The Polish experience in developing House Price Index

Renata Rechnio, Central Statistical Office of Poland

Abstract

The CSO of Poland started publishing the official House Price Index for Poland in 2015. Compilations of the HPI for Poland are based on stratification method and are carried out with the use of administrative sources. This paper presents the experience of the CSO of Poland in compiling and developing the HPI. The paper provides information on the used data sources and applied methods as well as discusses obtained results. Attention is given to problems connected with the use of administrative data. In particular, issues on cooperation with data holding authorities, completeness and quality of administrative sources and timeliness of obtained results are discussed. The paper also indicates areas in which further improvements can be made.

1. Introduction

The significant role of the housing market in the economy means that the important task of statistical authorities is developing reliable, good quality data on price changes of residential properties. However, the measurement of residential properties prices is challenging in practice. The literature on the subject indicates that the main problems with the construction of the HPI are related to the heterogeneity of the residential market and relative infrequency of transactions. At the same time practical experience shows that the crucial issue in the compilation of the HPI is access to the data sources containing information on transaction prices and detailed characteristics of sold properties.

The CSO of Poland started regular production of the HPI in 2012 and then the publication of the official HPI for Poland was launched in October 2015. Earlier attempts to develop price indices for residential properties were unsuccessful due to difficulties in obtaining data on real estate transactions. Development in this area was possible after implementation of electronic databases with real estate prices by administration authorities.

The purpose of the paper is to present the experience of the CSO of Poland in developing the HPI. In particular, issues connected with the use of administrative data are discussed. Moreover, the used methodology and obtained results are presented.

The paper is organized as follows: the second chapter describes experiences of the CSO of Poland with the use of administrative data for the compilation of the HPI. There is information on challenges connected with the use of this kind of data as well as methods of overcoming them. The third chapter presents information on the methodology used in the production of the HPI for Poland. Then, obtained results are presented in the fourth chapter. The fifth chapter indicates directions for further development.

2. Data sources

Calculations of the HPI for Poland are based on data from the administrative data source called the Register of Real Estates Prices and Values. The register is maintained in electronic form by 380 district governors offices (local administration units). The legal basis for the Register of Real Estate Prices and Values is the Act of May 17, 1989 The Geodetic and Cartographic Law (Journal of Laws on 2010, No. 193, entry 1287 with further amendments). The scope of the register is defined by the Minister of Regional Development and Housing Decree of March 29, 2001 on registration of land and buildings (Journal of Laws on 2001, no. 38, entry 454 with further amendments).

When deciding on the data source used for the HPI calculations, besides administrative data, the collection of data via a statistical survey was also considered. The major advantage of using

data from the Register of Real Estates Prices and Values compared to survey data is that it allows to reduce costs of data collections and eliminate the administrative burden on potential respondents. Another advantage of the Register of Real Estates Prices and Values is that it gathers data on transaction prices established in notarial deed. It is intended to include information on all transactions concerning complete property rights to real estate which constitute the majority of market transactions. Moreover the important attribute of the register is that in addition to price information it also gathers information on basic characteristics of properties which are relevant for the HPI calculations. In particular, it includes the following characteristics:

- symbol of voivodship (NUTS 2 level),
- symbol of powiat (LAU 1 level),
- symbol of gmina (LAU 2 level),
- type of gmina (urban, rural),
- transaction date,
- selling party (State Treasury, gmina, powiat, voivodship, natural person, legal person),
- purchasing party (State Treasury, gmina, powiat, voivodship, natural person, legal person),
- sale form (sale on free market, sale under tender procedure, sale without tender procedure),
- number of rooms (only for flats),
- storey (only for flats),
- price of transaction,
- price of property,
- square meters of floor space.

Despite unquestionable advantages of using administrative data for the HPI, the specificity of this kind of data causes that some limitations and challenges exist. Below are described inconveniences connected with the use of data from the Register of Real Estate Prices and Values for statistical purposes.

