OECONOMIA
COPERNICANA

VOLUME 9 ISSUE 1 MARCH 2018

p-ISSN 2083-1277, e-ISSN 2353-1827
www.oeconomia.pl

ORIGINAL PAPER

Citation: Głuszak, M., Czerski, J., & Zygmunt, R. (2018). Estimating repeat sales residential price indices for Krakow. *Oeconomia Copernicana, 9*(1), 55–69. doi: 10.24136/oc.2018.003

Contact to corresponding author: gluszakm@uek.krakow.pl; *Cracow University of Economics*, ul. Rakowicka 27, 31-510 Kraków, Poland
Received: 23 March 2017; Revised: 14 December 2017; Accepted: 7 January 2018

Michał Głuszak
Cracow University of Economics, Poland

Jarosław Czerski
Instytut Analiz Monitor Rynku Nieruchomości mrn.pl, Poland

Robert Zygmunt
University of Agriculture in Krakow, Poland

Estimating repeat sales residential price indices for Krakow

JEL Classification: *C18;C43; R31*

Keywords: *repeat sales index; Poland; housing market; house price; real estate*

Abstract
Research background: There are several methods to construct a price index for infrequently traded real estate assets (mainly residential, but also office and land). The main concern to construct a valid and unbiased price index is to address the problem of heterogeneity of real estate or put differently to control for both observable and unobservable quality attributes. The one most frequently used is probably the hedonic regression methodology (classic, but recently also spatial and quantile regression). An alternative approach to control for unobservable differences in assets' quality is provided by repeat sales methodology, where price changes are tracked based on differences in prices of given asset sold twice (or multiple times) within the study period. The latter approach is applied in renown S&P CoreLogic Case-Shiller house price indices.
Purpose of the article: The goal of the paper is to assess the applicability of repeat sales methodology for a major housing market in Poland. Previous studies used the hedonic methodology or mix adjustment techniques, and applied for major metropolitan areas. The most widely known example is the set of quarterly house price indices constructed by NBP

— especially for the primary and secondary market. The repeat sales methodology has not been adopted with significant success to date — mainly because of concern regarding relative infrequency of transactions on the housing market in most metropolitan areas (thus a potentially small sample of repeated sales).
Methods: The study uses data on repeat sales of residential transactions in Krakow from 2003 to 2015. We apply different specifications of repeat sales index construction and compare respective values to the hedonic price index for Krakow estimated by NBP.
Findings & Value added: Findings suggest that repeat sales house sales indices can be used to track price dynamics for major metropolitan areas in Poland. The study suggests problems that need to be addressed in order to get unbiased results — mainly data collection mechanism and estimation procedure.

Introduction

In order to construct a valid and unbiased real estate price index is to address the problem of heterogeneity. In other words, the key issue is to control for both observable (measurable) and unobservable attributes (location, neighborhood, structural and environmental). While difficult to build, real estate indices are extremely useful to investigate long-run economic processes. Among remarkable examples are Shiller's (2014, pp. 1486–1517, 2015) analysis of economic turbulences in US (1890–2014), Nicholas and Szczerbina's (2013, pp. 278–309) investigation into house price movement in 1920s and 1930s, not to mention historic studies of house prices in the long run in Netherlands (1628–1973) conducted by Eichholtz (1997, pp. 175–192), as well as China (1644–1840) conducted by Raff *et al.* (2013, pp. 368–386). Two former studies have one thing in common — in both cases repeat sales methodology was used.

Repeat sales method of index construction is a way to account for unobservable differences in a given asset's quality by investigating price changes between sales. The method was introduced by (Bailey *et al.*, 1963). The data requirements include the sample of real estate sold twice (or multiple times) within the study period. The modified repeat sales approach, based on the seminal paper by Case and Shiller (1987), has been applied in renowned S&P CoreLogic Case-Shiller house price indices. Although praised for theoretical soundness, the method has some limitations, one of them being the need for the relatively active real estate market (thus large sample of repeat sales). It is by no means easy to suffice this particular data requirement on thin real estate markets. As the result, repeat sales methodology is rarely used as a method for applied index construction. To date, the repeat sales house price index methodology has not been applied in Poland. The paper aims to address this gap, by analyzing properties of residential repeat sales index in Krakow.

