

## COMPARATIVE EVALUATION OF PUBLIC DATA SETS ON AVERAGE HOUSE PRICES IN POLAND

---

Sebastian Kokot

*University of Szczecin*

*e-mail: sebastian.kokot@usz.edu.pl*

*ORCID: 0000-0001-7312-0984*

Received 25.04.2024, Revised 28.02.2025, Accepted 5.03.2025

---

### Abstract

**Research background:** Access to official information on the real estate market, backed by the authority of public institutions, in particular on average housing prices, is an important element of the real estate market regime. In Poland, such information is published in various spatial, temporal and subject ranges by the National Bank of Poland (NBP) and the Central Statistical Office (CSO). Recipients of such data expect high quality data. This paper presents a comparison of data on the average prices of 1m<sup>2</sup> flats in 16 Polish cities published by these two state institutions.

**Purpose:** The aim of the research is to assess and verify the consistency of information published by the NBP and the CSO on listed prices of dwellings on the secondary and primary market with regard to the possibility of using them to make investment decisions and market analyses.

**Research methodology:** Relative and absolute differences between the average housing prices published by the CSO and the NBP were calculated. an analysis of the structure of these differences was carried out.

**Results:** Large differences in the data published by the CSO and the NBP make it necessary for these institutions to subject the entire process of creating information on average housing prices in individual local markets to a thorough analysis.

**Novelty:** The study indicates that public datasets should not be used for further analysis and research indiscriminately.

**Keywords:** real estate market, housing prices, public data sets, Poland

**JEL classification:** H49, K11, R30

## **Introduction**

The important function of the real estate sector in the national economy is unquestionable, and the situation in the real estate market plays a significant role in the life of almost everyone, as the living, professional, infrastructural, cultural and other needs of society are met thanks to real estate. These are sufficient reasons to make an effort to closely observe the phenomena occurring in the real estate market. For some professions, the analysis of these phenomena is the basis of the workbench. These professions include real estate appraisers, real estate agents, analysts working for financial and insurance institutions, investment advisers, developers, tax officials and, of course, investors themselves. For them, access to good quality information on real estate transactions is an essential condition for the reliability of their work. Irrespective of the needs of professionals, many private individuals are also interested in the real estate market situation (Francke, 2010) and make widespread use of the studies and analyses made available. In practice, however, access to reliable source data on the real estate market, and thus the possibility to carry out reliable analyses, is difficult. This is due to a number of conditions of various nature, which should be broadly divided into two categories. The first one results from the nature of the real estate market and concerns both the specifics of real estate as objects of trade and the very manner of concluding transactions. The second arises from organisational and technical problems of obtaining, processing, collecting and making available data on market transactions.

Real estate transactions occur relatively infrequently. The observation of prices is carried out in a specific way, as individual transactions concern diverse objects and, in addition, occur at irregular, one could even say somewhat random periods (Kokot, 2015). The relative infrequency of transactions on the real estate market is one of the characteristic features of this market and becomes a major problem when one wants to analyse the market with statistical tools. In the classical approach, the analysis of the dynamics of a given phenomenon is conducted on the basis of a time series, i.e. sequences of observations showing the development of the phenomenon under study in successive periods (days, months, quarters, years, etc.). Each period is assigned the value of the observed variable revealed in that period. If the observed variable is real estate prices, the matter becomes more complicated, as in a given period, e.g. a quarter, on a given observed market, there are several, several dozen or more transactions recorded, which concern various objects, of various characteristics – location, area, standard, technical condition, etc. How, then, based on a set of such dispersed variables, can the observed phenomenon be analysed? How, then, on the basis of a collection of such dispersed information,

can one build a valid time series? In practice, a single value representative of the period is sought. The typical procedure for this is to use the arithmetic average of the prices recorded on the market. However, such averages are tainted by the randomness of the transactions made during the period. For example, the fact that the transactions of properties with an attractive location dominated in a given period by chance, or the fact that a developer has just completed a development and is selling flats. As a result, we observe fluctuations in the calculated average prices that are not explained by market conditions.

In addition to the problems related to the analysis of prices on the real estate market resulting directly from the specifics of the real estate market and real estate as objects of trade on the market, we also face problems resulting from the legal, organisational and technical conditions of the principles of keeping proper registers of real estate prices. Pursuant to the Act of 17 May 1989 Geodetic and Cartographic Law and the Ordinance of the Minister of Development, Labour and Technology of 27 July 2021 on land and building registers, *starosts (local authorities)* or city presidents maintain the registers of real estate prices. Their source is the notarial deeds of real estate sales contracts sent by notaries. The adopted system of collecting and processing data on transaction prices generates further problems. The main ones are the poor and non-standardised property information contained in notarial deeds and the long delays, sometimes reaching over a year, in entering information from notarial deeds into the relevant systems. Transactions requiring two contracts are also a certain problem, which we encounter in particular when analysing the so-called primary market, i.e. involving the transactions of newly built houses or premises. The buyer's decision to purchase a given property is often taken while the construction is still in progress, sometimes not very far advanced. It results in the conclusion of a so-called preliminary contract, in which the developer undertakes to build and transfer the ownership of a specific property to the buyer, and the buyer undertakes to pay the agreed price. The transaction price is thus determined at the stage of concluding the preliminary agreement, and the agreement to sell the property usually takes place many months or even years later. During this time, market conditions have changed, which may mean that if the decision to purchase the property was made 'today' the contract would be for a different amount. Formally, however, the date of the contract transferring ownership of the property, which shows the price agreed in the past, is considered to be the date of the transaction.

Problems resulting from the legal, organisational and technical conditions of conducting real estate price registers cause that, due to a combination of many circumstances, published market data may be erroneous, unreliable, unreliable, outdated, and thus affect the quality

of analyses prepared on their basis. Meanwhile, conducting analyses of the residential market requires having reliable and complete data. Sources of data on prices, but also on the characteristics of dwellings, can be broadly divided into two categories (Konowalczyk, Ramian, 2012):

- primary data – data collected by analysts at the point of recording individual market events,
- secondary data – data that have been collected as primary data by a specific entity and subsequently processed and made available in that form.

Other classifications may refer to the subject and geographical divisions of real estate and markets, e.g. the market for flats in a municipality, the market for flats as separate objects of ownership, the market for co-operative ownership rights to dwellings, the rental market, or types of prices (market prices, non-market prices, and offer prices). The primary data source used for property market analyses should be primary data on market transactions. Transaction price data can be obtained from county offices within the structures of which there are geodetic documentation centres that maintain the registers of real estate prices. Different centres in Poland maintain their registers in different IT systems, which makes it difficult to obtain data systematically from different centres. Usually, the data are made available in electronic form in pdf or xls format files, but it happens that the sharing format requires a special application. The method of sharing usually prevents the data from being used for analysis directly, which means that the data requires very labour-intensive ‘manual processing’. In practice, therefore, data from this source can only be effectively used for the analyses of selected local markets. Another significant shortcoming of the data collected in price registers, especially from the point of view of analyses of the housing market, is that they do not include prices recorded in the trade of flats as objects of cooperative ownership rights to premises. Such information goes to housing cooperatives, which are not obliged to make it available to researchers. Even assuming the good will of housing cooperatives to make the data available, obtaining complete data would require cooperation with all housing cooperatives operating in the local market.

In addition, data on real estate transactions can be found at tax offices, land and mortgage courts, notary offices and estate agency offices. However, they do not maintain systematic transaction databases. In practice, however, these are sources that are not available to real estate analysts. Some of them – notary’s offices and estate agents’ offices have information only on those transactions which they themselves have handled. The difficulty of obtaining complete primary data, the time-consuming nature of its processing and sometimes the high costs of obtaining such data often lead analysts to use secondary sources, particularly publicly

available sources. The main drawback of data from such sources is the inability to gain an insight into the mechanism of their generation and thus into the quality of the primary data on which they are based. They also do not offer the possibility of freely regulating the scope of the data in terms of subject matter and geography. Usually secondary data are in the form of statistical parameters (e.g. arithmetic mean) calculated for a specific period (e.g. a quarter) and geographical area (e.g. a city). Currently, secondary data on housing prices in Poland can be obtained freely and free of charge from the following public sources:

- Central Statistical Office (CSO),
- National Bank of Poland (NBP).