Technical problems

Although there are national guidelines on maintaining the Register of Real Estates Prices and Values, the register is established and maintained at regional level. The units responsible for maintaining the Register of Real Estates Prices and Values are district governors offices. There are 380 such units in Poland. Each of these units is obliged to gather data on real estate transactions in electronic databases, however common information system does not exist. The district governors offices use systems provided by various producers and databases created at regional level are not merged in one national database. Using various information systems entail differences in the regional databases

structure and used symbols. As a result, gathering data from different systems and matching it to create one comprehensive database in timely manner is challenging.

Another technical difficulty which arises when using the Register of Real Estates Prices and Values, is connected with the fact that data from the register are exported in the SWDE or GML format. Using these formats allow for presenting spatial objects in a text file. However, at the same time using the SWDE and GML files does not allow for quick and easy reading of the data they contain. Therefore it was necessary to develop software that converts the SWDE and GML files into database format.

Differences in used concepts

The Register of Real Estate Prices and Values contains data which is not primarily collected for statistical purposes. Consequently, the concepts, definitions and variables used in the register differ to some extent from those required for statistical purposes. Differences in concepts result in the necessity of coverage adjustments. For example, the purpose of the HPI is to measure market transactions, however the register, beside market transactions, covers also non-market transactions such as sales between related parties, donations or sales of public dwellings at very low price to sitting tenants. Therefore it is necessary to identify and exclude any non-market transactions before using the administrative data in the calculations of the HPI.

Quality issues

The use of administrative data rises concern about quality of source data. This concern is connected with the fact that the CSO has no control over data collection and registration procedures. In particular some variables which are essential for the calculations of the HPI are of minor significance for administration authorities. Hence, some variables are of poor quality and thus cannot be used in calculations of the HPI. For instance, in case of single-family houses owners of the Register of Real Estate Prices and Values focus on providing information on the area of land while the information on the useful floor area of a building is very often skipped. The lack of information on the area of buildings makes it impossible to compile price index for single-family houses. As a result, the HPI for Poland is based on information on flats (single-family houses are excluded from calculations).

Timeliness

The way the data collection and registration is organised by the administrative units strongly influences timeliness of administrative data which has impact on timeliness of the HPI. The Register of Real Estate Prices and Values includes data from notarial deeds which shall be forwarded by notaries to district governors offices within two weeks of the date on which the notarial deed was signed. Then the owner of the register is obliged to introduce the data to the database not later than 30 days after receipt of the notarial

deed. However in some cases time lags are even longer. This is particularly for the biggest cities in Poland where the time lags between the date on which the notarial deed was signed and the date on which the information on the transaction was introduced to the register exceed 3 months. As a consequence the HPI for Poland is published about 100 days after the reporting period.

Additional costs
The use of data from the Register of Real Estate Prices and Values on the one hand allows to reduce costs of data collection but on the other hand specific characteristics of this data source entail costs which have to be incurred in order to address quality issues appropriately. In particular it was necessary to develop the electronic system of data collection, the software to convert the SWDE and GML files and the computer system for data validation.

Problems that arise in using administrative data can, at least to some extent, be tackled. The Polish experience shows that it is possible to improve the usability of administrative data by two types of actions: on the one hand close cooperation with administrative data owners and, on the other hand, development and implementation of appropriate data editing and imputation procedures.

Regular contacts with owners of administrative data allow to better understand how registers are being maintained. Moreover these contacts provide a good opportunity to communicate the CSO needs, explain how data are used and emphasise the importance of submitting timely and good quality data. Cooperation with administrative units takes several different forms:
- working contacts with persons supplying data (via e-mail, telephone); these contacts take place whenever any questions are raised concerning supplied data;
- working contacts with software suppliers who provide administrative systems;
- formal letters to authorities responsible for determining rules of maintaining registers as well as to authorities responsible for maintaining registers; this type of contacts aims to provide feedback on the register, inform on the CSO needs and suggest possible improvements;
- meetings with data suppliers; this form of cooperation was used several times in the past and concerned situation when completeness and timeliness of the given register was far from optimal;
- active participation in relevant government programs; in particular, the CSO participate in the government project aimed at introducing the Integrated System on Real Estate Information which envisages unification of data sets maintained at local level and introduction of notarial deeds in electronic form.