The article is organized as follows. Section 2 discusses the previous research on repeat sales method of index construction. In particular, it reviews the theoretical underpinnings, as well as advantages and disadvantages of the method. It provides a brief resume of previous papers applying the method in the context of residential properties — both foreign and Polish. Section 3 describes the data gathering process and introduces the econometric properties of the repeat sales regression. In section 4 we analyze the repeat sale residential real estate index in Krakow. We assess the validity of the results by comparison to other measures of house price volatility – such as National Bank of Poland house price indices.

Literature review

There is by no means a consensus regarding the methodology of real estate price dynamics measurement, save for the commonly shared opinion that property price indices are difficult to estimate. Nevertheless, in most cases two competing approaches, to deal with the observed and unobserved heterogeneity of real estate have been used in the literature — i.e. hedonic and repeat sales methods. The differences between both methods have drawn the attention of economist studying real estate price movements, mostly because in many cases they tend to produce incomparable insights into market volatility. In a US study, authors used a big data (1.1 mln sales) only to find that the result suggests significantly different cycles in metropolitan areas (Dorsey *et al.*, 2010, pp. 87–105). Nevertheless, in most cases results remained inconclusive regarding the choice of appropriate house price index methodology (Case *et al*., 1991).

The extensive summary of repeat sales methods of property price index construction provides a list of advantages and disadvantages of the technique (Prud'homme & Diewert, 2011, pp. 1–10). The list of former contains: limited data requirements, relative calculation ease, ease of reproduction, control for salient real estate characteristics (Prud'homme & Diewert, 2011, pp. 1–10), while latter can be contributed to: lack of depreciation correction, sample selection bias, data inefficiency, continuous revision problem (Prud'homme & Diewert, 2011, pp. 1–10).

The basic comparison of hedonic, repeat sales and hybrid methods are presented in a table (Table 1).

The shift from classic house price indices towards a hybrid approach (Jones, 2010, pp. 95–97; Nagaraja *et al*., 2014, pp. 23–46), has been followed by other innovative solutions to the old problem of measuring the house prices dynamics.

Bourassa *et al.* (2006, pp. 80–97) tested sale price appraisal ratio (SPAR) method in New Zeland and compared it to both hedonic and repeat-sales indices based on the same data. In the SPAR, the ratio between the transaction price and previously assessed property values is calculated. They suggest that while it gives fairly similar results to repeat sales index, it has relative simplicity advantage. The SPAR house price index was also constructed using Dutch data (de Vries *et al.*, 2009, pp. 214–223). Others have addressed depreciation problem and suggested techniques to disentangle pure time effect from property depreciation (Cannaday *et al.*, 2005, pp. 320–342; Englund *et al.*, 1998, pp. 175–192). In yet another article Francke (2010, pp. 24–52) analyzed the methods to estimate repeat sales index for a thin market

Recently scholars have begun advocating for a more robust method of index estimation, especially in presence of outliers. Zhang and Yi used quantile regression approach to construct repeat sales index in Beijing (Zhang & Yi, 2017, pp. 85–96), and Gwangju, Korea (Yeon, 2016, pp. 260–267), while the others experimented with pseudo-repeat sales techniques based on matched data (Guo *et al.*, 2014, pp. 20–38).

Both hedonic and repeat sales indices are difficult to build and update on regular bases. Institutional background of constructing real estate price index was described based on French experiences (Gouriéroux & Laferrère, 2009, pp. 206–213).