The CSO publishes databases on many areas of the country's social and economic life. From the point of view of the main subject of this work, data on the median and average prices of 1 m<sup>2</sup> flats are valuable. Such data are published quarterly for counties, divided into primary and secondary markets. As the CSO indicates, the presented data on sales volume and prices of real estate have been compiled on the basis of information obtained from Real Estate Price Registers maintained by the county offices and mayors of cities with county rights, using information from notarial deeds. The data refers to sales of a market nature, i.e. sales on the free market (including, e.g. sales between related parties) and sales by tender. The data refers to dwellings located in multi-family buildings (colloquially flats in blocks of flats), and does not refer to information on single-family houses. As such Real Estate Price Registers do not record the transactions of co-operative ownership rights to premises, and which are traded, as it were, in parallel to the premises separated as a unit property, it should be assumed that these statistics do not take into account the trade in cooperative premises. As a result, the actual number of units sold is likely to be higher and average prices do not entirely correctly reflect market conditions. Nevertheless, with such caveats, it seems that they could be successfully used to, for example, investigate price dynamics and differentiation. The data have been published since 2010. The advantage of the CSO data is that, unlike other publicly published data, they are related to the district level. By using these data, researchers can save themselves a great deal of work associated with the extremely time-consuming acquisition and processing of individual data themselves. Especially if their intention is to conduct comparative analyses on, for example, counties, which would involve obtaining data from multiple county offices.

Another useful source of data on the residential market for researchers is the NBP, which publishes quarterly reports "on the situation on the residential and commercial real estate market in Poland" on its website as part of its cyclical analytical materials. These publications are accompanied by the availability of the residential real estate price database comprising

average unit prices in selected cities and groups of cities as well as indices of the price of 1 m<sup>2</sup> flats determined on their basis. The NBP database of residential property prices is created thanks to the submission of data by real estate agents and developers with the involvement of the NBP Regional Branches. The database constitutes valuable information material that can be used in analytical work in the area of this segment of the real estate market. It should be presumed that, in contrast to the CSO database, the data also include the transactions of co-operative ownership rights to residential premises. Moreover, and apart from the average transaction prices, the database also contains average offer prices, which gives an additional research field. On the other hand, the shortcomings of this database include the fact that it concerns the markets of only 17 large Polish cities. In addition, it is based on surveys, which in a way may also weaken its relevance.

In conclusion, when data on average unit prices in local markets is sufficient for analyses, one can use data offered by the CSO or the NBP, and if detailed data on individual transactions is necessary, one should refer to the district offices keeping records of such prices.

Analyses of housing market prices should clearly distinguish between two of its sub-segments: the primary market and the secondary market. Despite the fact that both markets offer, so to speak, the same goods – the rights to residential premises, they differ with respect to a number of features directly translating into differences in residential prices. The primary market is the market for new flats, just handed over for use, usually in the so-called developer's standard to be finished and arranged by buyers themselves. These are flats built using modern technologies, with the use of modern material and functional solutions, meeting the current standards. In addition to strictly market conditions, their prices are influenced by current land prices, prices of materials and construction work. The secondary market includes a wide range of already used flats, built in different years – also in the pre-war period, executed in various technologies.

Both sub-segments are influenced by the same main factors of a macroeconomic nature, such as the economic situation, inflation, the labour market situation, the income of the population, interest rates and the availability of credit, etc., which should be analysed separately. Nevertheless, they should be the subject of separate analyses and their failure to distinguish should be regarded as a serious error in analytical assumptions. In both sub-segments, we can clearly observe differences in unit price levels. The relationship between housing prices in the primary and secondary markets is not permanent and fluctuates relatively large and irregularly over time (Kokot, 2022; Doszyń, 2023).

Having at their disposal the two databases of average prices of 1m<sup>2</sup> flats, characterised in detail above, i.e. the CSO database and the NBP database, researchers are faced with the dilemma of which one to use for analytical work. Both databases are backed by the authority of a state institution. Theoretically, both databases present the same, or at least very similar, information. The differences relate to the method of obtaining input data. The CSO obtains them from price registers, while the NBP obtains them from surveys. Moreover, the CSO database does not use data on transactions in co-operative ownership rights to residential premises. For these reasons, some minor and natural differences regarding published prices should be expected. A natural way to verify their compliance is to compare published values and analyse their differences. This is the task behind the presented study, the aim of which is to verify the compliance of the information published by these institutions on the listed prices of residential units. Thus, the study is a comparative evaluation of public data sets on the average prices of flats in Poland, coming from two sources, from the point of view of their usability for further analyses. Analyses using these datasets are carried out in a number of academic centres in Poland and address a variety of issues, including price indices (Tomczyk, Widłak, 2010; Kokot, 2015), price convergence (Dittmann, 2014; Matysiak, Olszewski, 2019; Tomal, 2019), price anchoring (Kokot, 2023), price factors (Kokot, 2020; Melnychenko et al. 2022), price forecasts (Belej, 2023), comparison and evaluation of price differentiation (Widłak, Nehrebecka, 2011; Dittmann, 2013, Doszyń, 2023), housing availability (Matel, Marcinkiewicz, 2017; Czerniak et al., 2022; Zakrzewska-Półtorak, Pluta, 2023).

There have also been published papers assessing various aspects of the quality of data used for property market analyses (Nawrocka, 2013; Konowalczyk, 2014; Tomczyk, Widłak, 2010). The conclusions drawn from these are rather pessimistic. In the course of the literature studies, no research involving a direct comparison of property price data from the two sources considered here was encountered. In this context, the study should be regarded as innovative, as it indicates that public datasets of market information should not be used for further analysis and research indiscriminately and that it is legitimate to subject them to evaluation.

The aim of the research undertaken is to assess the datasets on the average transaction prices of dwellings published by the Central Statistical Office and the National Bank of Poland in terms of their compatibility and usability for market analyses. The research hypothesis stated: data on average housing transaction prices published by the CSO and the NBP do not differ significantly. The hypothesis formulated in this way reflects the desired outcome of the study, since the recipients of the data expect that both sources under consideration are reliable.

## 1. Material and methods

The study used data:

1. On average quarterly prices per m<sup>2</sup> of residential units sold in market transactions published by the CSO. These data are compiled on the basis of information obtained from the Property Price Register (until 31 July 2021 of the Property Price and Value Register).
2. About the average quarterly transaction prices of 1 m<sup>2</sup> derived from the National Bank of Poland's residential property price database, which is created through the submission of data by real estate agents and developers with the involvement of the district branches of the NBP.

In both cases, quarterly information from 2010 to 2022 was obtained from the secondary and primary markets. Prices were expressed in PLN. The research period is determined by the overlap of the periods of publication of data by both institutions.

The following analytical work was carried out as part of the study:

1. The average prices of 1 m<sup>2</sup> flats published by the CSO and the NBP in the quarters for each city – separately for the secondary and primary market – are presented by means of line graphs.
2. A t-test was conducted for average prices published by the CSO and the National Bank of Poland for each city.
3. Differences between the average prices of 1 m<sup>2</sup> of a dwelling published by the CSO and the NBP for individual cities ( $R_{ti}$ ) were calculated separately for the secondary and primary market using the formula:

$$R_{ti} = C_{ti}^{CSO} - C_{ti}^{NBP}$$

where:

$C_{ti}^{CSO}$  – average price in period  $t$  in the  $i$ -th city published by the CSO,

$C_{ti}^{NBP}$  – average price in period  $t$  in the  $i$ -th city published by the NBP.