It is assessed that significant progress in terms of completeness and quality of the Register of

Real Estate Prices and Values has been achieved in recent years. The further positive developments are expected from 2017 onwards.

The usability of administrative data is also improved by developing and implementing data editing and imputation procedures. The calculation of the HPI is preceded by data validation which covers the following actions:

- detection of incomplete data, incorrect data format, recording practices inconsistent with the rules specified in the law;
- error detection and correction;
- detection and removing non-market transactions;
- detection and removing transactions which relate to sales of a share in the ownership of a property;
- detection and removing duplicated records and records lacking basic information (date, price of property, square meters of floor space);
- checks of consistency between variables;
- creating new variables (e.g. price per square metre, the variable which specifies if the given property is new or existing);
- imputing missing values

3. Methodology

The used methodology was developed taking into account data availability and requirements specified in the Commission Regulation (EU) No 93/2013 of February 2013 laying down detailed rules for the implementation of Council Regulation (EC) No 2494/95 concerning harmonised indices of consumer prices, as regard establishing owner-occupied housing price indices. Moreover, recommendations included in the "Detailed Technical Manual on Owner-Occupied Housing" as well as in the "Handbook on Residential Property Prices Indices" were considered.

The HPI for Poland is based on information on flats (single-family houses are excluded), both new and existing. Data are compiled quarterly on the basis of transaction prices as specified in notarial deeds. Only market prices are taken into account, non-market transactions are ruled-out from calculations. The HPI covers the whole territory of Poland. It is compiled for the country as a whole, voivodships (NUTS 2 level) and the biggest cities.

Calculations of the HPI for Poland are carried out using stratification method. The choice of this method was dictated by the availability of data. As already mentioned the used administrative sources include limited information on dwelling characteristics. The lack of sufficiently detailed description of properties results in inapplicability of the hedonic regression method which is considered as the best technique for constructing a constant quality HPI.

Stratification

The stratification method consists in dividing transactions into possibly homogeneous strata. Hence, the quality of results obtained by this method depends to a large extent on appropriate grouping of properties. Very detailed stratification based on many variables allows to increase homogeneity of defined strata. At the same time, more detailed stratification reduces the number of transaction per strata which can result in unreliable results. This is why, when compiling the HPI for Poland, efforts were made to ensure the balance between the number of defined strata and the number of transactions per strata. Moreover, stratification was carried out in such a manner that it allows to compile price index for each voivodship (NUTS 2 level).

The applied by the CSO of Poland stratification is based on market segment (new/existing flats), geographical location and property size. Separating the market of newly built flats and the market of second-hand flats is necessary as properties sold on these two markets differ significantly in terms of price levels. Moreover, the breakdown into new and existing flats allows to calculate price indices for these two markets, which is the information that meet data user needs. The next stage of the stratification, which is the breakdown by geographical location, is based on the NUTS and LAU classifications. Within each NUTS 2 level territorial unit (voivodship) strata were constructed by combining together LAU 1 units with comparable long-term average price levels. It is observed that for some voivodships it is sufficient to distinguish two strata: one for the capital city of the voivodship and another one for the rest of the voivodship. However, there are also voivodships with more considerable diversity of price levels among LAU 1 units which required more detailed stratification. As a result of the applied geographical stratification 35 strata for the market of newly built flats and 94 strata for the market of second-hand flats were identified. Importantly, there is separate geographical strata for each voivodship capital. The carried out geographical stratification was further extended taking into account the size of sold flats. To this end, three categories of flats were defined: flats with 1 or 2 rooms, flats with 3 rooms and flats with 4 and more rooms.

Outliers identification

Before starting the price indices calculations, the procedure of detecting outliers takes place. Outliers are identified with the use of algorithm based on the interquartile range. Transactions for which prices per square metre are outside limits: Q1-1.5IQR and Q3+1.5IQR (Q1 - lower quartile, Q3 – upper quartile, IQR – interquartile range) are marked as outliers. The final decision on rejection of transactions marked as outliers is taken on a case by case basis.