Real estate price index construction has been discussed in Polish economic literature. The selection of articles discussing various methodological issues includes Foryś (2012, pp. 41–52), Kokot (2014, pp. 14–27; 2015, pp. 84–100), and Trojanek (2010, pp. 5–14) and Trojanek and Trojanek (2012, pp. 74–84). The most discussed and well-known house price index in Poland has been published by National Bank of Poland. It consists of several indices for asking and sales prices on primary and secondary residential market. It has three major variants: (1) average, (2) median-based as well as (3) hedonic index. The latter was discussed in Widlak and Tomczyk (2010, pp. 203–227). Recent developments on hedonic index construction, include the regression splines and spatial methods to addresses spatial autocorrelation and smoothing (Trojanek *et al*., 2017; Widłak *et al*., 2015). To authors' best knowledge repeat-sales approach has not been applied on housing data in Poland, although an interesting adoption of the method to art market was proposed by Kempa and Witkowska (2011, pp. 181–186).

To conclude, to our best knowledge, there is no evidence on the validity of repeat sales price index construction method in Poland. We argue that due to data limitations, and the idiosyncratic character of residential mar-

kets in major Polish cities no valid analogies cannot be drawn from the various application of the method in other countries. The empirical part of the paper aims to address this gap, thus we believe the research is justified.

Research method

Data sources and management

Although there are multiple sources of information used to construct real estate price indexes — for example asking prices, sale prices, valuations (Pollakowski, 1995), for obvious reasons the choice is limited in case of the repeat-sales index. The most valid and reliable source of information, although not always efficient in terms of information provided, are notary acts.

Our empirical data comes from a database managed by Institute of Real Estate Market Institute mrn.pl, which is a professional organization of property valuers (Polish chartered surveyors), property market consultants and market experts. Currently, the database covers information of 296.5 thousand real estate transaction, 113.7 thousand of them in Krakow. The database is created on the basis of information obtained from notary acts and is the complete information source on the housing market in Krakow since 2002. It covers approx. 99% of housing transactions on the primary and secondary housing market, surpassing the other real estate market databases (along with National Bank of Poland database).

Based on the address and land register, we have matched all transactions involving the same properties and identified repeat sales from 4 quarter 2002 to 3 quarter 2015. As the research is focused on the secondary market, we dropped all sales from the primary market (sold by developers) from the sample. As the result, we have identified 2584 properties sold 2 times, 246 properties sold 3 times, 23 properties sold 4 times and 1 property sold five times during the study period. We decided to recode the data and match all transactions sold multiple times in pairs (first-second transaction). For example, information on 246 properties sold 3 times allowed us to create 492 repeat sales pairs (Table 2).

After re-coding the initial dataset, we addressed the quality of the data, in order to eliminate all possible sources of bias. We used three-step sampling procedure.

Firstly, we eliminated all observations where the time difference between sales was less than 180 days. This step in the procedure is in line with the commonly accepted view that short turnover period between sales

can be attributed mostly to atypical price movements — speculative behavior, cognitive biases, peculiar property conditions (Englund *et al*., 1998, p. 195) or even fraud. We recorded 343 repeat sales where the turnover period was 180 days and less (Figure 1).

As it can be easily seen from the picture, time differences in our repeat sales sample follow right-skewed distribution (possibly log-normal type) — thus short turnover periods being more frequent. This finding must be treated with caution, as on rare occasions the turnover period was less than 7 days. The repeat sales like these should be removed from the sample.

To investigate whether the short turnover period influenced the recorded price difference between sales, we explored the distribution of price ratios between in each pair. The results of the investigation for a subsample of properties sold again within 90 days from the first purchase is presented in a figure (Figure 2).

Data depicted on the figure (Figure 2) suggest that although in case of a substantial number of properties resale price was comparable (or even equal) to previous price (a result perfectly in line with sticky price hypothesis) in many cases the price change certainly did not reflect the market price movements. There are numerous examples of properties resold for a 150% higher price, within only 90 days of acquisitions. On rare occasions, the ratio was less than 0.5, which means that property was resold for less than half of the initial price.

In the second stage, we checked whether apartments did not change significantly between sales. Notably, we verified floor areas, to check whether the apartment was not significantly remodeled. Additionally, we traced for equity in additional premises being sold (parking places, storage rooms, cellars, etc.). All pairs where properties have undergone significant improvements were dropped from the sample.