4. The differences calculated in point 2 – separately for the secondary and primary markets – are presented by means of bar charts.
5. Relative differences between average prices of 1 m<sup>2</sup> flats published by CSO and NBP ( $R_{ti}^{rel}$ ) for individual cities separately for the secondary and primary market by adopting the formula:

$$R_{ti}^{rel} = \frac{C_{ti}^{CSO} - C_{ti}^{NBP}}{C_{ti}^{CSO}} \times 100\%$$

6. The relative differences calculated in point 4 for each city are presented using line graphs – common graphs for secondary and primary markets.
7. Arithmetic means, standard deviations and ranges of price differences of 1 m<sup>2</sup> flats published by the NBP and the CSO were calculated for individual cities – separately for the secondary and primary market.
8. The box-and-whisker diagrams of the measures calculated in point 6 for each city are presented.
9. Arithmetic averages of positive and negative differences were calculated for each city.
10. The numbers of positive and negative differences, the numbers of positive and negative differences exceeding  $\pm 5\%$  and exceeding  $\pm 10\%$  were counted.
11. The results obtained in points 8 and 9 are presented in tables – separately for the secondary and primary markets.

Due to the impossibility of objectively indicating acceptable levels of differences, when discussing the results, attention was paid to absolute differences exceeding 500 and 1,000 PLN/m<sup>2</sup> and relative differences exceeding 5 and 10%. Differences between the average prices of 1 m<sup>2</sup> flats, which are the subject of the analyses, should in principle be as small as possible. The classification into positive and negative differences used is of a conventional nature, as it results from assumptions as to which prices have been adopted as negative and which as positive. If the opposite assumption were made in this respect, numerically opposite results would be obtained. However, this would not change the essence of the study, the interpretation of the results and the evaluation of the phenomena under investigation.

## 2. Results and discussion

The graphs presented in Appendix A show the average prices of 1 m<sup>2</sup> flat in individual quarters of 2010–2022 on the secondary market for the cities under consideration published by the Central Statistical Office and the National Bank of Poland. We conclude from them that the data published by these institutions, although reflecting similar long-term trends, relatively often do not coincide, and in certain periods even diverge. Analogous data for the primary market are presented in the charts in Appendix B. Here, the degree of non-overlap between the average prices of 1 m<sup>2</sup> flats published by the CSO and the NBP is even greater. Table 1 contains the results of the t-test for average prices published by the CSO and the NBP for each city in the secondary and primary markets. For the secondary market, at a significance level of  $\alpha = 0.05$ , the average difference turns out to be significant for 1 the city of Rzeszow. However, for the

primary market, such significant differences are already observed for three cities – Katowice, Szczecin and Rzeszów.

Table 1. Results of the t-test for average prices published by the Central Statistical Office and the National Bank of Poland for individual cities on the secondary and primary market

| City         | Secondary market |             |         |        | Primary market |             |         |        |
|--------------|------------------|-------------|---------|--------|----------------|-------------|---------|--------|
|              | average CSO      | average NBP | t       | p      | average CSO    | average NBP | t       | p      |
| Białystok    | 4,823.81         | 4,689.02    | 0.6093  | 0.5437 | 4,954.15       | 5,255.73    | -1.7100 | 0.0903 |
| Bydgoszcz    | 4,166.71         | 4,302.47    | -0.6912 | 0.4910 | 5,116.85       | 5,486.52    | -1.8575 | 0.0661 |
| Gdańsk       | 6,678.27         | 6,696.42    | -0.0470 | 0.9626 | 6,813.31       | 7,178.59    | -1.1487 | 0.2534 |
| Katowice     | 3,978.50         | 4,036.33    | -0.3030 | 0.7625 | 4,740.18       | 5,696.04    | -3.9377 | 0.0002 |
| Kielce       | 4,374.58         | 4,247.80    | 0.7892  | 0.4319 | 5,076.73       | 5,157.49    | -0.5089 | 0.6119 |
| Kraków       | 7,040.58         | 7,026.33    | 0.0464  | 0.9631 | 7,258.67       | 7,362.58    | -0.3836 | 0.7021 |
| Lublin       | 5,274.67         | 5,312.61    | -0.1737 | 0.8624 | 5,424.37       | 5,613.96    | -0.9386 | 0.3501 |
| Łódź         | 4,171.06         | 4,235.18    | -0.3185 | 0.7507 | 5,218.10       | 5,429.89    | -1.1300 | 0.2611 |
| Olsztyn      | 4,719.71         | 4,773.55    | -0.2988 | 0.7657 | 5,232.63       | 5,374.79    | -0.8055 | 0.4224 |
| Opole        | 4,384.27         | 4,488.28    | -0.5899 | 0.5566 | 4,794.54       | 5,100.22    | -1.7307 | 0.0865 |
| Poznań       | 5,766.29         | 5,920.52    | -0.6880 | 0.4930 | 6,511.27       | 6,806.21    | -1.6972 | 0.0927 |
| Rzeszów      | 4,823.44         | 5,284.48    | -2.0606 | 0.0420 | 4,889.06       | 5,256.57    | -2.0802 | 0.0401 |
| Szczecin     | 4,593.79         | 4,772.63    | -0.7868 | 0.4332 | 5,179.82       | 5,733.97    | -2.1958 | 0.0305 |
| Warszawa     | 9,002.35         | 8,682.42    | 0.9215  | 0.3589 | 8,634.38       | 8,497.50    | 0.4939  | 0.6224 |
| Wrocław      | 5,848.33         | 6,326.52    | -1.6150 | 0.1094 | 6,477.62       | 6,822.23    | -1.4514 | 0.1497 |
| Zielona Góra | 3,767.35         | 3,758.22    | 0.0484  | 0.9615 | 3,991.20       | 4,339.61    | -1.8594 | 0.0659 |

Source: own study.

The results indicate that differences in prices can be significant, and this does happen for some cities. However, from a practical point of view, what matters is not only that the average prices for individual cities published by different institutions differ, but that the individual prices recorded for the periods in question differ. This is because it may turn out that in a given city the t-test does not show the significance of differences, but such differences appear incidentally in some periods. Therefore, in the following section, individual differences were calculated and analyzed for each city and period.

Appendixes C and D are presented by means of bar charts the differences between the average quarterly prices of 1 m<sup>2</sup> flats published by CSO and NBP on the secondary and primary markets, respectively. Appendix E, on the other hand, presents the relative differences between

these prices on the secondary and primary markets. We can see that in different cities and in different periods these differences are very irregular both in terms of direction and value. As these differences are analysed in detail later in the research, a detailed discussion of the graphs presented in these appendices is dispensed with at this point, resting instead on the formulation of the above general conclusion. This conclusion is at the same time the reason for developing the research.

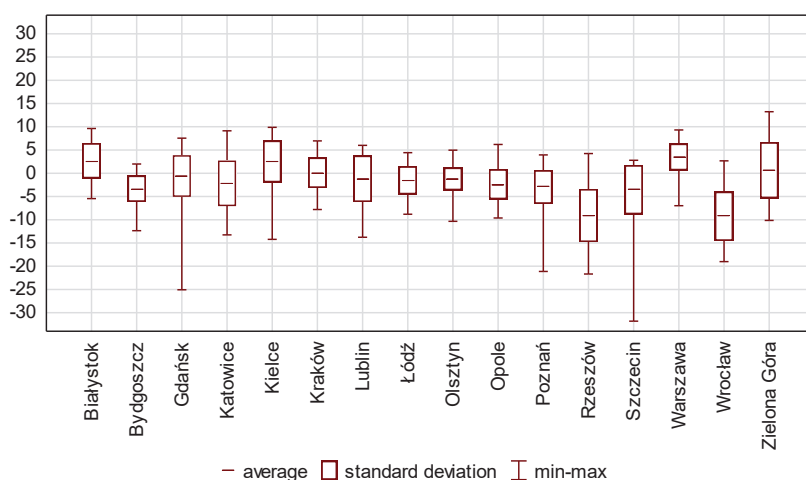


Figure 1. Averages, standard deviations and ranges of relative price differences of 1 m<sup>2</sup> dwellings published by the NBP and the CSO on the secondary market (%)

Source: own study.