Calculations

Calculations of the HPI for Poland are carried out in two stages. First, price indices are calculated for each strata, and then price indices at stratum level are combined to obtain price indices at

higher aggregation levels.

Stage 1: Calculation of elementary price indices at stratum level

In the current period the average price per square meter is calculated for each strata based on data on transactions classified to the given strata. Corresponding calculations for the reference period are also carried out. Elementary price indices at stratum level are calculated using Dutot formula i.e. by dividing the average price in the current period by the average price in the reference period which is the last quarter of the previous year.

$$I_i^{\overline{\frac{t,Y}{Q4,Y-1}}} = \frac{P_i^{t,Y}}{P_i^{Q4,Y-1}}$$

$I_i^{\overline{\frac{t,Y}{Q4,Y-1}}}$ — index for stratum *i* in the current quarter *t* in the year *Y*;

$P_i^{t,Y}$ - average price per square meter for stratum *i* in the current quarter *t* in the year *Y*;

$P_i^{Q4,Y-1}$ - average price per square meter for stratum *i* in the last quarter of the previous year.

Stage 2: Calculation of aggregated indices

Price indices for higher aggregation level are obtained using Laspeyres type formula i.e. as weighted average of price indices at strata level. The weights used to calculate the HPI are based on the value of transactions in the previous year (*Y-1*). They are changed annually and price-updated to the 4th quarter of the year *Y-1*.

The index for the aggregate *A* containing n strata can be written as follow:

$$I_A^{\overline{\frac{t,Y}{Q4,Y-1}}} = \sum_{i=1}^{n} w_i^{Y-1(Q4)} * I_i^{\overline{\frac{t,Y}{Q4,Y-1}}}$$

$I_A^{\overline{\frac{t,Y}{Q4,Y-1}}}$ - price index for aggregate *A* which contain n strata;

$I_i^{\overline{\frac{t,Y}{Q4,Y-1}}}$ - index for stratum *i* in the current quarter *t* in the year *Y*;

$w_i^{Y-1(Q4)}$ - weight price-updated to the 4th quarter of the year *Y-1*.

The obtained price indices based on the last quarter of the previous year are chain linked in order to create index series on the same scale.

Currently, the reference base year for all indices is 2010=100.

4.   Results

The results of the calculations are the HPI indices for Poland in the breakdown into new and existing flats as well as the HPI indices at regional level.

The HPI for Poland starts in 2010 and at the moment it covers 6 years. Starting from the beginning of 2011 prices on residential market were continually decreasing until the end of 2013. A slight upward tendency is observed from the beginning of 2014.

The HPI for new and existing flats show broadly similar trends over time. However, the HPI for new flats is lagged behind the HPI for existing flats in the major turning points. Moreover, price decreases at the primary market were lower and increases were higher than at the secondary market. As a result, in Q3 2015 prices of new and existing flats decreased respectively by 0.5% and 6.7% as compared with average prices in 2010.