Thirdly, we eliminated all properties with abnormal values or with additional clauses that significantly affect price (distressed properties, municipal sales, the sale between relatives).

After the data cleaning procedure, we got our final sample (effective) of 2704 repeat sales. The major differences between full transaction dataset (containing information of all residential transactions from 4 quarter 2002 to 3 quarter 2015), full repeat sales dataset (including abnormal observations) and final effective repeat sales database are presented in a Table 3.

The comparison of basic descriptive statistics between sales and repeat sales samples reveals interesting differences. On average, the apartments in both repeat sales subsamples were smaller (average floor area of 46.6 or 47.1 sqm of usable space compared to 55.5 in the whole sample). The geographic distribution of observation shifted significantly as well. While most

apartments being sold in Krakow from 2012 to 2015 were located in Podgórze (32.1%), in our repeat sales subsamples most observations came from Środmiescie (32.5–32.9%), and the share of observations from Podgorze dropped to approx. 22%. The results indicate the potential sample selection problem — residential properties sold twice or multiple times in Krakow differ significantly from typical transactions during the study period.

Estimation procedure

During the data management phase, we prepared a dataset on residential properties *n* sold more than once over the study period $t=0,\ldots T$. Within each pair, we compare the price at first sale p_n^s to the price at resale p_n^t. Due to data limitation (sample size), we decided to construct a quarterly residential properties price index. In line with the literature, we assume that multiple sales are treated as a set of independent matched pairs. Repeat sales regression is based on the following specification (Balk *et al.*, 2011, p. 66):

$$\ln\left(\frac{p_n^t}{p_n^s}\right) = \sum_{t=0}^{T} \gamma^t D_n^t + \mu_n^t \tag{1}$$

where D_n^t is a time dummy variable (it takes value -1 in the time period of the first sale, 1 in the time period of the second sale, and 0 otherwise), whereas μ_n^t is an error term.

Repeat sales index can be calculated based on regression coefficients (Balk *et al*., 2011, p. 67):

$$I^{0t} = \exp\left(\hat{\gamma}^t\right) \cdot 100 \tag{2}$$

The model is subject to heteroskedasticity (for discussion see: Case & Shiller, 1987, pp. 45–56), thus several methods, alternative to standard OLS, were suggested to overcome this problem. Following guidelines from the literature, we estimated three repeat sales indices, based on:

- OLS estimation (referred to as the baseline or the unweighted model).
- WLS estimation to correct for possible heteroscedasticity (the weighted model), the three-step procedure introduced by Case and Shiller (1987, pp. 45–56).
- Quantile estimation suggested by McMillen and Thorens (2006, pp. 567–584), referred to as the quantile model. The method should be robust to outliers and produce more stable results.

We compare the results with NBP Hedonic House Price Index as well as simple weighted average house price index based on the original apartments' sales mrn.pl database. The result is discussed in the following section.

Results and discussion

We have investigated the price dynamics in Krakow using repeat sales dataset. The analysis reveals a number of outliers, where price differential was bigger than suggested by the time difference. In most cases, this finding reflects timing effect — as some agents/speculators were able to buy and sell at a right time. The price ratio between the first and the second sale in particular cases was higher than 5 (Figure 3).

Based on estimation procedures described in the previous section, we have calculated three alternative repeat sales indices (RSIs) for residential properties in Krakow: baseline RSI_OLS, weighted RSI_WLS and quintile (median) RSI_QL. Then we compared the results with the values of two benchmark indices — hedonic house price index for a secondary market in Krakow calculated by (HPI), and simple weighted average residential price index based on a full sample of mrn.pl data (WAI). The weighting was location-based. Average prices were calculated based on transaction subsamples in 4 major districts in Krakow (Nowa Huta, Podgorze, Krowodrza, and Srodmiescie). Weighting was based on transaction volume in the first period (4 quarter 2002). The comparison period ranged from 4Q 2002 to 3Q 2007. In case of HPI, the first index values come from 3Q 2006, which was set as a base period (I=100) for all indices.