Figure 1 shows the arithmetic means, standard deviations and ranges (max-min) of percentage differences between average prices published by the CSO and the NBP in individual cities calculated from quarterly data for the period from 2010 to 2022 for the secondary market. Overall, the average percentage differences appear to be at a reasonable level – for 14 cities they do not exceed the –5 to 5% range, and for none of the cities do they exceed the –10 to +10% range. For two cities – Rzeszów and Wrocław – they approach –10%, which can also be considered as not glaring. Negative differences prevail, which, given the formula adopted for their calculation, means that usually the average prices of 1 m<sup>2</sup> published by the NBP were higher than those published by the CSO (11 times). The levels of standard deviations should be considered as definitely less reasonable, as for the majority, because for 13 cities, the percentage of  $\bar{x} + s$  or  $\bar{x} - s$  was respectively greater than 5 or less than –5%, which means that these differences are characterised by high variability and that for a significant

number of observations the differences between the published values were greater than 5%. The analysis of extreme values though leads to worrying conclusions. It turns out that there are observations for which percentage differences in the average prices of 1 m<sup>2</sup> flat published by two state institutions exceed 20%. This happened for data concerning Gdańsk, Poznań, Rzeszów and Szczecin. These are already very large differences and, despite the fact that they occur relatively rarely, one should ask the question about their reasons. This is because they do not seem to be caused solely by differences in the way in which measurements are taken.

Table 2 summarises the results accordingly, broken down into positive and negative deviations. The number of observations for individual cities was, as a standard, 52 (13 years of 4 quarters each), but due to occasional gaps in the data, there were sometimes fewer observations. The main conclusion that can be drawn from the analysis of the presented data is that it is difficult to find unambiguous regularities here. There were positive and negative deviations in all of the cities. There are cities with a clear advantage of positive deviations (Kielce, Warsaw), but there are also cities with a clear advantage of negative deviations (Białystok, Bydgoszcz, Opole, Poznań, Rzeszów, Szczecin, and Wrocław). There are no cities for which deviations exceeding 5% (positive or negative) were not observed, although there are some in which such deviations were relatively rare – less than 10 were observed for Bydgoszcz, Gdańsk, Krakow, Łódź, and Olsztyn. There are also cities where deviations exceeding 5% were extremely frequent – Rzeszów, and Wrocław, where they occurred more than 40 times and were always negative. Obviously, deviations exceeding 10% occurred less frequently – for 5 cities they did not occur at all, for another 5 they occurred once (3 positive, 2 negative). One cannot fail to notice, however, that for Rzeszów and Wrocław the number of such deviations occurred for almost half of the observations.

Table 2. Averages, numbers and numbers above 5% and above 10% positive and negative percentage differences for each city in the secondary market

| City      | Mean deviation + | Mean deviation – | Number of deviations + | Number of deviations – | Number of deviations + over 5% | Number of deviations – over 5% | Number of deviations + over 10% | Number of deviations – over 10% |
|-----------|------------------|------------------|------------------------|------------------------|--------------------------------|--------------------------------|---------------------------------|---------------------------------|
| 1         | 2                | 3                | 4                      | 5                      | 6                              | 7                              | 8                               | 9                               |
| Białystok | 4.26             | –2.57            | 12                     | 40                     | 2                              | 13                             | 0                               | 0                               |
| Bydgoszcz | 1.15             | –3.70            | 4                      | 48                     | 9                              | 0                              | 1                               | 0                               |
| Gdańsk    | 2.46             | –2.84            | 22                     | 30                     | 2                              | 3                              | 1                               | 0                               |
| Katowice  | 2.83             | –5.13            | 20                     | 32                     | 2                              | 15                             | 0                               | 4                               |
| Kielce    | 4.55             | –3.50            | 39                     | 13                     | 15                             | 3                              | 1                               | 0                               |

| 1            | 2    | 3      | 4  | 5  | 6  | 7  | 8 | 9  |
|--------------|------|--------|----|----|----|----|---|----|
| Kraków       | 2.97 | -2.33  | 24 | 28 | 4  | 3  | 0 | 0  |
| Lublin       | 2.73 | -4.98  | 26 | 25 | 15 | 13 | 3 | 0  |
| Łódź         | 1.92 | -3.00  | 15 | 36 | 0  | 6  | 0 | 0  |
| Olsztyn      | 1.07 | -2.43  | 18 | 34 | 0  | 2  | 0 | 1  |
| Opole        | 1.75 | -3.47  | 11 | 41 | 1  | 9  | 0 | 0  |
| Poznań       | 1.22 | -3.60  | 7  | 45 | 0  | 10 | 0 | 1  |
| Rzeszów      | 2.75 | -10.25 | 3  | 47 | 0  | 43 | 0 | 23 |
| Szczecin     | 1.64 | -4.21  | 6  | 46 | 0  | 12 | 0 | 3  |
| Warszawa     | 4.02 | -5.03  | 49 | 3  | 15 | 1  | 0 | 0  |
| Wrocław      | 2.67 | -9.48  | 1  | 51 | 0  | 40 | 0 | 25 |
| Zielona Góra | 5.88 | -4.33  | 25 | 26 | 14 | 11 | 4 | 1  |

Source: own study.

Figure 2 shows the averages, standard deviations and ranges of relative price differences of 1 m<sup>2</sup> flats published by the National Bank of Poland and the Central Statistical Office on the primary market in relative terms. It turns out that there are cities for which even the average differences exceed or approach 10%. Typical values of the modules  $\bar{x} + s$  or  $\bar{x} - s$  for as many as 14 cities exceed 10%, and in three cases even 20%. Very alarming are the levels of extreme values of price differences, which for Katowice and Szczecin exceed 70%, for Białystok – 60%, and for Bydgoszcz, Warsaw and Zielona Góra – 30%. It should be noted that in the case of the primary market, negative differences prevail, which means that prices published by the CSO are generally lower than those published by the NBP. Such high differences are certainly not solely the result of methodological differences. They point to the legitimacy of undertaking work to identify their causes. Annex No. D presents in detail the differences between the average prices of 1 m<sup>2</sup> flats published by the CSO and the NBP on the primary market in the examined cities in particular quarters of the years 2010–2022. From the charts presented there, it may be concluded that for the majority of cities, particularly large and usually negative differences were observed in the last several – more than ten quarters of the analysed period. Perhaps the reason for this is the fact that the contract for the purchase of a dwelling on the primary market is concluded in two stages. In the first stage, the developer undertakes to build a dwelling and sell it to the purchaser for a specified price, and in the second stage, often several months later, the constructed dwelling is actually sold. This mechanism leads to contracts for the sale of flats at prices agreed in the past. Moreover, as is well known, the primary market is a specific market, as transactions on it take place only when development investments are undertaken by specialised companies, and these investments can take various forms – from budget,

low-cost flats located on the outskirts of cities to prestigious premium class investments in top locations. But does this explain such large differences in primary market housing prices published by various institutions?