House Price Index for Poland, 2010=100
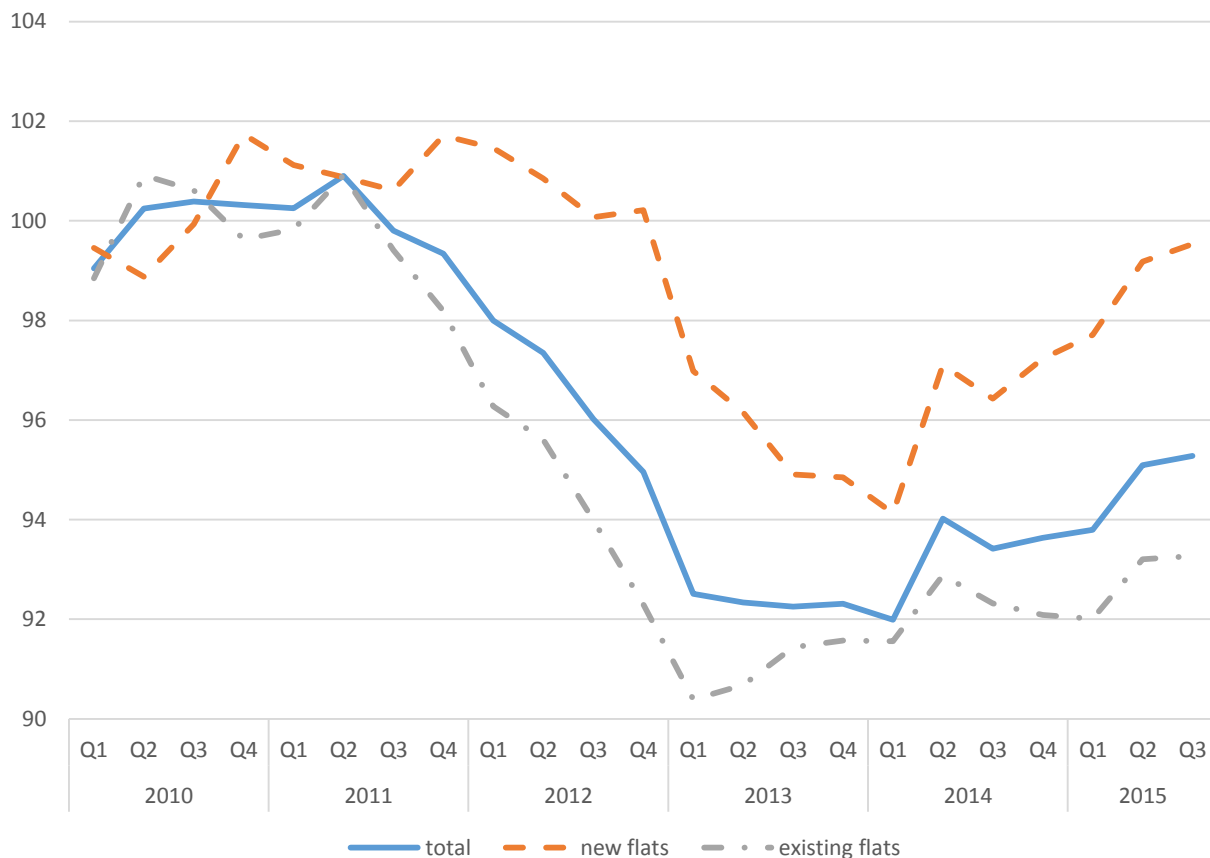
The HPI indices for voivodships (NUTS 2 level) allow observation of regional differences in price changes of residential properties. In Q3 2015 the highest annual increases in residential properties prices were observed in pomorskie (4.5%), mazowieckie (4.4%) and podlaskie (4.4%).

In turn, the highest decreases were observed in zachodniopomorskie (-3.9%) and świętokrzyskie (-3.3%).

