Acta Sci. Pol., Administratio Locorum 20(1) 2021, 5–18.
https://czasopisma.uwm.edu.pl/index.php/aspal plISSN 1644-0749 eISSN 2450-0771 DOI: 10.31648/aspal.6111

ORIGINAL PAPER Accepted: 20.01.2021

THE REGIONAL SPATIAL DIVERSITY OF HOUSING PRICES AND MARKET ACTIVITY – EVIDENCE FROM POLAND

Radosław Cellmer[1], Aneta Cichulska[2], Mirosław Bełej[3]

[1] ORCID: 0000-0002-1096-8352
[2] ORCID: 0000-0001-6627-5259
[3] ORCID: 0000-0001-8650-6990
[1,2,3] University of Warmia and Mazury in Olsztyn
Prawochenskiego street 15, 10-720 Olsztyn, **Poland**

ABSTRACT

The aim of this study is to identify the factors that significantly determine the regional spatial differentiation of housing prices as well as housing market activity in Poland. The present research makes the assumption that average housing prices and market activity (number of transactions) are regionally shaped by economic, social, infrastructural and environmental conditions which can be described as a set of diagnostic features ascribed to a given area, constituting a statistical unit. Furthermore, it is assumed that individual effects may appear, being tied to the idiosyncrasies and significance of the given area. The time horizon of the research is 2008–2018. Based on data sourced from the Central Statistical Office a panel data was prepared for each of 380 Polish districts (poviats). Next, parameters were estimated for a single-factor panel model, as well as a two-factor model in which the constant term is different for different time periods and different units. This resulted in a model encompassing both average price determinants, and individual effects which reflect certain regularities of their spatial distribution. Moreover, the research will result in a set of cartograms made with Geographic Information System tools, depicting the random effects resulting from estimates of panel models using the Nerlove and Swamy-Arora transformations.

Keywords: spatial diversity, spatial autocorrelation, panel data modelling, geomodelling, housing market, geographical information system

INTRODUCTION

Residential construction has been a booming business in Poland (2008–2018), reflected in the number of new residential units, building permits issued and building works started. The trading frequency on the real estate market is also increasing, specifically including residential housing. The increase in supply and market size in the real estate segment is caused by an increasing demand for housing units, as well as for built-up land and land intended for housing. That demand, in turn, is the result of the significant development of the national economy, including a dynamic growth in Gross Domestic Product (GDP), indicating the growing wealth of the population. A growth trend in average monthly earnings has held for a number of years, although accompanied by rising prices of goods and services

rcellmer@uwm.edu.pl, aneta.cichulska@uwm.edu.pl, miroslaw.belej@uwm.edu.pl

(food prices increased by 4.6%, home upkeep costs – by 1.6%, transport – by 3.8%). Deflation or low inflation enabled constant interest rates (the central bank's (NBP) reference rate has stood at a record low 1.5% for the last 3 years), thereby increasing the stability of the money market, and indirectly capital market and accessibility of mortgage loans (a nearly 10% growth in new housing loans in 2014–2017 and a nearly 8% growth in new housing loans value), and ultimately – the financial accessibility of real estate purchases by the public. Economic development coupled with continued improvements on the labour market (lowest unemployment rate in many years: 4.5% according to Labour Force Survey, 6.6% according to Central Statistical Office), are borne out in the enduring growth trend of the housing market.

The aim of the research is to identify a set of factors that significantly determine the regional spatial differentiation of housing prices as well as housing market activity in Poland. The poviat was adopted as the statistical unit, in accordance with the statistical nomenclature of territorial units adopted in Poland, developed on the basis of the European Nomenclature of Territorial Units for Statistics (NUTS). Poland has a three-stage administrative division, with 2477 municipalities, 380 poviats and 16 voivodeships. The choice of a poviat as the average level of division of the country is a compromise between too much detail (municipalities) and too much generality (voivodeships). For the purposes of this research we have assumed that the spatial differentiation of residential property prices is determined by regional levels of economic, social, infrastructural and environmental factors. Therefore, the effect of the research will only indicate significant regional determinants of the price heterogeneity of apartments at the level of 380 poviats. The research will use time series of regional economic, social, infrastructural and environmental data and average prices of apartments and market activity (number of transactions) from 2008–2018. The research covers a relatively calm period after the 2008 crisis and before the current "coronacrisis". The use of panel data in research allows for more extended modelling than in case of classical cross-sectional data, because they contain more variability and less collinearity between variables. Moreover, the research will result in a set of cartograms of, so called, random effects by district resulting from estimated models for housing prices with the use of Nerlove's and Swamy-Arora transformations. The research used a statistical analysis software (Statistica, R, Gretl) and Geographic Information System tools (QGIS, ArcGIS Pro).

The article consists of an introduction, three chapters and a conclusion. A review of the literature on spatial modelling of the factors determining price volatility in residential markets has been conducted in the first chapter. The second chapter presents the data description and brief description of methodology. The third chapter presents detailed results together with a discussion of the results obtained. The article closes with conclusions.

LITERATURE REVIEW

The real estate sector is one of the principal foundations of the national economy, not only stimulating economic growth in a keyway, but also creating an environment favourable for dynamic development. The fundamental nature of this sector consists in its role as a repository of fixed assets, as it creates spatial conditions for the development of various other sectors of the economy and presents a very important way to allocate capital. The market economy creates a system in which most sectors are correlated to others, whereby a given sector is susceptible to changes outside of itself and may influence the conditions in other sectors. This complexity of links and processes is also present in the housing market, necessitating constant scrutiny of events arising in it, in order to predict future market conditions [Adams & Füss, 2010]. According to Urbanavičienė et al. [2009] the growth or decline of the housing sector considerably affects the general growth or decline of a country's economy. One of the fundamental tasks of the state and local authorities is fostering conditions conducive for satisfying the residential needs of the society. Housing market requires its efficiency in allocating

