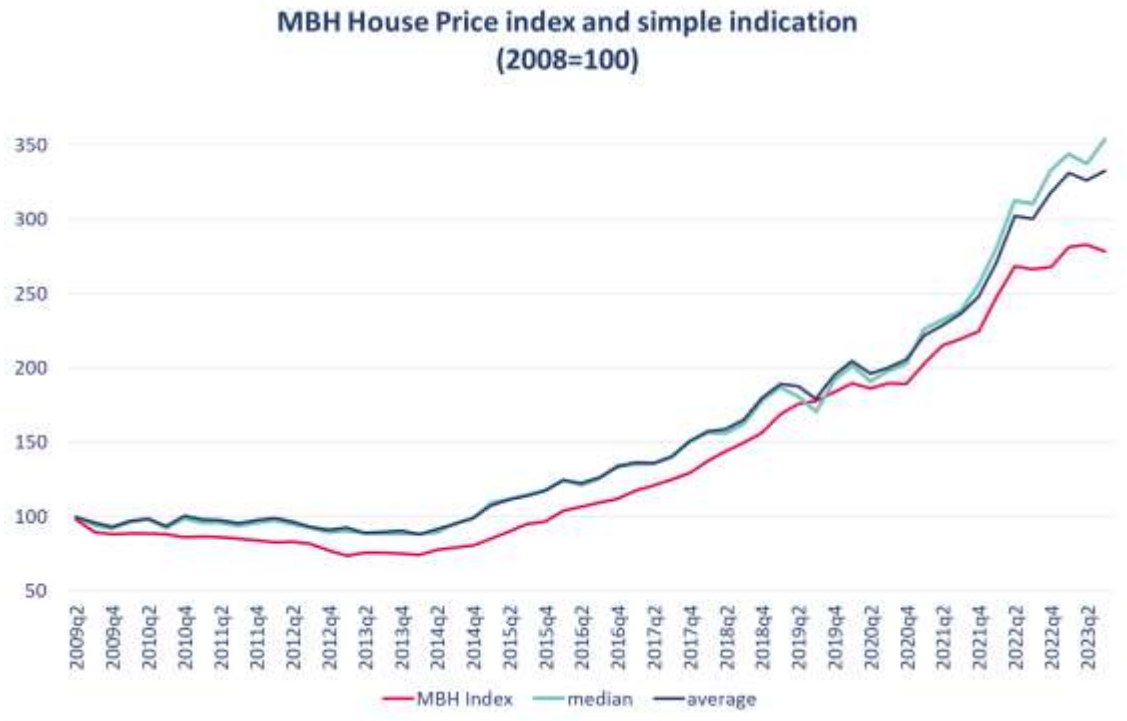



Figure 1: run of MBH House Price Index and simple indicators (2000=100)

MBH JELZÁLOGBANK

# 4. Conclusions, remarks

In Figure 2, yearly changes of the previous indicators can be seen. For us, the difference between 2000 and 2003 was especially instructive. During these two years, the mean prices rose similarly, while the value of the hedonic index rose significantly more in 2000.

Having analysed the data, it turned out that the number of flats in Budapest – which are relatively more expensive – was much lower in the database in 2000 than usual. As a result, although house prices increased significantly both in the countryside and in Budapest, there were relatively more properties from the countryside (cheaper properties) among the observations, and this change in the composition decreased the rise of the average price.

The biggest differences are to be seen in data from 10 years ago, as in this time period the supply of data was even less complete than today.