The results were presented in Figure 4.

The results depicted in Figure 4 imply substantial differences between indices. Both RSI_OLS and RSI_WLS were in line with simple weighted average residential price index WAI. On the other hand, all three indices differed significantly from the hedonic price index (HPI) calculated by NBP. The main difference can be attributed to higher price appreciation during housing boom period, and less apparent price decrease since the

market peak in 2 or 3 quarter 2007. A quantile (median) repeat sales index (RSI_QL) yielded different results and suggested more rapid price decrease from the market peak in 3Q 2007. The differences between the results require further examination. Although quantitative analysis of differences was not a subject of this paper, a potential expansion could follow the work of Nagaraja *et al.* (2014).

The results should be treated with caution for two complementary reasons: (1) sample differences between NBP and mrn.pl transaction datasets, (2) possible problems regarding the sample selection problem in repeat sales indices.

Conclusions

In the paper have explored the properties of repeat sales index for residential properties in Krakow. Based on a sample of 2704 repeat sales drawn from a larger transaction database, we have estimated three repeated sales house price indices. The results were later compared to renown National Bank of Poland hedonic house price index for the secondary housing market in Krakow and simple weighted average index based on a mrn.pl transaction database.

The results imply significant differences between repeat sales indices and NBP hedonic index, especially regarding market cycle. Repeat sales indices suggest that housing market peaked about 1 quarters later (in 3 quarter 2007). The results should be treated with caution. Although mrn.pl database in more detailed and significantly larger, the sample alone does not explain the differences found.

One concern regarding repeat sales index in Krakow is a sample selection problem. We have found that properties sold twice or multiple times differed significantly from full sample of housing transactions, both in terms of size and location. This result is in line with Clapp and Giacotto (1992, pp. 300–306). On average they were smaller, and central locations were overrepresented. There are two basic explanations of the latter phenomenon. Firstly, smaller apartments are frequently bought by young households, who are fairly active in a housing market, and tend to change their apartment following family lifecycle (especially when the family grows). That would explain the higher turnover rate in the case of smaller flats. Secondly, small flats in central locations may be more popular as alternative investment assets (and used rental apartments), thus having a higher turnover rate. As we have not tested that, we conclude that further research is needed. One obvious direction for further empirical research

involves using hybrid index construction methods. Additionally, as Krakow is a fairly unique housing market (prone to speculative behavior, due to touristic attractiveness), it would be interesting to replicate the results on a data from the other major metropolitan market in Poland.

References

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, *58*(304). doi:10.1080/01621459.1963.10480679.

Balk, B., De Haan, J., & Diewert, E. (2011). *Handbook on residential property prices (RPPIs)*. doi:10.2785/34007.

Bourassa, S. C., Hoesli, M., & Sun, J. (2006). A simple alternative house price index method. *Journal of Housing Economics*, *15*(1). doi: 10.1016/j.jhe. 2006.03.001.

Cannaday, R. E., Munneke, H. J., & Yang, T. T. (2005). A multivariate repeat-sales model for estimating house price indices. *Journal of Urban Economics*, *57*(2). doi:10.1016/j.jue.2004.12.001.

Case, B., Pollakowski, H. O., & Wachter, S. M. (1991). On choosing among house price index methodologies. *Real Estate Economics*, *19*(3). doi: 10.1111/1540-6229.00554.

Case, K., & Shiller, R. (1987). Prices of single-family homes since 1970: new indexes for four cities. *New England Economic Review*, September/October.

Clapp, J. M., & Giacotto, C. (1992). Estimating price indices for residential property: a comparison of repeat sales and assessed value methods. *Journal of the American Statistical Association*, *87*.

De Vries, P., de Haan, J., van der Wal, E., & Mariën, G. (2009). A house price index based on the SPAR method. *Journal of Housing Economics*, *18*(3). doi: 10.1016/j.jhe.2009.07.002.