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

the existing resources and the possibilities of creating new supply adequate in terms of quality and quantity to the notified demand constitute the outcome of quite a numerous group of economic and non-economic factors. Economic changes on the macro, meso and local levels influence the situation in the residential market to different degrees. Doubtless, a vital role in shaping the real estate market is played by the financial sector [Beltratti & Morana, 2010], including WIBOR (Warsaw Interbank Offered Rate) interest rates, accessibility of mortgage loans, formal and legal factors, including the legislation in force and general principles shaping the market, as well as economic trends, such as GDP dynamics. A relatively large number of research studies on macroeconomic determinants of real estate prices can be observed [Wang et al., 2017, Panagiotidis & Printzis, 2016, Belke & Keil, 2018]. Over the years, the relationship between macroeconomic indicators and the dynamics of real estate prices has been studied. One of the earliest is an analysis of real estate prices in the USA based on cross-sectional data from 1977 to 1991 [Hendershott & Abraham, 1992]. The findings establishing a close relationship between prices and interest rates were confirmed in research by Himmelberg et al. [2005] and those presented by Iacoviello and Minetti [2003], the latter of which focused on the influence of financial liberalisation on the relationship between monetary policy and real estate prices in Finland, Sweden and Great Britain. In turn, Brunnermeier and Julliard [2008] found that the level of real estate prices is shaped by inflation, with its probable influence on future economic recession. Adams and Füss [2010] turned special attention to the dependence of housing demand on interest rates, specifically mortgage interest rates and the costs of real estate developers. Demand models presented by Attanasio et al. [2012], and Eichholtz and Lindenthal [2014] prove that the age structure and educational structure of a population are also of significant importance. Research with a similar focus was undertaken, among others, by Nguyen and Wang [2010] – focusing on the GDP indicator and the prices of goods and services, Meidani et al. [2011] – the GDP indicator, prices of goods and services, currency exchange rates, Lastrapes [2002] – monetary shocks, or Tsatsaronis and Zhu [2004] – GDP, inflation, interest rates, and Englund and Ioannides [1996] – GDP and interest rates.

The housing market is incontrovertibly local, and hence is primarily shaped by local conditions [Hendershott & Abraham, 1992] which include, among others, indicators characterising the labour market (unemployment rate, average wage, etc.), factors relating to market size influencing price volatility – the number of inhabitants and households, age structure, marriage rate, etc. Nevertheless, it is the factors characterising the labour market, with their ability to shape demand and reflect the aggregate societal wealth, which are pointed to as the main determining factors of the housing market and its price levels [Adams & Füss, 2010]. An example of the utility of these factors is the research conducted by Żelazowski [2011], which relied on regional indicators such as: unemployment rate, average wage, population, extant and newly built housing. On the macroeconomic side, the research tracked GDP per capita.

The sheer number of hypotheticals determining factors in residential real estate prices burdens an individual assessment of their impact with a high likelihood of estimation error. Current research makes increasing use of analytical methods, including models based on panel data. Such data encompasses a set of independent variables along with dependent variables from more than one period. In other words, they are produced by aggregating time series of indicators for particular units of analysis. Therefore, they exhibit characteristics of cross-sectional data from a single period, as well as of time series pertaining to a single variable over different periods. Panel models also reflect the defining impact of the time factor on the dependent variable using lagged dependent variables [Liu et al., 2018]. There has been little research done in Poland using panel models to identify determining factors in housing prices, especially spatial aspects (more about panel models in 3.1).

As it can be observed from the review of literature, the research presented in the article is consistent with the current trend of research on the housing market, which takes into consideration social and economic factors.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

DATA DESCRIPTION AND METHODOLOGY

Data description

The analysis conducted was based on data concerning the housing property market in Poland. A district (poviat) was adopted as a statistical unit, in line with the nomenclature of statistical territorial units adopted for statistical purposes in Poland, prepared on the basis of the European Nomenclature of Territorial Units for Statistics (NUTS). The analysis of factors determining differences in average residential real estate prices was performed using transactional and statistical data for 380 districts in Poland from 2008 to 2018, which were made available by the Central Statistical Office (GUS) and its Local Data Bank (BDL) via the website [https://bdl.stat.gov.pl/BDL, date: 10.10.2020].

The initial intent was to analyse the most recent 11 years by quarter or half-year periods. Ultimately, due an incomplete of such data, we decided to analyse over time periods of one year, which was also due to a greater availability of other statistical data. The table below presents variables which were analysed as a select set of indicators of environmental, demographic and economic conditions. The use of most of these variables is borne out in the literature, while in other cases the deciding motivation was an intention to find new and under-researched correlations. A summary of information on the adopted variables is presented in Table 1. The dynamics of average housing prices (PRC) and the average

Table 1. Description of variables

Symbol	Description	Unit
PRC	Average housing price*	PLN/m^2
NB	Number of transactions	number/1000 population
POP	Population density	persons/km^2
BRTH	Birthrate	persons/1000 population
MOB	Share of mobile working age population in total population	%
MIGR	Migration rate	persons/1000 population
SAL	Average salary	PLN/month
EMP	Registered unemployment rate	%
BSN	Entities entered into register of business entities (REGON)	number/1000 population
POL	Emission of dust pollution PM10	t/km^2
AV	Average usable floor area of a dwelling unit per person	m^2
DWL	Completed new dwellings	units/1000 population

* 1 PLN ≈ 0.22 EUR
Source: own research.

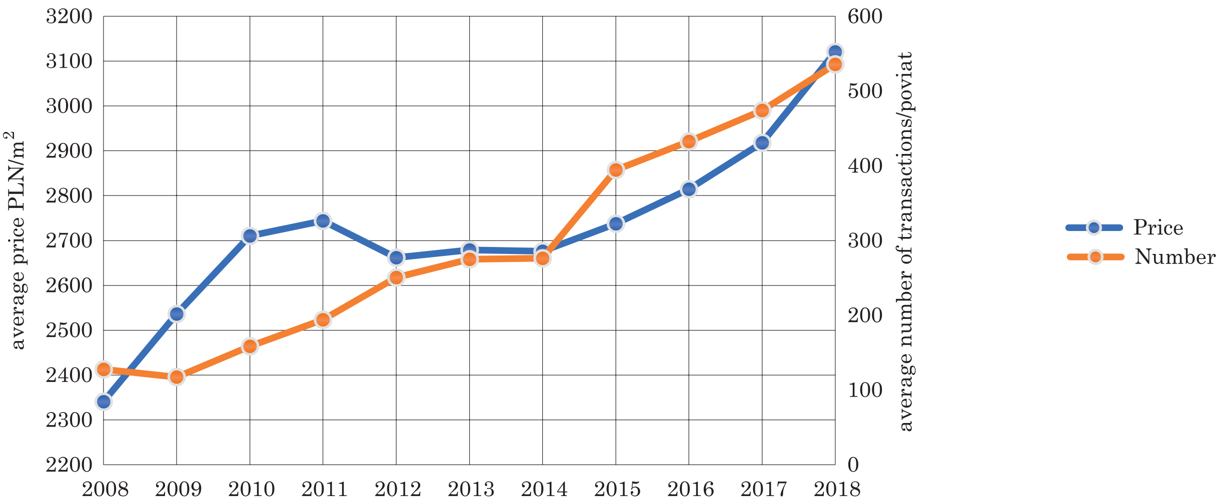

Fig. 1. Average price of 1 square metre of apartments and average number of sales in 2008–2018
Source: own research.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