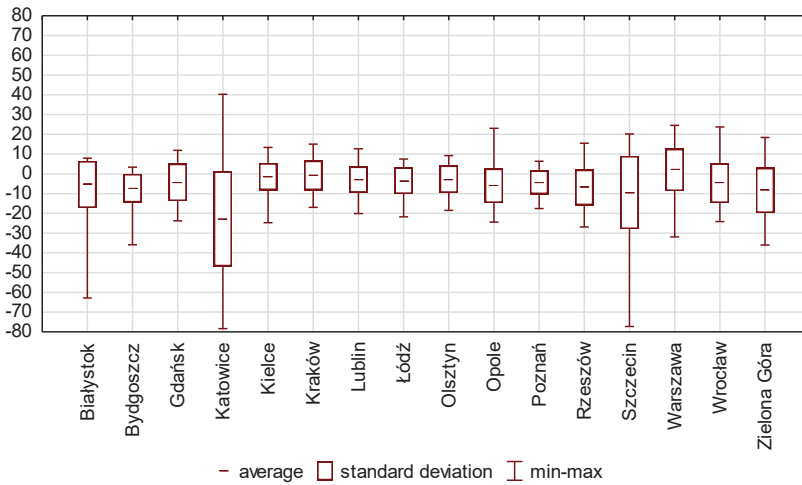


Figure 2. Averages, standard deviations and ranges of relative price differences of 1 m<sup>2</sup> dwellings published by the NBP and the CSO on the primary market (%)

Source: own study.

To better illustrate the magnitude of the occurrence and intensity of the differences in question, Table 3 shows the recorded averages, numbers, numbers above 5 and above 10% positive and negative percentage deviations for each city in the primary market.

Table 3. Averages, numbers and numbers above 5 and above 10% positive and negative percentage differences for individual cities on the primary market

| City      | Mean deviation + | Mean deviation - | Number of deviations + | Number of deviations - | Number of deviations + over 5% | Number of deviations - over 5% | Number of deviations + over 10% | Number of deviations - over 10% |
|-----------|------------------|------------------|------------------------|------------------------|--------------------------------|--------------------------------|---------------------------------|---------------------------------|
| 1         | 2                | 3                | 4                      | 5                      | 6                              | 7                              | 8                               | 9                               |
| Białystok | 2.90             | -8.75            | 15                     | 37                     | 3                              | 19                             | 0                               | 8                               |
| Bydgoszcz | 1.27             | -8.58            | 7                      | 45                     | 0                              | 29                             | 0                               | 15                              |
| Gdańsk    | 5.20             | -10.26           | 20                     | 32                     | 9                              | 22                             | 3                               | 15                              |
| Katowice  | 22.29            | -29.97           | 6                      | 44                     | 6                              | 42                             | 5                               | 39                              |
| Kielce    | 4.31             | -5.33            | 21                     | 31                     | 7                              | 11                             | 3                               | 4                               |

| 1            | 2    | 3      | 4  | 5  | 6  | 7  | 8 | 9  |
|--------------|------|--------|----|----|----|----|---|----|
| Kraków       | 6.67 | -5.86  | 21 | 31 | 13 | 15 | 2 | 6  |
| Lublin       | 3.16 | -5.86  | 17 | 35 | 2  | 14 | 1 | 6  |
| Łódź         | 4.74 | -6.18  | 13 | 39 | 7  | 17 | 0 | 9  |
| Olsztyn      | 3.53 | -7.33  | 22 | 30 | 6  | 20 | 0 | 9  |
| Opole        | 6.85 | -8.72  | 9  | 43 | 5  | 28 | 2 | 17 |
| Poznań       | 3.51 | -6.37  | 11 | 41 | 4  | 21 | 0 | 11 |
| Rzeszów      | 5.70 | -9.95  | 9  | 41 | 4  | 28 | 1 | 17 |
| Szczecin     | 3.50 | -18.10 | 20 | 30 | 4  | 20 | 1 | 16 |
| Warszawa     | 7.72 | -8.91  | 34 | 18 | 17 | 11 | 9 | 6  |
| Wrocław      | 5.18 | -10.43 | 19 | 33 | 8  | 24 | 1 | 15 |
| Zielona Góra | 6.74 | -13.07 | 12 | 39 | 7  | 34 | 3 | 22 |

Source: own study.

The data in the Table above shows that the average positive and negative deviations for the primary market are significantly higher than for the secondary market. There are also significantly more significant deviations – above 5% and above 10%. As in the case of the secondary market, it is also difficult to find clear patterns here. In the case of differences calculated for the primary market, negative differences prevail in all cities except Warsaw. The number of negative differences exceeding 10% is astonishing – there is no city for which such an observation would not be recorded, and in the case of Katowice it is as high as 39 out of 50 observations. For the other cities, the numbers are also usually significant.

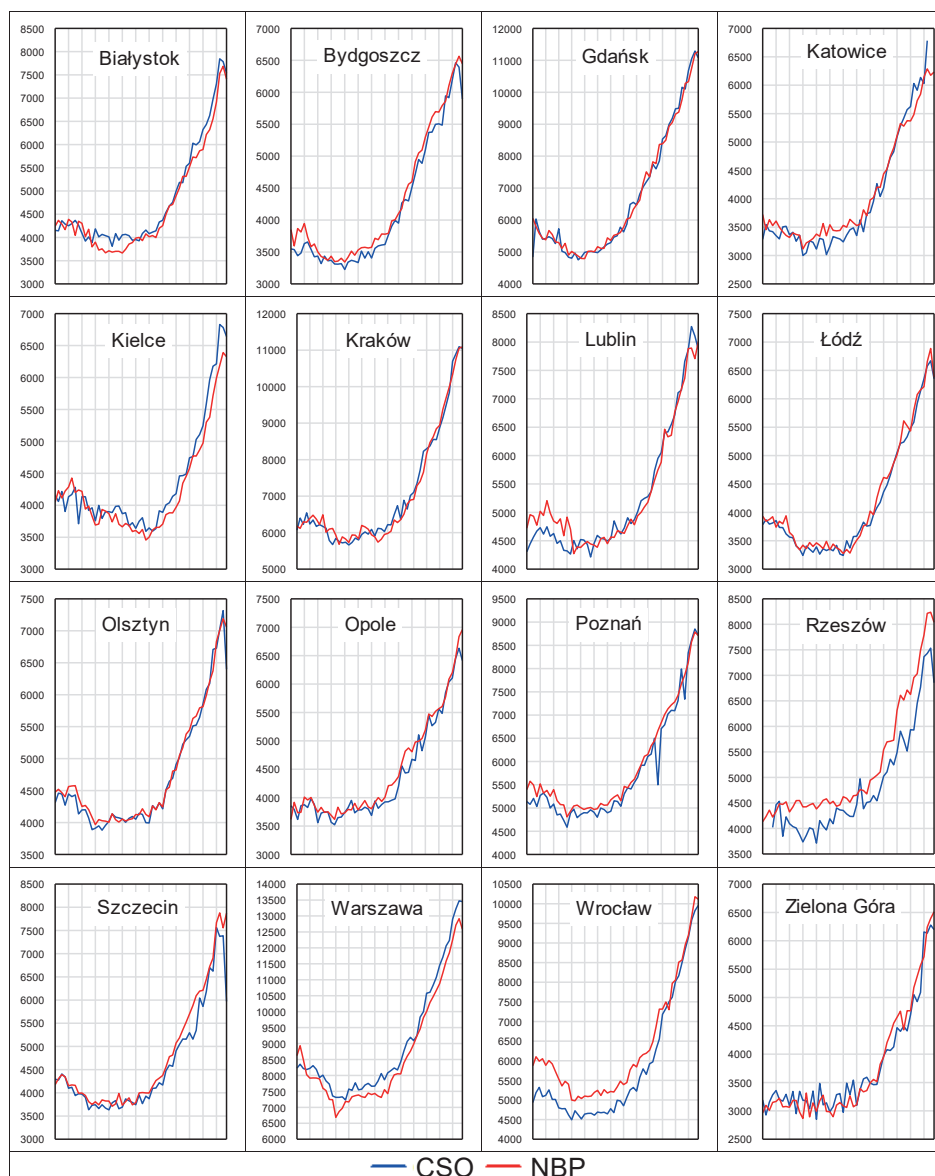
## Conclusions

The results obtained concerning the differences between average prices of 1m<sup>2</sup> for 16 cities in Poland published by two state institutions – the Central Statistical Office and the National Bank of Poland – lead to disturbing conclusions. Indeed, these differences are very large. This is especially the case for the primary market, where the t-test has already shown the significance of differences for three cities. Moreover, it was found that the differences between the published average prices in a significant number of observations exceed 10%, but there are also differences of 30, 40 and even more than 70%. The observed differences are large and irregular. It is difficult to discern any regularity in them concerning, for example, higher or lower average prices of flats published by one institution in relation to data published by another, a characteristic level of differences, or groups of cities for which such regularities could be identified. Therefore, it cannot be concluded that the average prices of flats published