Dorsey, R. E., Hu, H., Mayer, W. J., & Wang, H. Chen. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, *19*(2). doi: 10.1016/j.jhe.2010.04.001.

Eichholtz, P. M. A. (1997). A long run house price index: the Herengracht index, 1628–1973. *Real Estate Economics*, *25*(2). doi: 10.1111/1540-6229.00711.

Englund, P., Quigley, J. M., & Redfearn, C. L. (1998). Improved price indexes for real estate: measuring the course of Swedish housing prices. *Journal of Urban Economics*, *44*(2). doi: 10.1006/juec.1997.2062.

Foryś, I. (2012). Metoda mix-adjustment wyznaczania indeksów cen nieruchomości mieszkaniowych na przykładzie lokali spółdzielczych. *Studia I Materiały Towarzystwa Naukowego Nieruchomości*, *20*(1).

Francke, M. K. (2010). Repeat sales index for thin markets. *Journal of Real Estate Finance and Economics*, *41*(1). doi: 10.1007/s11146-009-9203-1.

Gouriéroux, C., & Laferrère, A. (2009). Managing hedonic housing price indexes: the French experience. *Journal of Housing Economics*, *18*(3). doi: 10.1016/j.jhe.2009.07.012.

Guo, X., Zheng, S., Geltner, D., & Liu, H. (2014). A new approach for constructing home price indices: the pseudo repeat sales model and its application in China. *Journal of Housing Economics*, *25*. doi: 10.1016/j.jhe.2014.01.005.

Jones, C. (2010). House price measurement: the hybrid hedonic repeat-sales method. *Economic Record*, *86*(272). doi: 10.1111/j.1475-4932.2009.00596.x.

Kempa, K., & Witkowska, D. (2011). Returns from the art market. Price index evaluated for the most-traded Polish painters. *Quantitative Methods in Economics*, *15*(2). doi: 10.4135/9781849209397.

Kokot, S. (2014). The analysis of differences in residential property price indices. *Real Estate Management and Valuation*, *22*(3).

Kokot, S. (2015). Diversified price dynamics in some sub-segments of the housing market. *Folia Oeconomica Stetinensia*, *84*. doi:10.1515/foli-2015-0030.

McMillen, D. P., & Thorsnes, P. (2006). Housing renovations and the quantile repeat-sales price index. *Real Estate Economics*, *34*(4). doi: 10.1111/j.1540-6229.2006.00179.x.

Nagaraja, C., Brown, L., & Wachter, S. (2014). Repeat sales house price index methodology. *Journal of Real Estate Literature*, *22*(1). doi: 10.5555/reli.22.1.f22430v375351169.

Nicholas, T., & Scherbina, A. (2013). Real estate prices during the roaring twenties and the great depression. *Real Estate Economics*, *41*(2). doi: 10.1111/j.1540-6229.2012.00346.x.

Pollakowski, H. O. (1995). Data sources for measuring house price changes. *Journal of Housing Research*, *6*(3).

Prud'homme, M., & Diewert, E. (2011). Repeat sales methods. In *Handbook on residential property price indices*.

Raff, D., Wachter, S., & Yan, S. (2013). Real estate prices in Beijing, 1644 to 1840. *Explorations in Economic History*, *50*(3). doi: 10.1016/j.eeh.2012.10.003.

Shiller, R. J. (2014). Speculative asset prices. *American Economic Review*, *104*(6).

Shiller, R. J. (2015). *Irrational exuberance*. Princeton University Press.

Trojanek, R. (2012). An analysis of changes in dwelling prices in the biggest cities of Poland in 2008-2012 conducted with the application of the hedonic method. *Actual Problems of Economics*, *7*(2).

Trojanek, R., Tanas, J., Raslanas, S., & Banaitis, A. (2017). The impact of aircraft noise on housing prices in Poznan. *Sustainability*, *9*(11). doi: 10.3390/su9112088.