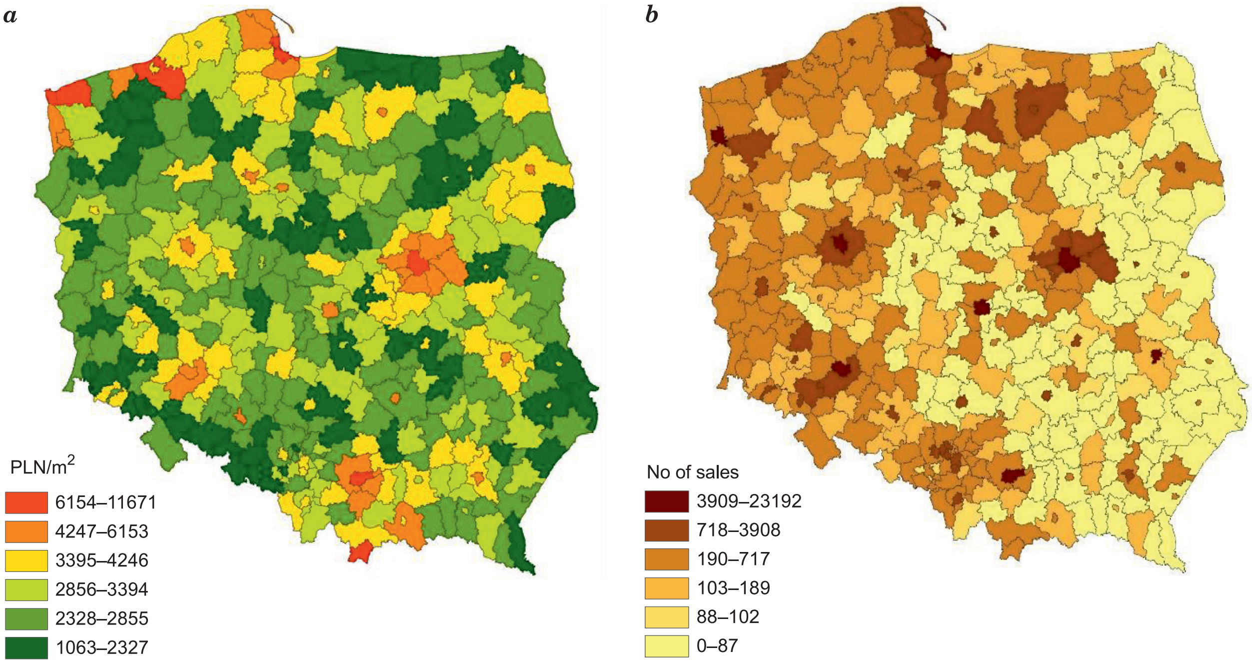

Fig. 2. Average housing price per m^2 (a) and the number of apartment sales (b) in 2018, by district
Source: own research.

number of transactions (NB) in Poland in the years 2008–2018 has been shown in Figure 1.

Since 2008, the average price per m^2 of apartment has been steadily rising (in 2017, the average selling price per m^2 was 10.4% higher from that of 2014), along with the average number of apartment sales in Polish districts in the recent years. As an example, Fig. 2 presents the spatial distribution of the PRC variable (average price per m^2 of apartment) and spatial distribution of the NB variable (number of transactions) on example on data from 2018.

Values concerning the average price of 1m^2 of an apartment and the number of apartment sales in 2018 divided by district in natural clusters highlight differences between metropolitan areas and typical areas with low investment.

METHODOLOGY

Data encompassing cross-sectional and time dimensions, which are used in creating panel models contain both cross-sectional information (a description of a population in a single period), and time-series information (a description of a unit in many periods) [Mátyás & Sevestre, 2013]. Panel models relate their dependent variables not only to independent variables, but also to unmeasurable, constant factors individual for each unit, known as effects. Panel models are uniquely able to discern differences between periods for the same object, as well as between different objects in the same period. The model is formulated as below [Baltagi, 2008]:

$$y_{it} = \beta_0 + \sum_{k=1}^{k} \beta_{kit} x_{kit} + \alpha_i + v_t + \varepsilon_{it} \qquad (1)$$

where:

y_{it} – dependent variable
x_{kit} – independent (explanatory) variable
β_0 – constant term
β_{kit} – structural parameter
α_i – individual effects (part of variation of variable y characteristic for i-th object)
v_t – period effects (part of variation of variable y characteristic for period t)
ε_{it} – random disturbance component
$i = 1, \ldots, N$ – sequence of objects

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

t = 1, …, T – sequence of time periods
k – independent variable number.

Panel data analysis is performed by the classical least squares estimation method, the fixed effects model, or the random effects model [Mátyás & Sevestre, 2013]. Estimation of panel data model with the ordinary least squares method (OLS) is employed if all objects are uniform and empirical data departs from hypothetical values of a dependent variable only as a result of the random component [Baltagi, 2008]. Such an estimation is permissible if there is no individual effect and the panel is a cross-sectional data set. A model based on OLS is formulated as follows [Wooldridge, 2002]:

$$y_{it} = x_{it}\beta + v_{it} \tag{2}$$

where:
y_{it} – dependent variable
x_{it} – independent variable (altogether, independent variables column vector)
β – vector of N structural parameters
v_{it} – aggregate random error, composed of the purely random component ε_{it} and individual effect u_i pertaining to the specific i-th unit of the panel.

Another method of panel data analysis is the panel model with fixed effects (FE – Fixed Effects Model), which eliminates fixed individual and period effects by averaging them over time (t index) [Wooldridge, 2002]. Its formulation is as follows:

$$y_{it} = \alpha_i + \sum_{k=1}^{k} \beta_{kit} x_{kit} + \varepsilon_{it} \tag{3}$$

where parameter α_i is taken as an individual characteristic of each unit, the estimation of which encompasses the impact of all characteristics not reflected in the observable variables vector. In the random effects panel model (RE), each unit is ascribed a random variable responsible for the individual effect in the given period. Ultimately, individual effects are not treated in the form of parameters, and the model assumes the following form:

$$y_{it} = \beta_0 + \sum_{k=1}^{k} \beta_{kit} x_{kit} + v_{it} \tag{4}$$

The model assumes that independent variables and random components, and their individual effects, are independent for all individuals. Consequently, assuming that constants are fixed results in the random component reflecting the discrepancies between objects and periods (one-way model). In the reverse situation, where the factor is variable relative to different periods and objects, the resulting model is two-factor (two-way model).

