A MODEL OF INTERNAL MIGRATION: AN EXTENDED NEO-CLASSICAL MIGRATION MODEL AND EVALUATION OF REGIONAL MIGRATION DETERMINANTS IN POLAND

Pavlína Hejduková¹, Lucie Kureková²

¹ University of West Bohemia, Faculty of Economics, Department of Finance and Accounting, Czech Republic, ORCID: 0000-0003-3387-1198, pahejdu@kfu.zcu.cz;
² University of Economics, Prague, Faculty of Economics, Department of Economics, Czech Republic, ORCID: 0000-0002-7611-0463, xkurl06@vse.cz.

Abstract: The goal of this paper was to carry out an empirical verification of the theoretical model of internal migration and evaluate the influence of selected determinants on internal migration in Poland. In order to achieve this goal, an analysis of internal migration in Poland and an econometric analysis were carried out. The analysis was processed using data from 16 voivodeships (i.e. on the NUTS2 level); the analyzed period was from 2003 to 2017. Based on selected theoretical concepts of migration and present studies on internal migration, the theoretical model of internal migration was specified. A total of five models were estimated; first, four models were estimated for all voivodeships and subsequently a GLS estimate was carried out with the so-called “strong region”. Results of the econometric analysis show that regional wage differences together with differences in the unemployment rate determined internal migration, which is in agreement with the assumptions of neo-classical economy. The positive relationship was indicated between the degree of migration and its delayed variable last year, which points simultaneously to the validity of the assumption of the theory of cumulative causation. The test of the dummy variable has shown that the degree of regional migration increased in the period of crisis. From the perspective of push and pull factors, pull factors in the host region were the following determinants: wage, demand for work in the industrial sector of the economy, and foreign human capital; on the contrary, push factors included the increasing unemployment rate, housing costs and the degree of urbanization.

Keywords: Regional migration, model, panel data, econometric analysis, Poland.

JEL Classification: O15, R11, R15, J6, C5.


Introduction
Population migration continues to be a current topic linked to a wide spectrum of various external and internal factors on both international and regional levels. In contemporary literature, there is a whole score of empirical studies that deal with international migration, its determinants and impacts on the economy. However, there are only few empirical studies that deal primarily with solely regional (i.e. internal) migration in comparison to the large number of studies analyzing international migration, which is one of the main reasons for the selection of the topic of this study and its focus on internal migration and thus on movements that take place within one geopolitical entity, usually a nation-state (for more on the definition of internal migration, see, e.g., Fendel, 2014; Royuela &
Ordóñez, 2016). So-called “gravity models” stemming from an analogy to Newton’s law of gravity and Ravenstein’s laws of migration are often used for modelling internal migration and the study of it, or for the analysis of the main determinates that impact these internal fluctuations of citizens; however, these gravity models of migration are often criticized for their insufficient theoretical foundation. The second group of models that can be used to study internal migration are so-called “theory driven models”, which are often based on assumptions and knowledge from neo-classical economy and are often elaborated upon using other knowledge of theoretical concepts such as “push and pull” factors (Amara & Jemmali, 2018; Akarca & Tansel, 2018) or “human capital theory” (Gutiérrez-Portilla, Maza, & Hierro, 2018).

The beginnings of the study of internal migration can be found in the research of Ravenstein (1885), who defined the rules of migration (or “laws of migration”), which were deduced by collecting data on internal migration in England. Through his work, Ravenstein made a contribution to the future development of the research of migration, for example by revealing the inverse relationship between the volume of migration and the distance between initial and target regions. Many other authors then built upon his work. Here we can also name e.g. Zipf (1949), who with the help of a mathematic-statistical system bases his work upon the migration laws defined by Ravenstein and, based on the conclusions of his empirical research, in which he studied the movement of inhabitants between cities, he formulated the hypothesis that the content of migration is indirectly proportionate to the distance which migrating individuals must travel and directly proportionate to the size of the population of the place of their departure and the site of their destination. The conclusions that Zipf drew were thus in accordance with Ravenstein concerning the preference of migration for shorter distances. Although migration laws are categorized as empirical-inductive and are often criticized for their limited validity in space and time, they are still given attention today, specifically and most often concerning the application of gravity models of migration (see, e.g., Hagerstrand, 1957; Grogger & Hanson, 2011; Beine, Docquier, & Özden, 2009; Paleta & Jandová, 2010).