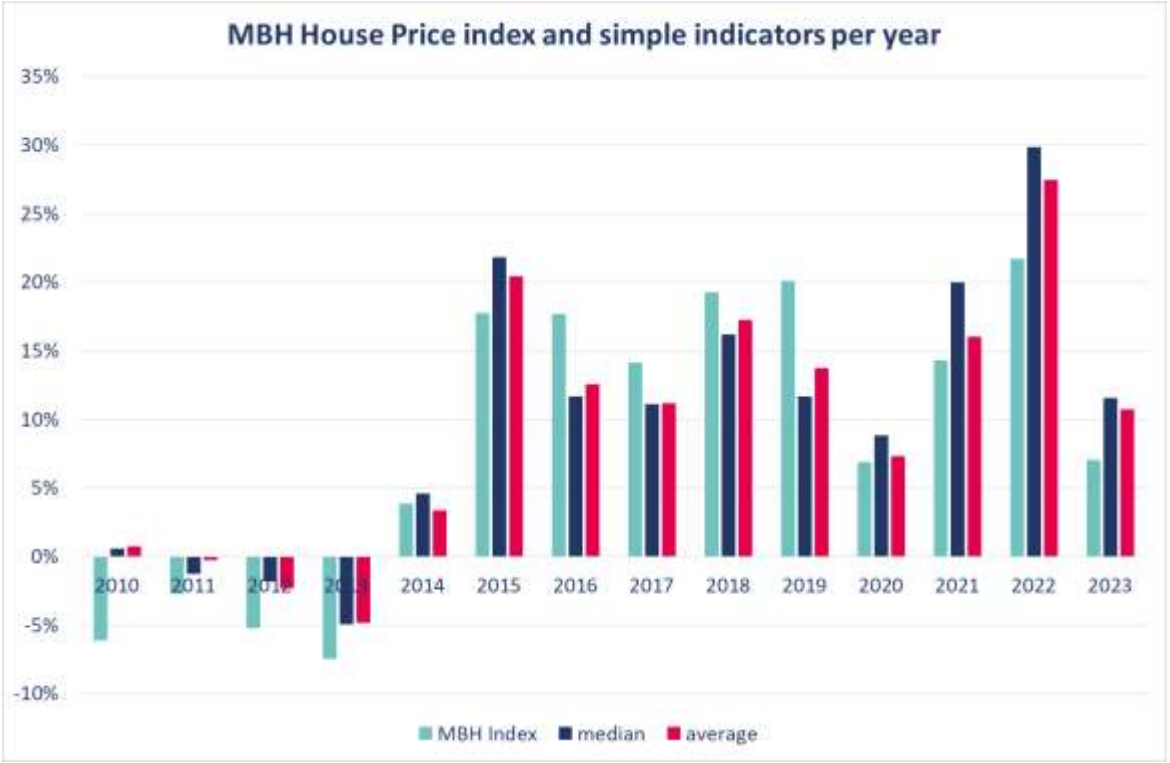



Figure 2: run of MBH House Price Index and simple indicators compared to previous year (%)

MBH JELZÁLOGBANK

# 4. Conclusions, remarks

As a second conclusion, we would like to mention that having analysed the seasonality of the index, we found that the index contains only a minimal seasonality, so we decided to publish seasonally unadjusted data. Figure 3 shows what result the seasonal adjustment with the Tramo-Seats method brings.

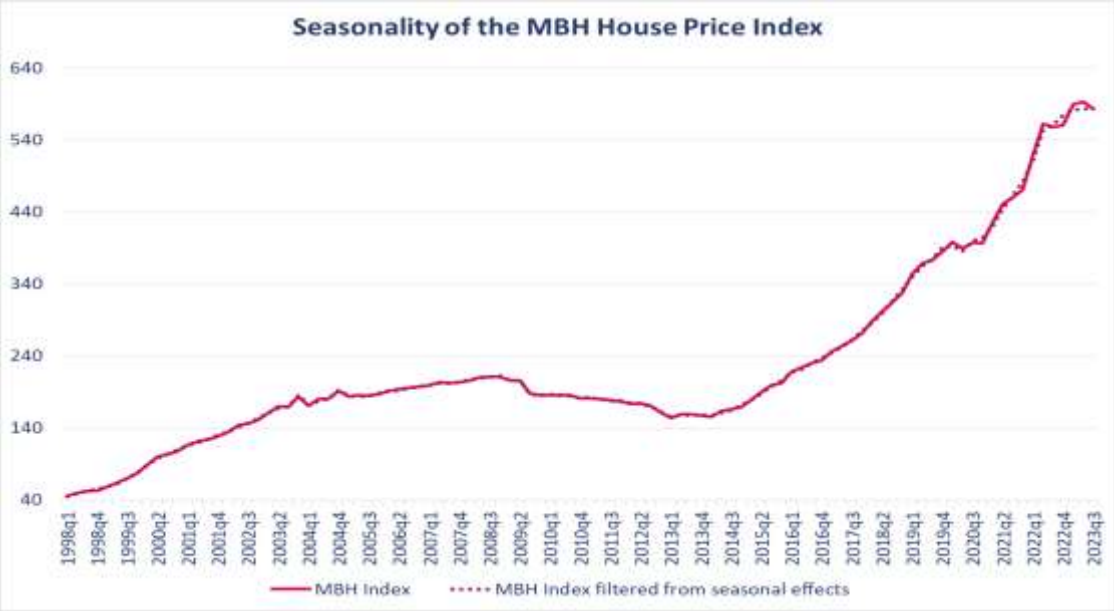



Figure 3. run of MBH House Price Index and its seasonally adjusted version (2000=100)

MBH JELZÁLOGBANK