Estimation of the panel data model can make use the classical least squares method if the condition of estimator consistency for total error and pure random error is satisfied, and a correlation between an individual effect u_i and explanatory variable x_{it} does not exist [Baltagi, 2008]. The model with random effects assumes that the random component contains both individual and periodical effects. The Breusch-Pagan test shall be used to investigate whether the variance of random components for all observations is constant, based on statistics from a sample of the following form (Lagrange multiplier test):

$$LM = \frac{nt}{2(t-1)}\left[\frac{\sum_{i=1}^{n}(\sum_{t=1}^{t}\varepsilon_{it})^2}{\sum_{i=1}^{n}\sum_{t=1}^{t}\varepsilon_{it}^2} - 1\right]^2 \tag{5}$$

where n is the number of observations, t denotes the number of time units, while ε_{it} are the residuals of the total regression model. With the truth of the null hypothesis, the above statistics has a chi-square distribution with one degree of freedom.

Including group and time effects in the panel models makes it necessary to use specific estimation methods. The use of the classical method of least squares encounters difficulties resulting from the fact that the assumptions of Gauss-Markov concerning the properties of a random component are not usually satisfied. Since random components in the RE model are correlated, in this situation, a generalized least squares estimator of structural parameters of the following form is used to estimate model parameters:

$$\hat{\beta}_{RE} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} y \quad (6)$$

where X is the matrix of explanatory variables, y is a vector of response variables, while Ω denotes a reversible variance-covariance matrix of the total random error [Baltagi, 2008].

The decision to choose the appropriate model form (FE or RE) is made on the basis of Hausman's test, which consists in comparing the values of estimated parameters obtained with the use of both estimators. The null hypothesis H_0 then states that both estimators (FE and RE) are not biased, but in such a situation RE is more efficient, with the alternative hypothesis H_1 under which the FE estimator is not biased and the RE estimator is biased, or an error in the model specification occurred. The test statistic is defined by the following formula:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})[var(\hat{\beta}_{RE}) - var(\hat{\beta}_{FE})]^{-1}(\hat{\beta}_{RE} - \hat{\beta}_{FE}) \quad (7)$$

and has a chi-square distribution with the number of degrees of freedom equal to the number of parameters estimated in both models.

Generalised least squares estimation (GLS) requires estimating the variance within each cross-sectional unit (σ_ε^2) and variance between units (σ_ν^2), referred to simply as "within" and "between". There are several methods of estimating those values, whose detailed descriptions can be found in Baltagi [2008], Swamy and Arora [1972], and Nerlove [1971]. The Swamy and Arora method of estimation concerns the transformation of the model based on the following formula:

$$\bar{y_i} - y = \beta_1(\bar{x}_{1i} - \bar{x}_1) + \cdots + \beta_k(\bar{x}_{ki} - \bar{x}_k) + \bar{v}_i - \bar{v} \quad (8)$$

In a similar way, variables for individual periods of time are transformed. Alternative estimators are proposed by Nerlove [1971] with the use of following transformation:

$$y_{it}^* = \beta_1 x_{1it}^* + \cdots + \beta_k x_{kit}^* + v_{it}^* \quad (9)$$

where

$$z_{it}^* = z_{it} - \theta_1\, \bar{z}_i - \theta_2\, \bar{z}_t + \theta_3\, \bar{z}_{it}$$

and

$$z_{it} \in \{y_{it}, x_{1it}, \ldots, x_{kit}, x_{kit}\ \}$$

and

$$\theta_1 = 1 - \frac{\sigma_\varepsilon}{\sqrt{T\sigma_\mu^2 + \sigma_\varepsilon^2}}, \quad \theta_2 = 1 - \frac{\sigma_\varepsilon}{\sqrt{N\sigma_\lambda^2 + \sigma_\varepsilon^2}},$$

$$\theta_3 = \theta_1 + \theta_2 + \frac{\sigma_\varepsilon}{\sqrt{T\sigma_\mu^2 + N\sigma_\lambda^2 + \sigma_\varepsilon^2}} - 1 \quad (10)$$

where μ_i is the individual-specific error component and λ_t is period-specific error component.

Panel models are widely used for econometric analysis or for economic analysis, but also for studies of road safety, dependencies of demographic phenomena on economic factors, or environmental pollution and economic growth. The importance of panel models is also highlighted by Griliches and Intriligator [2007] and Hsiao [2003].

RESULTS AND DISCUSSION

Panel modelling was preceded by an analysis of correlations between the adopted variables. High correlation between explanatory variables is an unfavourable phenomenon, although in the case of a relatively large number of variables it is difficult to avoid this phenomenon. The correlation matrix is presented in Table 2.

The variables most strongly correlated with the average price (PRC) are BSN and DWL. For the number of transactions, the variables most closely correlated with it are POP, SAL and BSN. A clear correlation also applies to the explained variables themselves. Some inconvenience may be the relatively high correlation between MIGR and DWL as well as between BSN and DWL variables. These variables are not directly related to each other in terms of content and at the same time they can bring important information to the model. Therefore, it was decided to include them in further analyzes. However, it should not significantly reduce the quality of the estimated models. The results of OLS modelling for the explained variables PRC and NB are presented in Table 3. It should be noted that this model was used only for preliminary dependency assessment.

rcellmer@uwm.edu.pl, aneta.cichulska@uwm.edu.pl, miroslaw.belej@uwm.edu.pl

Table 2. The correlation matrix

	POP	BRTH	MOB	MIGR	SAL	EMP	BSN	POL	AV	DWL	PRC	NB
POP	1	-0.073	0.010	-0.101	0.269	-0.288	0.355	0.458	-0.100	0.133	0.381	0.475
BRTH	-0.073	1	0.351	0.372	-0.171	-0.050	0.236	-0.060	-0.212	0.407	0.152	-0.020
MOB	0.010	0.351	1	0.092	-0.374	0.344	0.132	0.137	-0.384	0.059	-0.145	-0.192
MIGR	-0.101	0.372	0.092	1	0.052	-0.273	0.361	-0.140	0.467	0.677	0.318	0.085
SAL	0.269	-0.171	-0.374	0.052	1	-0.467	0.214	0.141	0.374	0.211	0.420	0.497
EMP	-0.288	-0.050	0.344	-0.273	-0.467	1	-0.173	-0.120	-0.456	-0.314	-0.400	-0.265
BSN	0.355	0.236	0.132	0.361	0.214	-0.173	1	0.121	0.221	0.530	0.602	0.491
POL	0.458	-0.060	0.137	-0.140	0.141	-0.120	0.121	1	-0.153	-0.029	0.084	0.117
AV	-0.100	-0.212	-0.384	0.467	0.374	-0.456	0.221	-0.153	1	0.398	0.370	0.133
DWL	0.133	0.407	0.059	0.677	0.211	-0.314	0.530	-0.029	0.398	1	0.559	0.374
PRC	0.381	0.152	-0.145	0.318	0.420	-0.400	0.602	0.084	0.370	0.559	1	0.489
NB	0.475	-0.020	-0.192	0.085	0.497	-0.265	0.491	0.117	0.133	0.374	0.489	1

Source: own research.