Trojanek, M., & Trojanek, R. (2012). Profitability of investing in residential units: the case of real estate market in Poland in the period. *Actual Problems of Economics*, *2*(7).

Widlak, M., & Tomczyk, E. (2010). Measuring price dynamics: evidence from the Warsaw housing market. *Journal of European Real Estate Research*, *3*(3). doi: 10.1108/17539261011094722.

Widłak, M., Waszczuk, J., & Olszewski, K. (2015). Spatial and hedonic analysis of house price dynamics in Warsaw. *NBP Working Paper*.

Yeon, K. (2016). A study on repeat sales house price index based on penalized quantile regression. *Procedia Computer Science*, *91*. doi: 10.1016/j.procs.2016.07.071.

Zhang, L., & Yi, Y. (2017). Quantile house price indices in Beijing. *Regional Science and Urban Economics*, *63*. doi: 10.1016/j.regsciurbeco.2017.01.002.

Acknowledgements

Michal Gluszak acknowledges financial support from the funds allocated to development of research potential of the Faculty of Economics and International Relations of the Cracow University of Economics for 2017

Annex

Table 1. Comparison of selected econometric methods used to house price index

Method	Explanation
1. Hedonic method	Price dynamics is tracked based on cross-sectional data on real estate transactions (offers). The use of regression allows controlling for housing quality observed differences between properties and to separate pure price index for given periods of time. Requires detailed data on housing attributes.
2. Repeat sales method	Price dynamics is analyzed based on observed differences in the price of a given property between sales. Allows taking account of unobserved quality differences. based. Requires relatively simple data.
3. Hybrid method	Price dynamics is tracked based on both sales and repeat sales. The method can account for observable quality changes between sales of a given property. Requires detailed data on housing attributes.

Table 2. Initial sample properties and recoding

Type	N (initial)	N* (pairs after recoding)
Double sales	2584 (90.54%)	2584
Triple sales	246 (8.62%)	492
Quadruple sales	23 (0.81%)	69
Quintuple sales	1 (0.03%)	4

Table 3. The full sample of residential property sales and subsamples of repeat sales in Krakow from 2002 to 2015

Metropolitan areas	Sales sample (full)	Repeat sales sample (full)	Repeat sales sample (effective)
Size (N)	58739	3149	2704
Mean floor area (sqm)	55.5	46.6	47.1
Location (district)			
Krowodrza	16,093 (27.4%)	772 (24.5%)	656 (24.3%)
Nowa Huta	10,314 (17.6%)	645 (20.8%)	561 (20.7%)
Podgorze	18,905 (32.1%)	697 (22.2%)	597 (22.1%)
Środmiescie	13,427 (22.9%)	1024 (32.5%)	890 (32.9%)

Figure 1. Repeat sales turnover period in Krakow from 2002 to 2015

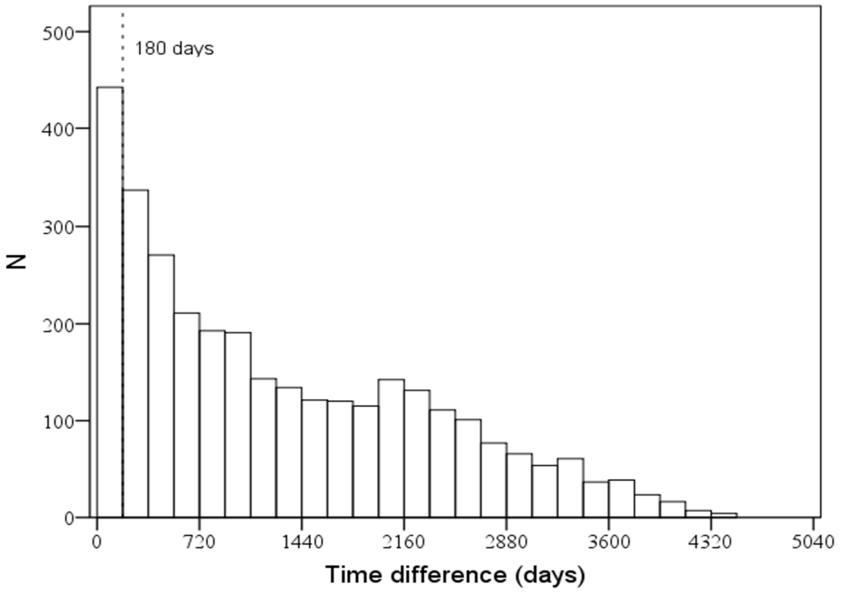

Figure 2. Price differences for short resale properties (turnover less than 90 days)