HPI at NUTS 2 level in Q3 2015, annual growth rate
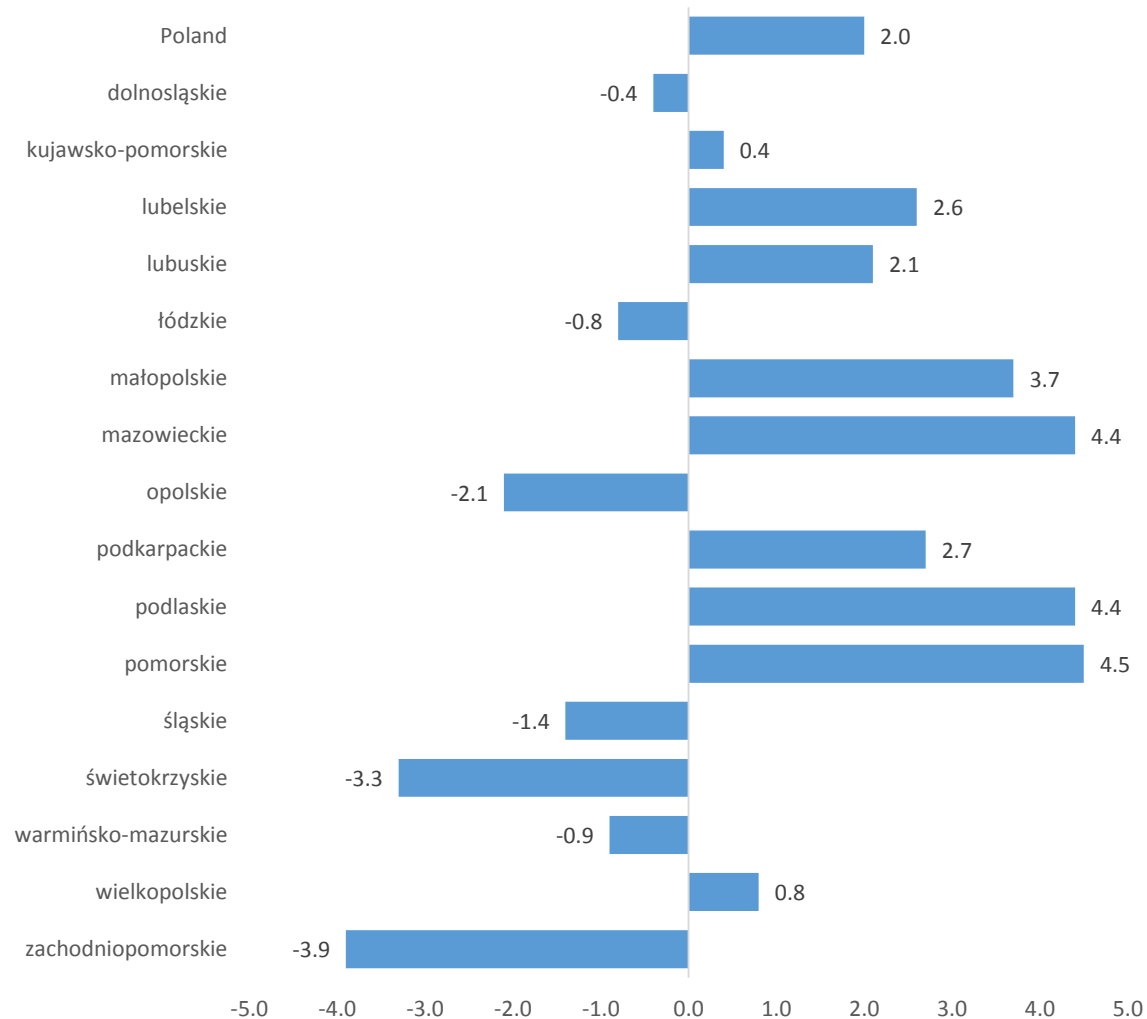
Another source of information on residential properties prices in Poland are price indices provided by the National Central Bank of Poland (NBP). The NBP data covers the biggest cities in Poland and are calculated using hedonic method. The chart below shows the HPI for Poland compiled by the CSO and the hedonic house price indices compiled by the NBP in the breakdown into Warsaw, 6 biggest cities and 10 remaining cities. It is notable that the CSO and the NBP data reveals similar long-run trends. However, the CSO data are not directly comparable with the NBP data because of differences in coverage and used methodology.