Table 3. Results of OLS modelling

	Dependent variable: PRC			Dependent variable: NB		
Variable	coef.	std. error	p-value	coef.	std. error	p-value
intercept	2686.400	508.718	<0.001	11.096	1.193	<0.001
POP	0.294	0.020	<0.001	0.001	<0.001	<0.001
BRTH	36.832	5.580	<0.001	-8.083	0.013	<0.001
MOB	-54.882	7.205	<0.001	-0.202	0.017	<0.001
MIGR	-2.546	0.430	<0.001	-0.005	0.001	<0.001
SAL	222.585	19.721	<0.001	1.143	0.046	<0.001
EMP	-8.732	2.202	<0.001	0.015	0.005	0.003
BSN	135.296	5.234	<0.001	0.268	0.012	<0.001
POL	-20.231	6.145	0.001	-0.113	0.014	<0.001
AV	48.796	5.516	<0.001	-0.143	0.012	<0.001
DWL	123.257	7.325	<0.001	0.295	0.017	<0.001
	$R^2 = 0.567$, $F (10, 4169) = 547.185$ p-value<0.001			$R^2 = 0.527$, $F (10, 4169) = 464.950$ p-value<0.001		

Source: own research.

The overall results of the OLS estimation can only be considered as a preliminary analysis. However, it shows which variables can actually be treated as predictors of prices and number of transactions. In this case, all variables turned out to be statistically significant at a significance level (p-value) of less than 0.001. On the one hand, this may indicate that all explanatory variables significantly affect prices and the number of transactions, while on the other, a large number of degrees of freedom should be taken into account, which certainly affects the p-value assessment. Therefore, a certain distance should be approached to assess the quality of classic regression models built on the basis of cross-sectional data, especially when we use the ordinary least squares method for estimation. The results of OLS modeling may, however, constitute an important premise to conclude that the set of variables characterizing socio-demographic, eco-

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

nomic and environmental conditions has been selected in an appropriate way.

An appropriate panel model and a relevant method of estimation were selected on the basis of the Breusch-Pagan test and the Hausman test. The results of tests carried out for models with explanatory variables PRC and NB are presented in Table 4.

Table 4. Results of the Breusch-Pagan test and the Hausman test

Test	dependent variable: PRC	dependent variable: NB
LM	9325.100 p<0.001	3395.750 p<0.001
H	183.518 p<0.001	718.102 p<0.001

Source: own research.

A low p-value in the Breusch-Pagan test (LM) counts against the null hypothesis that pooled OLS model is adequate, in favour of the random effects alternative. Although the Hausman test indicates the FE model as consistent, the RE model was used because it provides the opportunity to analyze the spatial variation of random effects. Therefore, it was decided to estimate the model taking into account random effects. Both Swamy-Arora and Nerlove estimators were used in the construction of the panel models [Swamy & Arora, 1972, Nerlove, 1971]. Table 5 shows the parameters of the RE models (one way) for dependent variable PRC, i.e. average prices.

Comparison of information criteria indicates that a slightly better model is the one obtained as a result of the Swamy-Arora transformation. This is also indicated by the standard error of the residuals. In both models similar values of coefficients were obtained, with the biggest differences concerning the constant of the model. All variables turned out to be statistically significant. The variables MOB, MIGR, EMO and POL are average price destimulants, while the others have a positive impact on prices.

The signification of individual variables and their hierarchy can be determined either by a model using standardised values of the variables or simply by dividing the coefficient by the variable span. It turned out that by far the most important variable is the SAL variable, meaning the average salary, which may translate directly into demand resulting from the purchasing power. The least significant is the population density in the adopted statistical units.

Table 5. Model estimation results (GLS) for dependent variable PRC using Nerlove's and Swamy-Arora transformations

	Nerlove transformation dependent variable: PRC			Swamy-Arora transformation dependent variable: PRC		
Variable	coef.	std. error	p-value	coef.	std. error	p-value
intercept	1431.230	549.749	0.009	1645.77	535.439	0.002
POP	0.255	0.021	<0.001	0.267	0.021	<0.001
BRTH	18.948	5.656	<0.001	24.671	5.647	<0.001
MOB	-38.018	7.939	<0.001	-39.649	7.656	<0.001
MIGR	-2.929	0.416	<0.001	-2.816	0.421	<0.001
SAL	336.082	22.338	<0.001	289.983	21.281	<0.001
EMP	-21.510	2.344	<0.001	-17.634	2.304	<0.001
BSN	145.850	5.522	<0.001	143.360	5.434	<0.001
POL	-23.586	5.663	<0.001	-22.718	5.804	<0.001
AV	46.277	5.616	<0.001	46.174	5.593	<0.001
DWL	114.463	7.151	<0.001	117.559	7.210	<0.001
	LogLik: -33111.92 Std. error of residuals: 667.655 AIC: 66245.85 Schwarz: 66315.57			LogLik: -33070 Std. error of residuals: 661.040 AIC: 66162.60 Schwarz: 66232.32		