In the neo-classical economic theory of migration, we can observe two approaches to explaining migration: (i) the micro-economic approach and (ii) the macro-economic approach. On the micro-economic level, the neo-classical model of migration assumes that each migrant measures up the costs and benefits of migration, and these costs and benefits differ depending on the intended destination of the migrant (see Borjas, 1989, 1990). Contrary to this is the so-called “new migration theory”, the founders of which are Stark and Bloom (1985), which assumes that households, not individuals themselves, make decisions on migration, and these households do not consider only the conditions of the labor market while making their decisions, but also take into consideration the conditions on other markets. Sjaadstad (1962) claims that the theory of human capital enriched the neo-classical model of migration with so-called socio-demographic characteristics, in which migration is significantly affected by the availability of human capital (e.g. skills, experiences, age, marital status, gender, profession, position on the labor market), preferences, and expectations. Bauer and Zimmermann (1999) assume that these characteristics are significant determinants for migration and claim that the probability of migration decreases with age and commonly increases with the level of acquired education.

The macro-economic model of migration explains migration primarily via geographic differences in supply and demand on the labor market. A territory with a high share of labor in capital will have a relatively low equilibrium wage and vice versa. Migration flows tend to be from less developed areas, where there is a labor surplus and relatively low wages to more developed areas where there is a shortage of labor and relatively high wages. The most basic model appears in the work of Hicks (1932) and also Lewis (1954). In the 1970s, Harris and Todaro (1970) elaborated upon the model with two more important factors – unemployment and time. This elaboration made it possible to explain the significant portion of migration flows from rural areas to urban agglomerations. In Harris and Todaro’s model (1970), migration is determined by expected income differential, i.e. the income differential adjusted for the probability of finding a job. The original neo-classical theory only used differences in
current wages in the decision to migrate. According to neo-classical theory, migration should contribute to the balancing of disparities between individual territories. However, not all theories agree with this assumption and, on the contrary, they claim that, thanks to internal migration and by extension international migration, inequalities may deepen. In addition to the positive effects of migration, Myrdal (1957) describes the operation of negative polarization (or back-wash) effects. One form of the negative effect can be migration from less developed regions to more developed regions. The result of migration is then not a state of general balance, but a greater deepening of disparities.

The goal of this paper is to carry out an empirical verification of the theoretical model of internal migration and evaluate the influence of selected determinants on internal migration in Poland. In order to achieve this goal, an analysis of internal migration in Poland and an econometric analysis were carried out. As was outlined above, it is highly important to keep in mind that there are many types and methods of the research of migration, and that this paper focuses on internal migration and its determinants. The present study is an important extension to the research topic of migration, mainly because the empirical part focuses on testing of the regional migration model (i.e. NUTS2 level), as well as on a detailed description of the procedure for econometric analysis using panel data, and finally, authors have attempted to introduce a new approach of capturing the space factor in the econometric model (pseudo difference). The use of panel data for the analysis of regional migration is a current trend in this area. Methods, estimation and testing options for verifying the suitability of selected procedures and the post-estimation tests for panel data are expanding fast.

In the first section, contemporary studies dealing with internal migration are introduced and subsequently a theoretical model of internal migration is formulated. In the second section, the methodological approach is explained, an economic model is formulated, and data used in the study are described. In the third section, the results of the migration analysis and econometric estimates are presented. In the conclusion of this paper, primary findings are presented in summary, including indications for further possible research in this given area of study.

1. Literature Review and Theoretical Framework
1.1 Contemporary Empirical Studies on the Issue of Internal Migration

In regard to the thematic focus of our paper, the chosen starting points for empirical verification primarily include contemporary empirical studies that deal with internal migration; their conclusions are then compared to the results of the econometric analysis.