House price indices compiled by the CSO and the NBP, annual growth rate
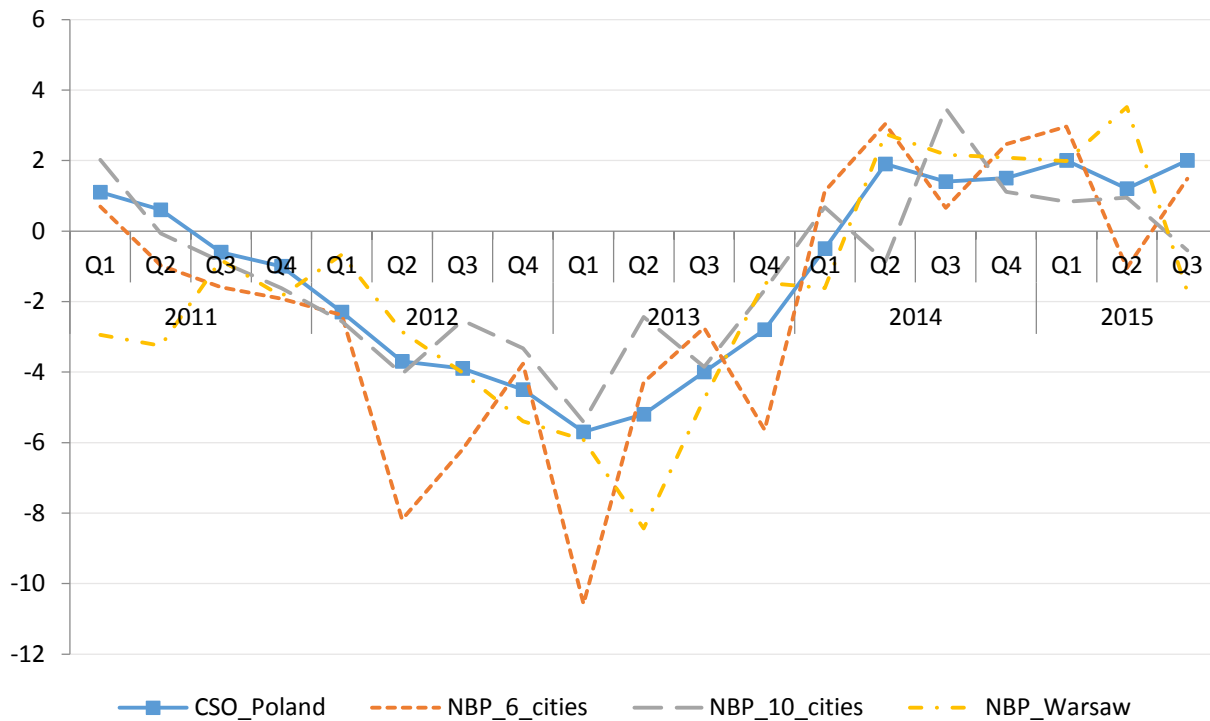
Source: CSO and NBP

5. Future developments

The CSO of Poland plans to further develop house price statistics. It is expected that it will be possible to improve the quality of obtained results thanks to planned changes in the Register of Real Estate Prices and Values. Specifically, the on-going government plan aimed at introducing the Integrated System on Real Estate Information envisages unification of data sets maintained at local level and introduction of notarial deeds in electronic form. It should ensure quick access to up-to-date and reliable information on real estate. However this improvements will be introduced in the long-term perspective.

In the interim, the works will be focused on finding solutions which allow to enrich information available in the Register of Real Estate Prices and Values. In particular, it is planned to merge information from the administrative source with information available in the Building Database which is maintained by the CSO of Poland on the basis of data from the Housing Census and other surveys on housing stock in Poland. It is expected that it will enable to extend the scope of the HPI to single-family houses. Access to more detailed information on properties characteristics should also allow the replacement of the stratification method with hedonic method which is the preferred solution. Moreover, it is necessary to improve timeliness of results which will be achieved by further improvements in organisation of the survey.

References

Act of May 17, 1989 The Geodetic and Cartographic Law (Journal of Laws on 2010, No. 193, entry 1287 with further amendments)

Commission Regulation (EU) No 93/2013 of 1 February 2013, laying down detailed rules for the implementation of Council Regulation (EC) 2494/95 concerning harmonised indices of consumer prices, as regards establishing owner-occupied housing price indices [2013] OJ L33/14

Decree of the Minister of Regional Development and Housing of March 29, 2001 on registration of land and buildings (Journal of Laws on 2001, no. 38, entry 454 with further amendments)

Eurostat, Detailed Technical Manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices, Draft 2.0, March 2012

Eurostat, Handbook on Residential Property Prices Indices (RPPIs), 2013