by the CSO or the NBP are subject to a specific systematic error. Awareness of the existence of such a systematic error in published data would certainly influence the attitude of the users to the analysed phenomena. Such large and irregular differences in the prices of flats published by the NBP and the CSO cause a huge threat to the recipients of these data, as certainly at least one of the indicated sources presents data distorting the reality. By making investment decisions based on information that does not correspond to reality, there is a significant risk that the investor will make a wrong decision. Inaccurate data also puts analysts who base their analyses on it and publish their results in a difficult situation, as they can thus provide the public with a distorted picture of the situation on the housing market. Distorted data may also lead to erroneous conclusions from scientific research conducted by various research centres using data behind which stands the authority of state, and public institutions. Such centres usually do not have the financial and technical- organisational capacity to obtain and process information on average prices in various markets on their own. On the other hand, these institutions have an adequate budget and teams of fulltime employees, as well as equipment and software. In addition, they benefit from the special powers given to them by the Law on Public Statistics. Therefore, the public has a right to expect that the data published by these institutions are reliable and credible. Different methods of obtaining and processing input data, i.e. data on prices quoted in individual transactions, may lead to differences in the results obtained and therefore differences in the published average prices. However, differences of more than a few, a dozen or even several tens of per cent seem unacceptable. Thus, the research carried out leads to the conclusion that there is no basis for the claim that the data on average housing transaction prices published by the CSO and the NBP do not differ significantly, which means that the research hypothesis has been negated. Work should therefore urgently be undertaken to clarify the reasons for such differences. Unfortunately, this is not possible from the position of the data recipient. Both the CSO and the NBP should subject the whole process of creating information on average prices of flats in particular local markets to reliable analysis, starting from the evaluation of methods of obtaining input data, evaluation of their quality in terms of completeness and reliability, and evaluation of the processes of processing and analysis of verified data. Future research should focus primarily on monitoring the differences in published data. In addition, the right direction seems to be to assess the impact of the differences in published average prices on other characteristic parameters of housing markets such as economic indicators and rates of return, as well as on the buying and selling decisions of market participants.

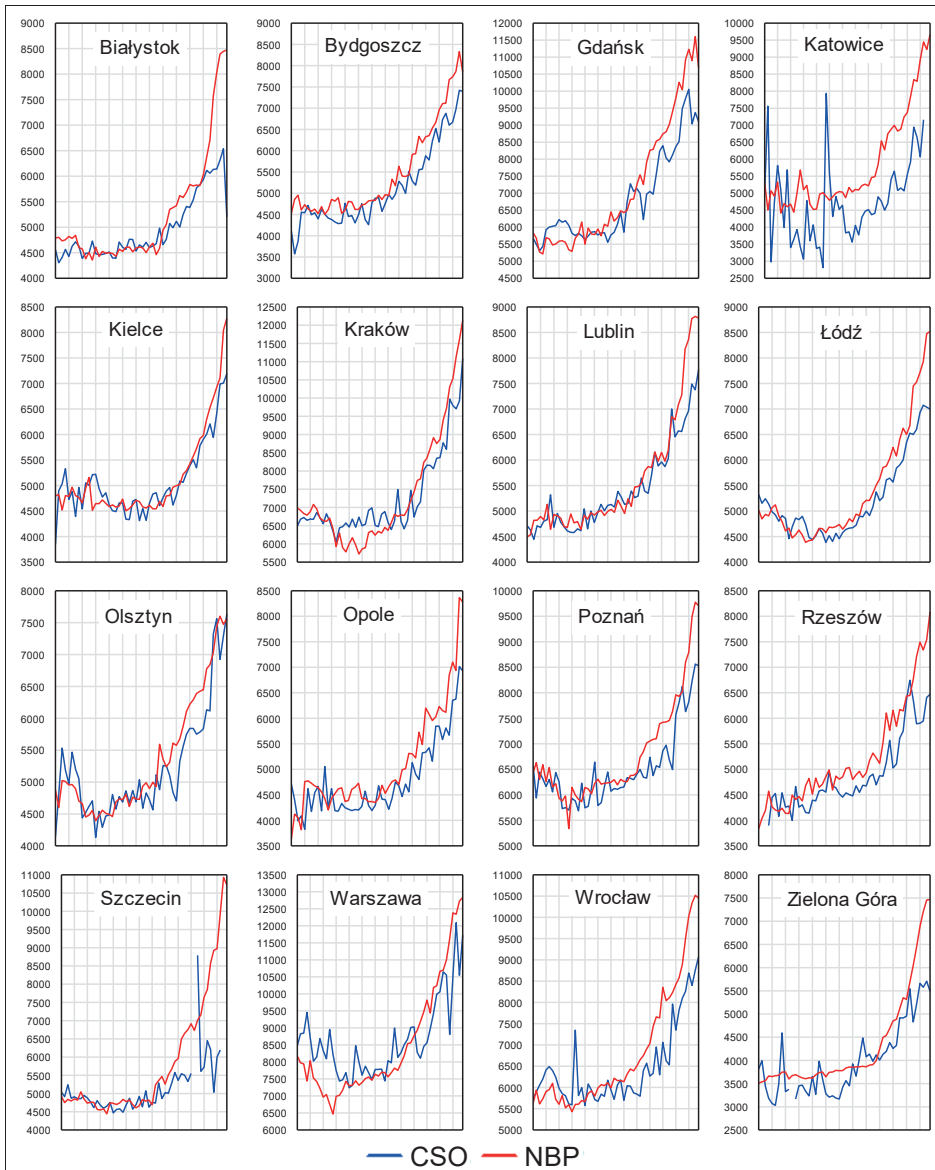
## Appendix

### A. Average prices of 1 m<sup>2</sup> flats on the secondary market in cities according to the Central Statistical Office and the National Bank of Poland between 2010 and 2022 (PLN)