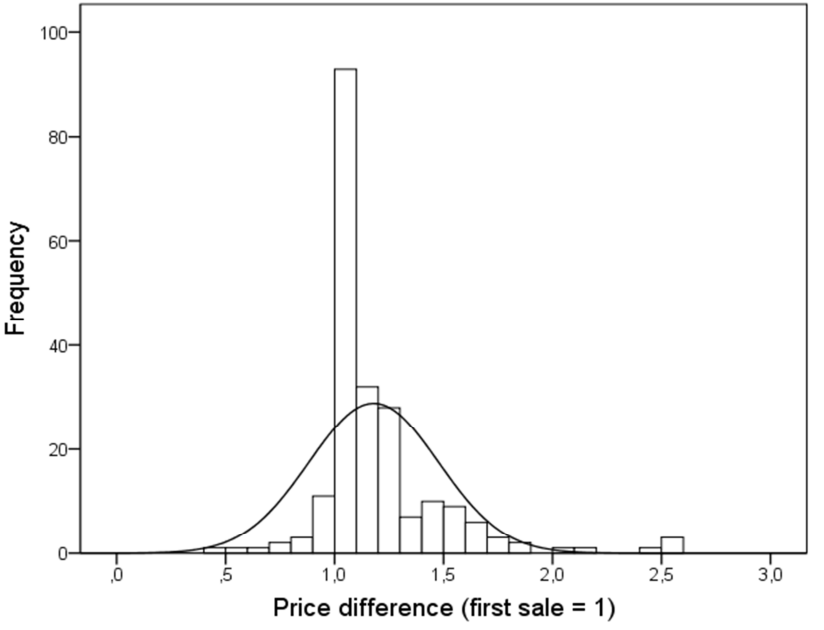

Figure 3. The relation between turnover period and price differences between sales

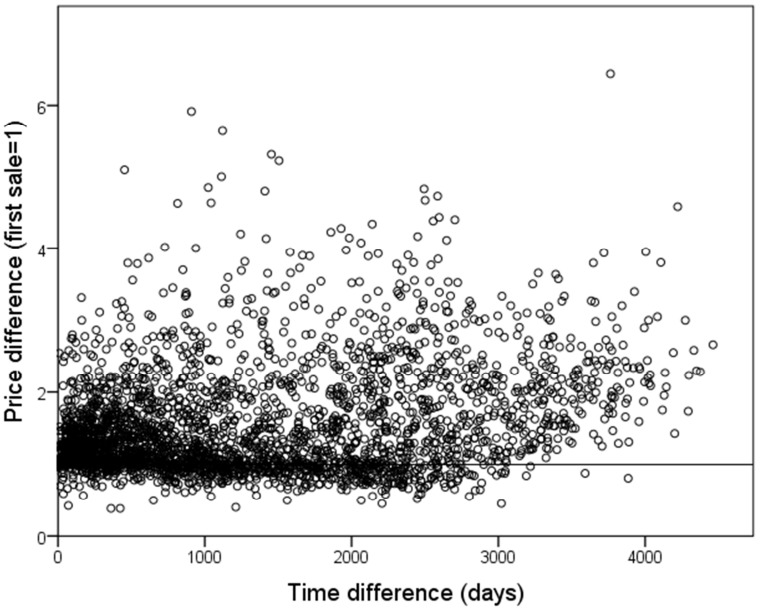

Figure 4. Comparison of residential real estate indices in Krakow (2002-2015)