According to An et al. (2016, p. 1), “the issue is important to policymakers, especially in middle-income countries with high economic volatility. Planners need to forecast population movements for a range of fiscal and infrastructure provision reasons, and therefore are interested in knowing how, in addition to earnings gaps, wage instability and economic booms and busts affect population movements”. Lehmer and Ludsteck (2011) attempt to count the benefits of interregional migration via expected salaries. Their study points to the fact that young workers have the highest earnings and also that high earnings stem from migration from rural to urban areas. These two studies listed above thus attempt to find an explanation why a sufficiently large degree of migration between regions is not taking place by using an estimate of costs and benefits. Fischer (2019) claims that the willingness of inhabitants to move is influenced by costs of migration and states that estimated average migration costs are in the range of six times the average annual wage, which appears sizable enough to prevent taking advantage of economic opportunities for workers. We can assume from his conclusion that if these costs decrease _ceteris paribus_ over time, we can expect the degree of internal migration to increase. His idea in the form of including costs linked to decisions about migration was also indirectly incorporated into our model of migration. In addition to wage disparities (WAGE), housing costs (HOUSING) and the supply of housing (FLAT) were also included into the econometric equation, as it can be assumed that the costs of housing and the probability of finding housing can influenced internal migration.

It has been proven that both international migration and regional migration can be determined by the probability of finding work in the region. Studies commonly focus on the differences in the regional unemployment rate
among regions or the unemployment rate and their influence on migration flows. For instance, Carlsen et al. (2013) have found a dependence between the differences in the level of unemployment and migration and furthermore have shown that the reactions of migration flows to regional shocks in the unemployment rate grows among population groups with higher education. We can infer from the conclusions drawn by Laamanen’s study (2014) that sector changes in employment are not significant for migration; however, interregional migration is to a certain degree influenced by regional differences in unemployment and available work opportunities for unemployed workers. On the other hand, authors Pop Silaghi and Ghatak (2011), just as Kureková (2015), have not found based on their results that internal migration is influenced by regional differences in the unemployment rate. However, these authors have confirmed the influence of wage differences on regional migration flows. In addition, Pop Silaghi and Ghatak (2011) point out that regions with long-term high salaries appear to be strong regions and internal migration flows are strongly drawn (i.e. concentrated) to these regions. In their words, high salaries are a large pull factor and the influence of other push factors in home regions is rather weak. Their results have led to the inclusion of selected independent variables representing push and pull factors into the model of internal migration. These factors can be represented by the degree of industrialization (INDUS) or the degree of urbanization (URB).

Another possible factor that appears in academic studies focusing on internal migration is the accumulation of human capital in a given region. Clemente et al. (2013) have studied the influence of human capital on internal migration via the size of the work force and its composition. Their studies have produced contradictory results, as in one observed period the regions that experienced a heightened ratio of qualified workers to unqualified workers became less attractive for internal migration; however, in the next period, they noted the opposite tendency in migration flows. Mlzé and Reinkowski (2011), similarly to Maza and Villaverde (2004), did not find differences in human capital to be highly useful for modeling internal migration; nonetheless, other authors have used this variable in their models (see, e.g., Napolitano & Bonasia, 2010; Kureková, 2015). Despite the unsatisfactory empirical results on the potential influence of human capital on internal migration, two variables that could appropriately represent the supply of human capital in a given region have been incorporated into our econometric model. The first variable represents domestic human capital (HC) as a percentage of the population with a university degree in the given region, while the second represents the accumulation of international human capital (FOREIGNERS) via the influx of migration into the region from abroad.

For the modeling of internal migration, it is also possible to consider non-traditional independent variables such as institutional factors (see Čermáková & Jašová, 2019; or Simionescu et al., 2019) or the influence of “intergenerational mobility” (heightening of social status). Kim and Lee (2019) have studied the role of intergenerational mobility via decision made by parents on their own migration and the migration of their children. Parents make decisions based on the future income for their children and are motivated to move to areas in which they will achieve the highest income and thus have a higher probability of reaching higher social status. The authors of the study have found that highly educated families with children of school age have a tendency to move to areas that show growing intergenerational mobility. Similar behaviour as in the case of parents can be observed also for students, the students’ emigration has probably a similar background (see Mishchuk et al., 2019). Such models of migration, however, are highly demanding in terms of gathering microeconomic data, and therefore the results of this study could not be directly compared to results of the analytical section of this paper. Nonetheless, the conclusions of this aforementioned study support the assumption of migration to regions with higher salaries or a higher supply of human capital (measured as the proportion of the population with university education).