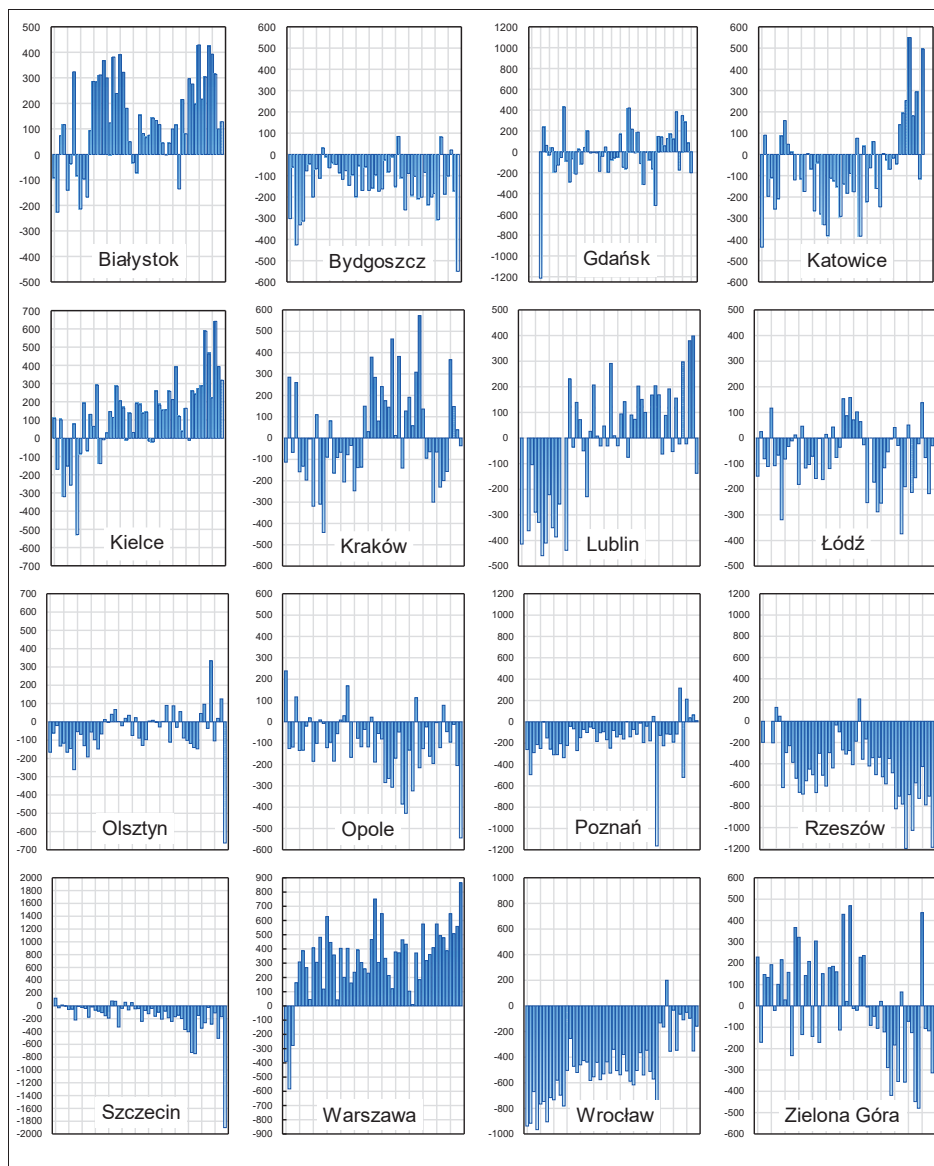
Source: own elaboration based on data published by CSO and NBP.

B. Average prices of 1 m<sup>2</sup> flats on the primary market in cities according to the Central Statistical Office and the National Bank of Poland in 2010–2022 (PLN)



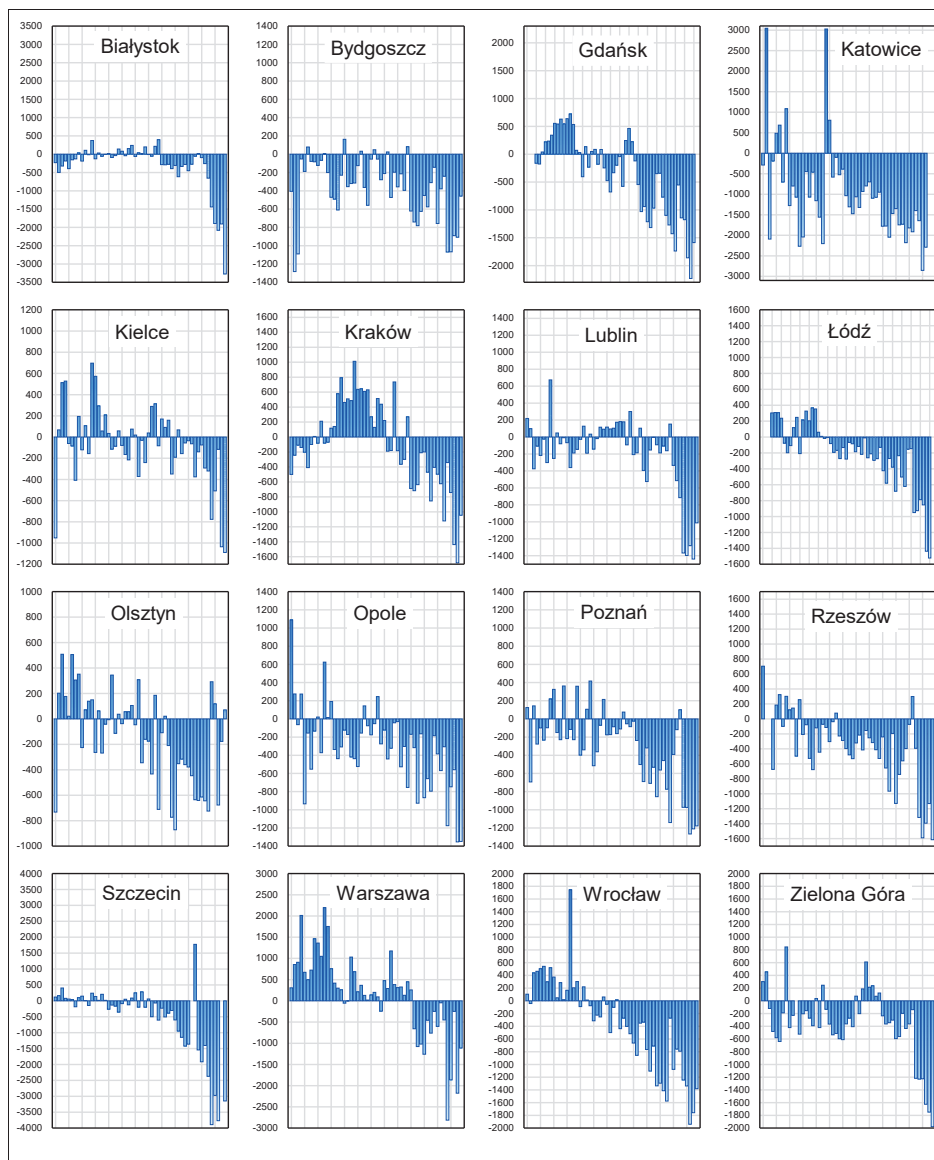
Source: own elaboration based on data published by CSO and NBP

C. Differences between the average prices of 1 m<sup>2</sup> of a flat published by the Central Statistical Office and the National Bank of Poland on the secondary market in cities between 2010 and 2022 (PLN)