Source: own research.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

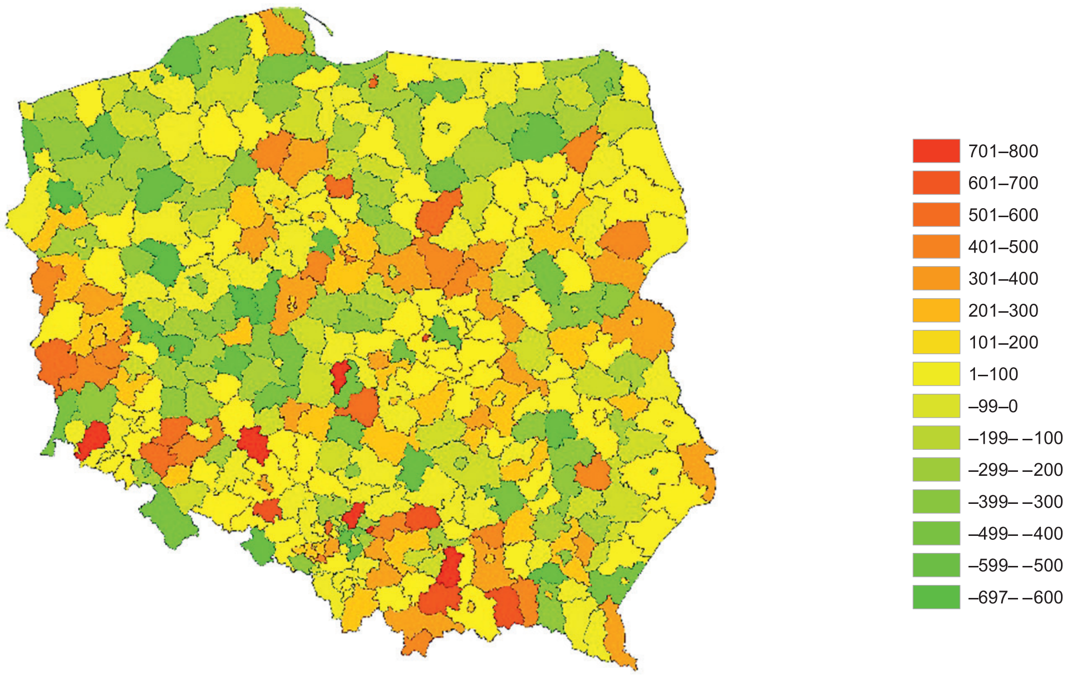

Fig. 3. Random effects by district resulting from estimated models for dependent variable PRC with the use of Nerlove's transformations
Source: own research.

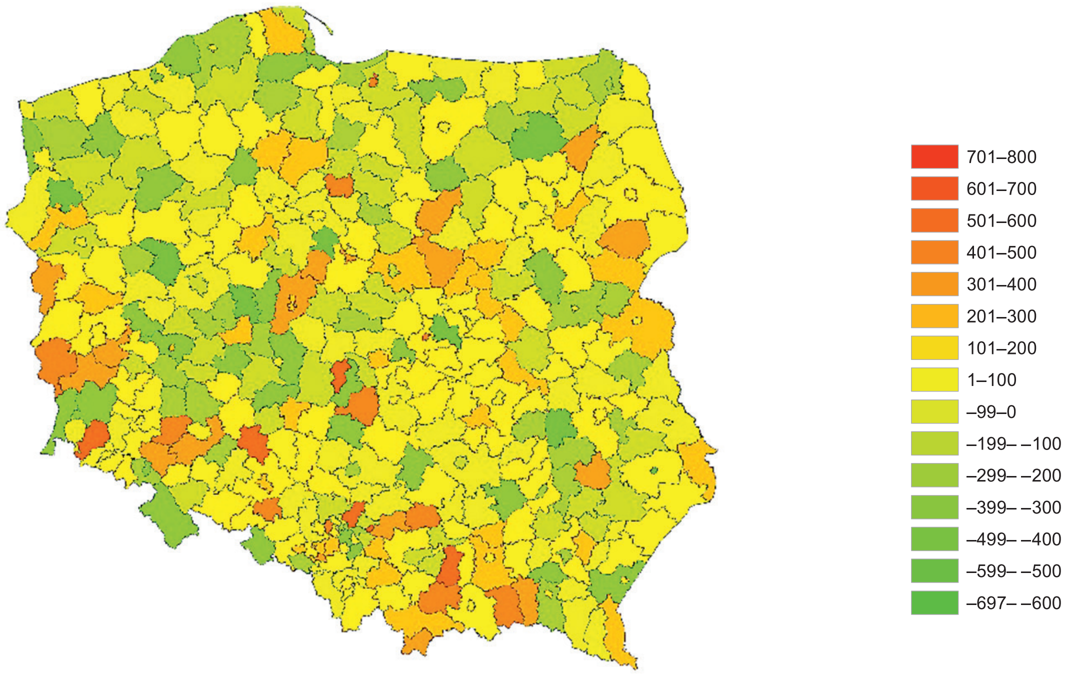

Fig. 4. Random effects by district resulting from estimated models for dependent variable PRC with the use of Swamy-Arora transformations.
Source: own research.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

Figures 3 and 4 present the distribution of random effects in individual districts for both models.

The distribution of random effects in both models looks similar. In the Nerlove transformation model, these effects are slightly greater in terms of absolute value. The highest positive values are found mainly in the southern and south-western part of Poland. This may indicate a large market potential, resulting from the relatively high industrialization of these regions, as well as their location near the main communication routes. While in the case of positive random effects it is difficult to observe spatial autocorrelation, in the case of negative effects they are concentrated in the northern and central-western part of Poland.

The modelling results for the dependent variable NB, i.e. the number of transactions, are presented in Table 6.

The comparative analysis of the presented models shows that the model in which the Swammy-Arora transformation was used is much better than the model in which the Nerlove transformation was used. This is indicated by both the value of the logarithm of the credibility function and the information criteria (AIC, Schwarz). This is indicated by both the value of the logarithm of the likehood function and the information criteria (AIC, Schwarz). Similar conclusions can be drawn from the standard error of the residuals. Among the significant differences one can also indicate the sign of the coefficient with the variable AV (average usable floor area of a dwelling unit per person). This variable is a stimulant in the Nerlove transformation model and a destimulant in the second one. Both models indicate that share of mobile working age population in total population and emission of dust pollution are factors that negatively affect the number of transactions. The other variables are stimulants. However, it does not have to mean cause-effect relationships (eg. unemployment rate).

All variables turned out to be statistically significant. The comparison of the results obtained with the range of values of the variables indicates that, similarly as in the case of average prices, the biggest influence on the market activity has the average monthly salary (variable SAL). Figures 5 and 6 present the distribution of random effects concerning the number of transactions in particular districts for both models.