A common focus among studies is the analysis of migration from rural (agricultural) areas to urban (industrial) areas, which is also sometimes called migration from less developed to more developed regions. In highly simplified terms, it is possible to claim that urbanized regions commonly have higher average salaries and a higher probability of finding work, and therefore migration towards these
regions can be studied. Christiaans (2017) has modeled migration from the countryside to industrial regions, and his conclusions suggest that in the future the gradual depopulation of the countryside will continue to take place, and therefore these conclusions strengthen pessimistic expectations concerning the rural exodus. In the context of urbanization, agglomeration effects that can influence the degree of internal migration are also mentioned. The results of Mitze and Schmidt (2015) support the opinion that agglomeration economies are in fact the key driving force for internal migration flows. Therefore, the measure of the degree of urbanization (URB) has been included in our econometric model to make it possible to study its influence on the degree of migration.

There is proof that international migration can be determined by heightening insecurity in the host country; in other words, migration flows can differ depending on an economic cycle (Beets & Willekens, 2009; Caro et al., 2018). Therefore, authors An et al. (2017) recommend exploring internal migration also in the context of economic development, as their results show that crisis has weakened the influence of wage differences and thus reduced the volume of internal migration. In light of these facts, estimates of the econometric model were carried out with a dummy variable representing an economic recession (CRISIS). This added dummy variable makes it possible to test whether internal migration in Poland has been influenced by the global economic crisis that began in 2008.

Some studies have attempted to predict the development of internal migration in the context of climate change, which is more typical of studies from the USA or other states that are relatively large in area (Hornbeck, 2012; Smith et al., 2006; Fan et al., 2016). The studies have predicted that climate change will lead to a redistribution of the population as people decide to move to regions that are less predisposed to extreme climates. In terms of their focus, these studies most likely exceed the thematic focus of our analysis; nonetheless, interesting conclusions are offered in studies by authors Fan et al. (2018), who claim that previous studies have ignored or insufficiently taken into account the fact that migration will be suppressed by changes in salaries and prices in housing as a consequence of migration. The conclusions of their study strengthen the significance of the variables such as the price of housing (HOUSING) that have been included in our econometric model.

1.2 Theoretical Model of Internal Migration

The theoretical model of internal migration has been specified by the authors of this paper based on an extended neo-classical migration model containing an independent variable representing the wage and employee differences between regions and contains other independent variables that can be considered to be “push and pull” factors; the factor of space or distance, which stems from the formulation of Ravenstein’s migration laws, is also incorporated into the model.

The general form of the equation for the theoretical model of internal migration can be expressed in the following:

$$MIG = f (W; π; PnPF; dist)$$

1.6.2020   16:39:23
the emigration of a population from a region will increase.

The symbol $P_{nPF}$ represents so-called push and pull factors – according to their character, these factors can attract individuals to a certain region or, on the contrary, push them from one. A traditional factor that appears in migration models is the degree of industrialization, which can make a region more attractive for potential migration, primarily in terms of an enduring demand for work in the industrial sector of the economy. We can assume that the influx of new populations into a region will grow along with the growing degree of the industrialization of a given region. On the contrary, we can also observe the opposite tendency of migration from industrial areas, and this tendency can be explained by two possible influences. The first possible influence can be the transformation of the economy in the sense of an increase in the significance of the tertiary sector compared to the secondary sector. The second influence can be the state of the environment in a given region, as the degree of industrialization can indirectly indicate the state of the environment, as the growth in the degree of industrialization is often linked to a relative deterioration in the state of the environment. This can cause migration of individuals from a given region, and such migration can be labeled "amenity migration". The factors that affect amenity migration can be divided into three groups: (i) socio-economic characteristics, (ii) natural factors and (iii) overall economic level (Kuentzel, 2005; Novotná et al., 2013; Mishchuk & Grishnova, 2015).

Another factor that can affect migration is the human capital in a given region. We commonly assume a direct dependence between the degree of migration and this factor. The accumulation of human capital is understood as a key determinant for the creation of innovations and additional socio-economic development of a region, and therefore migrants are likely to prefer areas with a higher concentration of human capital. On the other hand, the concentration of human capital in the form of the growing supply of a qualified work force heightens competition on the labor market and qualification standards for job positions may grow, a fact that for some potential migrants may represent added expenses for their own human capital and thus their willingness to migrate will decrease or they will prefer regions with a lower concentration of human capital.