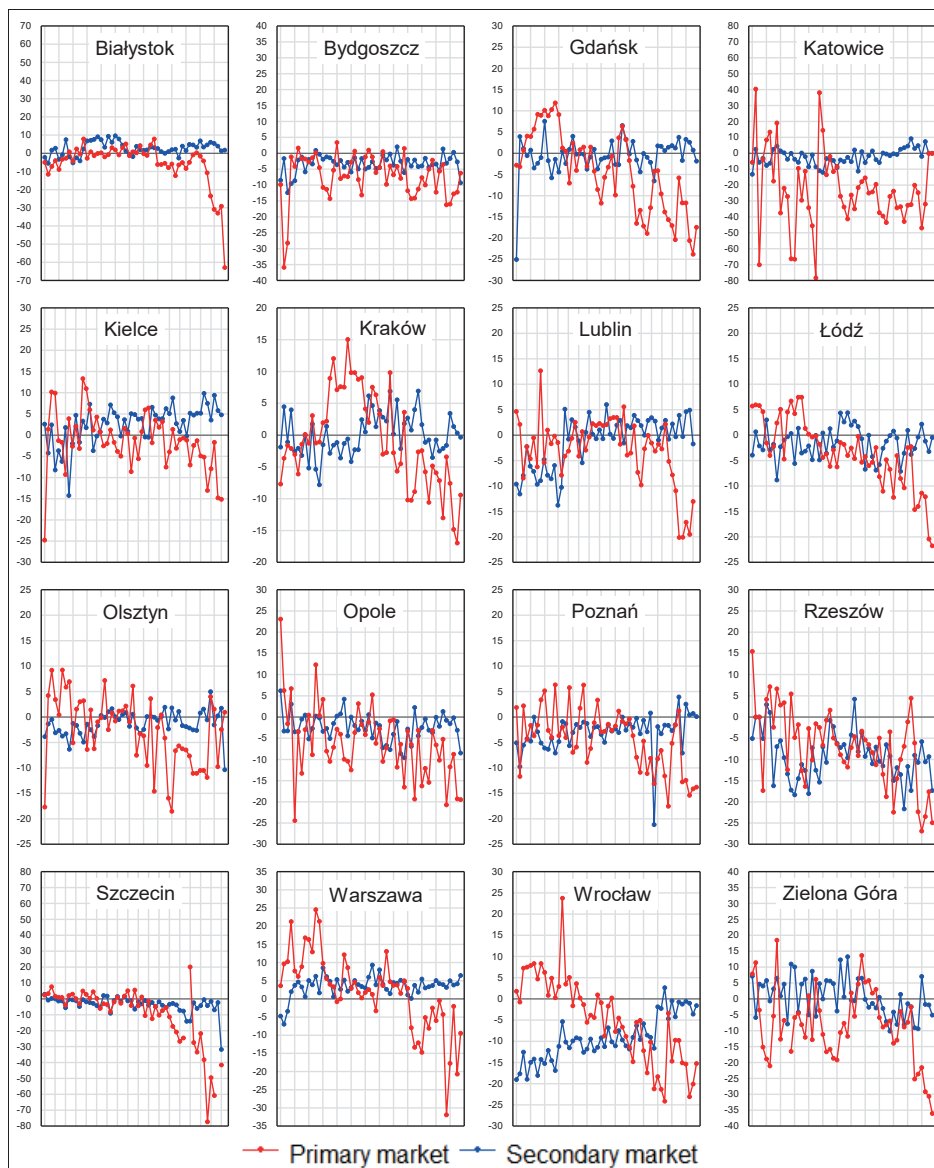
Source: own elaboration based on data published by CSO and NBP.

D. Differences between average prices of 1 m<sup>2</sup> of a flat published by the Central Statistical Office and the National Bank of Poland on the primary market in cities between 2010 and 2022 (PLN)



Source: own elaboration based on data published by CSO and NBP.

E. Relative differences between the average prices of 1 m<sup>2</sup> of a dwelling published by the CSO and the NBP on the primary and secondary markets in cities between 2010 and 2022 (%)



Source: own elaboration based on data published by CSO and NBP.

---

## References

---

- Belej, M. (2023). Predicting Housing Price Trends in Poland: Online Social Engagement – Google Trends. *Real Estate Management and Valuation*, 31(4), 73–87. DOI: 10.2478/remav-2023-0032.
- Bryx, M. (2001). *Finansowanie inwestycji mieszkaniowych*. Warsaw, Poland: Poltex.
- Bryx, M., Matkowski, R. (2001). *Inwestycje w nieruchomości*. Warsaw, Poland: Poltex.
- Czerniak, A., Czaplicki, M., Mokrogulski, M., Niedziółka, P. (2022). *Raport SGH i Forum Ekonomicznego 2022*. Dostępność finansowa mieszkań w krajach Europy Środkowo-Wschodniej na tle zmian parametrów polityki pieniężnej (pp. 85–108), Warsaw, Poland: Forum Ekonomiczne, Szkoła Główna Handlowa. DOI: 10.33119/978-83-8030-563-2.2022.3.85.108.
- Dittmann, I. (2014). Gamma konwergencja cen na lokalnych rynkach mieszkaniowych. *Studia Ekonomiczne*, 181, 195–207.
- Doszyń, M. (2023). Integration and cointegration of apartment prices on the primary and secondary market in Szczecin in the years 2006–2022. *Real Estate Management and Valuation*, 31(4), 36–44. DOI: 10.2478/remav-2023-0028.
- Ferry, D., Brandon, P. (1986). *Cost planning of buildings*. London, UK: Collins.
- Francke, M.K. (2010). Repeat Sales Index for Thin Markets. *The Journal of Real Estate Finance and Economics*, 41(1), 24–52.
- Kokot, S. (2015). Przyczynek do badań nad indeksami cen nieruchomości. *Wiadomości Statystyczne*, 647(4), 1–12. DOI: 10.5604/01.3001.0016.0850.
- Kokot, S. (2020). Socio-economic factors as a criterion for the classification of housing markets in selected cities in Poland. *Real Estate Management and Valuation*, 28(3), 77–90. DOI: 10.1515/remav-2020-0025.
- Kokot, S. (2022). Identification of regularities in relation between prices on primary and secondary housing market in selected cities in Poland. *Real Estate Management and Valuation*, 30(3), 45–60. DOI: 10.2478/remav-2022-0020.
- Kokot, S. (2023). The Effect of Price Anchoring on the Housing Market Based on Studies of Local Markets in Poland. *Real Estate Management and Valuation*, 31(3), 44–57. DOI: 10.2478/remav-2023-0020.
- Konowalczyk, J. (2014). The corporate real estate market in public statistics in Poland. *Real Estate Management and Valuation* 22(2), 41–51. DOI: 10.2478/remav-2014-0016.
- Konowalczyk, J., Ramian, T. (2012). Potrzeby i bariery informacyjne rynku nieruchomości z perspektywy wyceny nieruchomości w Polsce. *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 20(2), 105–118.

- Matel, A., Marcinkiewicz, J. (2017). Analiza dostępności mieszkań w miastach wojewódzkich Polski z wykorzystaniem metody topsis affordability. *Prace naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 475, 170–182. DOI: 10.15611/pn.2017.475.15.
- Matysiak, G.A., Olszewski, K., (2019). *NBP Working Paper*, 316. A panel analysis of Polish regional cities: residential price convergence in the primary market, Warsaw, Poland.
- Melnychenko, O., Osadcha, T., Kovalyov, A., Matskul, V. (2022). Dependence of Housing Real Estate Prices on Inflation as One of the Most Important Factors: Poland's Case. *Real Estate Management and Valuation*, 30(4), 25–41. DOI: 10.2478/remav-2022-0027.
- Narodowy Bank Polski, Departament Stabilności Finansowej (2022). *Raport o sytuacji na rynku nieruchomości mieszkaniowych i komercyjnych w Polsce w 2021 r.* Warsaw, Poland.
- Nawrocka, E. (2013). Analiza i ocena jakości krajowych zasobów informacji o nieruchomościach w aspekcie ich wykorzystania do opracowania indeksu cen. *Studia Ekonomiczne*, 155, 82–96.
- Tomal, M. (2019). House Price Convergence on the Primary and Secondary Markets: Evidence from Polish Provincial Capitals. *Real Estate Management and Valuation*, 27(4), 62–73. DOI: 10.2478/remav-2019-0036.
- Tomczyk, E., Widłak, M. (2010). Konstrukcja i własności hedonicznego indeksu cen mieszkań dla Warszawy. *Bank i Kredyt*, 41(1).
- Widłak, M., Nehrebecka, N. (2011). Wykorzystanie regresji kwantylowej w analizie zróżnicowania cen mieszkań. *Wiadomości Statystyczne. The Polish Statistician*, 600(5), 17–46.
- Zakrzewska-Półtorak, A., Pluta, A. (2023). Zmiany w zakresie atrakcyjności nieruchomości mieszkaniowych w polskich miastach w latach 2010–2020, ze szczególnym uwzględnieniem dostępności. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 67(3), 164–176. DOI: 10.15611/pn.2023.3.15.

## Citation

---

- Kokot, S. (2025). Comparative Evaluation of Public Data Sets on Average House Prices in Poland. *Folia Oeconomica Stetinensia*, 25(1), 157–179. DOI: 10.2478/fofi-2025-0008.