Table 6. Model estimation results (GLS) for dependent variable NB using Nerlove's and Swamy-Arora transformations

	Nerlove transformation dependent variable: NB			Swamy-Arora transformation dependent variable: NB		
Variable	coef.	std. error	p-value	coef.	std. error	p-value
intercept	27.830	1.402	<0.001	26.679	1.325	<0.001
POP	0.001	<0.001	<0.001	0.001	<0.001	<0.001
BRTH	0.095	0.020	<0.001	0.055	0.017	0.001
MOB	-0.533	0.019	<0.001	-0.469	0.018	<0.001
MIGR	0.011	0.002	<0.001	0.005	0.001	<0.001
SAL	0.476	0.087	<0.001	1.042	0.064	<0.001
EMP	0.047	0.006	<0.001	0.050	0.006	<0.001
BSN	0.046	0.014	<0.001	0.109	0.014	<0.001
POL	-0.131	0.021	<0.001	-0.115	0.019	<0.001
AV	0.167	0.026	<0.001	-0.042	0.019	0.026
DWL	0.182	0.016	<0.001	0.201	0.016	<0.001
	LogLik: -8726.190 Std. error of residuals: 1.954 AIC: 17474.38 Schwarz: 17544.10			LogLik: -8059.987 Std. error of residuals: 1.666 AIC: 16141.97 Schwarz: 16211.69		

Source: own research.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

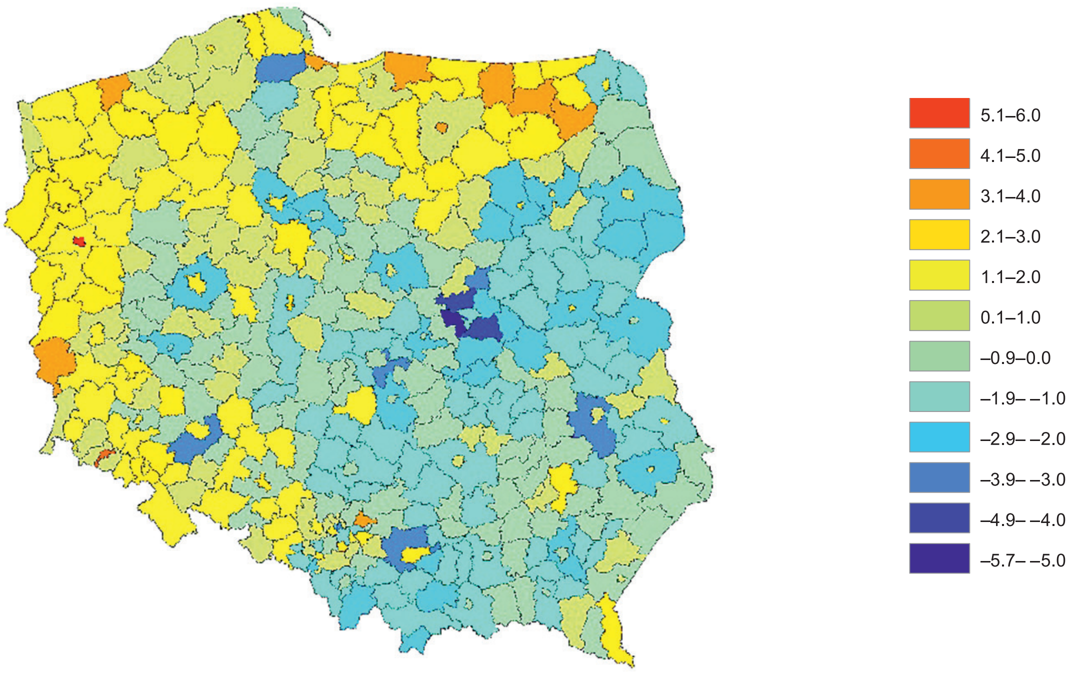

Fig. 5. Random effects by district resulting from estimated models for dependent variable NB with the use of Nerlove's transformations

Source: own research.

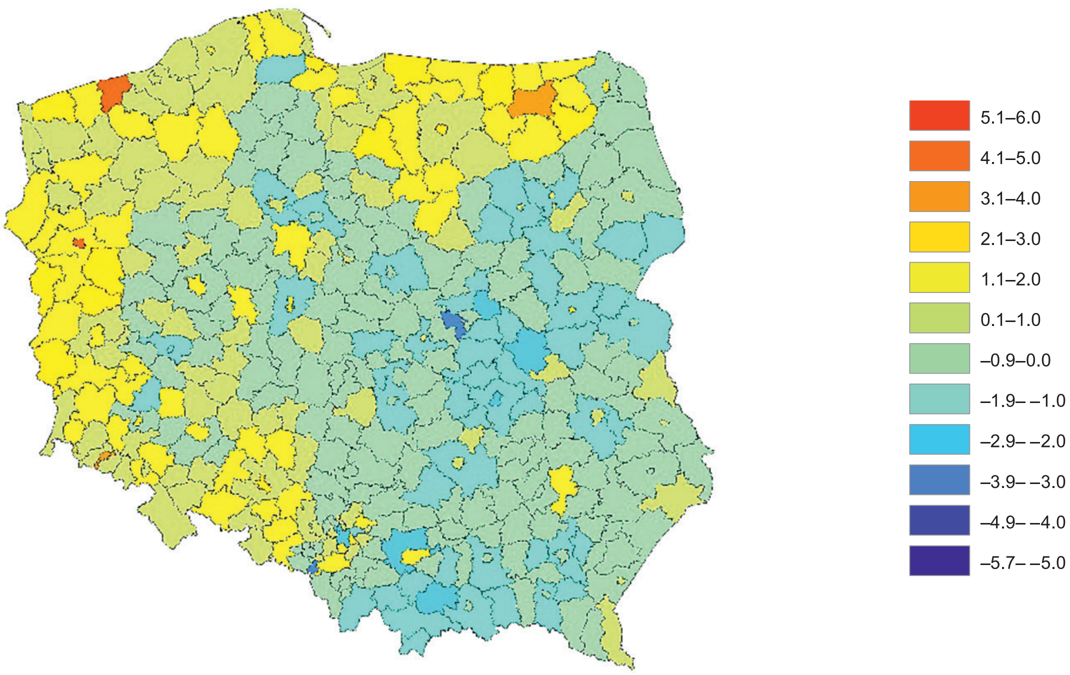

Fig. 6. Random effects by district resulting from estimated models for dependent variable NB with the use of Swamy-Arora transformations.

Source: own research.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

The spatial distribution of random effects in both transformations is similar. Negative random effects are concentrated mainly in the eastern and south-eastern part of Poland. This means that there may be additional factors, not included in the model, influencing market activity in these areas. The spatial distribution of random effects also indicates a fairly clear spatial autocorrelation and the presence of clusters dominated by positive (yellow in Figures 5 and 6) and negative values (blue in Figures 5 and 6).

CONCLUSIONS

The panel model employed in the analysis of cross-sectional time series data encompassing 380 spatial units (districts), in the context of 2 dependent variables and 10 independent variables, in a 11-year time series showed the utility of this kind of model for analysing the relationship between real estate prices and economic, social and spatial conditions. The process identified the set of variables which significantly impact both the housing unit price and the number of real estate transactions. Judging by the quality of the models produced, the present study indicates that all adopted variables turned out to be statistically significant with a significance level of less than 0.01. This means that the adopted set of variables characterizing socio-demographic, economic and environmental conditions to a large extent explains both price formation processes and the activity of the housing market. The variable that has the strongest impact on average prices and the number of transactions is the average monthly salary. It is also worth noting the significant impact of air quality expressed by the variable characterizing the emission of pollutants. Spatial distribution of random effects determined on the basis of average price models indicates medium spatial dependence. In models using the number of transactions as a variable explained by the spatial distribution of random effects points to clear clusters with low and high values.

The occurrence of spatial autocorrelation also indicates the possibility of continuing the research with the use of panel models taking into account spatial dependencies (spatial panel models). Although the presented models largely explain the studied phenomena, it should be noted that the dependent variables may also be affected by other factors not included in the model.

REFERENCES

Adams, Z., Füss, R. (2010). Macroeconomic Determinants of International Housing Markets. *Journal of Housing Economics*, 19(1), pp. 38–50. doi:10.1016/j.jhe.2009.10.005.

Attanasio, O.P., Bottazzi, R., Low, H.W., Nesheim, L., Wakefield, M. (2012). Modelling the Demand for Housing over the Life Cycle. *Review of Economic Dynamic*, 15(1), pp. 1–18. doi:10.1016/j.red.2011.09.001.

Baltagi, B. (2008). *Econometric Analysis of Panel Data*. John Wiley & Sons.

Belke, A., Keil, J. (2018). Fundamental Determinants of Real Estate Prices: A Panel Study of German Regions. *International Advances in Economic Research*, 24, pp. 25–45. doi:10.1007/s11294-018-9671-2.

Beltratti, A., Morana, C. (2010). International House Prices and Macroeconomic Fluctuations. *Journal of Banking & Finance*, 34(3), pp. 533–545. doi:10.1016/j.jbankfin.2009.08.020.

Brunnermeier, M.K., Julliard, C. (2008). Money Illusion and Housing Frenzies. *The Review of Financial Studies*, 21(1), pp. 135–180. doi:10.1093/rfs/hhm043.

Eichholtz, P., Lindenthal, T. (2014). Demographics, Human Capital, and the Demand for *Housing. Journal of Housing Economics*, 26, pp. 19–32. doi:10.1016/j.jhe.2014.06.002.

Englund, P., Ioannides, Y.M. (1997). House Price Dynamics: An International Empirical Perspective. *Journal of Housing Economics*, 6(2), pp. 119–136.

Griliches, Z., Intriligator, M.D. (2007). *Handbook of Econometrics*. Elsevier.

Hendershott, P.H., Abraham, J.M. (1992). Patterns and Determinants of Metropolitan House Prices, 1977–91. National Bureau of Economic Research. doi:10.3386/w4196.

Himmelberg, C., Mayer, C., Sinai, T. (2005). Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. *Journal of Economic Perspectives*, 19(4), pp. 67–92. doi:10.1257/089533005775196769.

Hsiao, C. (2003). *Analysis of Panel Data*. Cambridge, Cambridge University Press.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl

Iacoviello, M., Minetti, R. (2003). Financial Liberalization and the Sensitivity of House Prices to Monetary Policy: Theory and Evidence. *The Manchester School*, 71(1), pp. 20–34. doi:10.1111/1467-9957.00332.

Lastrapes, W.D. (2002). The Real Price of Housing and Money Supply Shocks: Time Series Evidence and Theoretical Simulations. *Journal of Housing Economics*, 11(1), pp. 40–74. doi:10.1006/jhec.2002.0309.

Liu, F., Liu, C., Ren, H. (2018). Urban Housing Price Fluctuations and Regional Systemic Financial Risks: Panel Spatial Economic Models in Jiangsu, China. *Sustainability*, 10(10), 3452. doi:10.3390/su10103452.

Meidani, A.A.N., Zabihi, M., Ashena, M. (2011). House Prices, Economic Output, and Inflation Interactions in Iran. *Research in Applied Economics*, 3(1), pp. 1–13, doi:10.5296/rae.v3i1.664.

Nerlove, M. (1971). Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections. *Econometrica, Econometric Society*, 39(2), pp. 359–382.

Nguyen, T.T., Wang, K. (2010). Causality between Housing Returns, Inflation and Economic Growth with Endogenous Breaks. *Journal of Chinese Economic and Business Studies*, 8(1), pp. 95–115. doi:10.1080/ 14765280903488397.

Panagiotidis, T., Printzis, P. (2016). On the Macroeconomic Determinants of the Housing Market in Greece: A VECM Approach. *International Economics and Economic Policy*, 13, pp. 387–409. doi:10.1007/s10368-016-0345-3.

Swamy, P., Arora, S.S. (1972). The Exact Finite Sample Properties of the Estimators of Coefficients in the Error Components Regression Models. *Econometrica, Econometric Society*, 40(2), pp. 261–275.

The Econometrics of Panel Data: Handbook of Theory and Applications. (2013). L., Mátyás, P., Sevestre. (eds.). Springer Science & Business Media.

Tsatsaronis, K., Zhu, H. (2004). What Drives Housing Price Dynamics: Cross-Country Evidence. BIS Quarterly Review, Bank for International Settlements, March.

Urbanavičienė, V., Kaklauskas, A., Zavadskas, E.K. (2009). The Conceptual Model of Construction and Real Estate Negotiation. *International Journal of Strategic Property Management*, 13(1), pp. 53–70. doi:10.3846/1648-715X.2009.13.53-70.

Wang, Y., Wang, S., Li, G., Zhang, H., Jin, L., Su, Y., Wu, K. (2017). Identifying the Determinants of Housing Prices in China Using Spatial Regression and the Geographical Detector Technique. *Applied Geography*, 79, pp. 26–36. doi:10.1016/j.apgeog.2016.12.003.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge–London, MIT Press.

Żelazowski, K. (2011). Regionalne zróżnicowanie cen i ich determinant na rynku mieszkaniowym w Polsce (Regional Differences In Prices and their Determinants on Polish Housing Market). *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 19(3), pp. 98–106.

✉rcellmer@uwm.edu.pl, ✉aneta.cichulska@uwm.edu.pl, ✉miroslaw.belej@uwm.edu.pl