Furthermore, the theoretical model includes the supply of a foreign work force in the given region. This factor can affect the degree of migration similarly to human capital; a growing supply of a foreign work force can increase the degree of migration, as the accumulation of foreign human capital occurs, which can be understood as a determinant of economic growth or, on the contrary, the domestic work force may be pushed out by the foreign work force. The influx of foreign work supply is linked to the segmentation of the labor market and the creation of ethnic groups.

Availability and housing costs are other factors that determine the degree of migration. We can assume that the growing supply of housing will increase the interest in a given region and, on the contrary, the increase in the price of housing will decrease the degree of immigration. Another important factor that the theoretical model takes into consideration is the degree of urbanization of a given territory. Similarly as in the case of the degree of industrialization, this factor can be assumed to have a direct relationship to the degree of migration, as urban regions can attract migration via an enduring demand for work and, on the contrary, we can observe the departure of populations via amenity migration.

The last but equally important push and pull factor is economic crisis, which can affect migration via changes in the supply of and demand for labor and in the probability of finding work. It is likely that, in a time of crisis, migration may increase thanks to the loss of work and the search for new employment. On the other hand, the decrease of net migration may decrease owing to thanks to the decrease of the probability of finding work; the resulting effect of the economic crisis variable can cause changes in net migration in both directions.

According to Ravenstein’s laws, the degree of migration decreases with increasing distance between destinations. For this factor of distance, we will thus assume that the degree of migration decreases with increasing distance, and so the factor of space or distance has been incorporated into the model. Spatial difference together with the assumption of balanced regional development of the inflation rate allows us at least to a certain degree to solve the problem of nominal and real values. For the sake of simplification, we assume that the regional inflation rate develops equally in all
Economics

regions, or it is the same on a national level, in which case spatial differences of nominal quantities should correspond to the differences of real quantities. This assumption had to be instated primarily because data on the development of price growth on a regional level was not available, although it is quite unlikely that the inflation rate always develops in all regions in the same manner, which is evident in the differentiated growth of housing costs among individual regions.

2. Methodology and Data
2.1 Specifications of the Econometric Model

Based on findings of contemporary empirical studies (Ch. 1.1) from the theoretical migration model (Ch. 1.2), the econometric model of internal migration in Poland was specified. The aim of this paper is to carry out an empirical verification of the theoretical model of internal migration and evaluate the influence of selected determinants on internal migration in Poland. In order to reach this goal, it is necessary to analyze internal migration in Poland and carry out an econometric analysis.

Before the estimate of the econometric model of migration itself, it was first necessary to carry out an analysis of internal migration and foreign migration in Poland. Migration was depicted via cartograms and modified Webb diagrams, using net internal and foreign migration. The depiction of internal migration via an identity matrix of migration efficiency ratio and node diagrams cannot be carried out, as information on the flows of populations between individual regions is missing.

As has already been mentioned, the creation and application of the model of internal migration stems not only from the formulated theoretical model of internal migration, but also from contemporary empirical studies, the closest of which is a study by Kureková (2015), in which the internal model of migration for the Czech Republic was estimated. Significant elaborations carried in this study include new variables, which have been incorporated into the model of internal migration in Poland, and the new method of computing independent variables, thanks to which the factor of space, called a “pseudo_diff”, is implemented in the econometric model (see below for an explanation). The final specification of the econometric model for Poland’s internal migration is as follows:

\[
MIG_n = \alpha + \beta_1 MIG_{n-1} + \beta_2 WAGE_n + \beta_3 UNEMPL_n + \\
\beta_4 INDUS_n + \beta_5 FOREIGNERS_n + \beta_6 HC_n + \\
\beta_7 FLAT_n + \beta_8 HOUSING_n + \beta_9 URB_n + \\
\beta_{10} CRISIS_n + \mu_n
\]  

(2)

The data gained are on the NUTS2 level; in our dataset there is a total of 16 voivodeships, index \( r \) takes values from 1 to 16, the studied period is 15 years long and between the years 2003 and 2017, and the values of the index thus take on a maximum value of 15.

The dependent variable of migration \( MIG_n \) expresses the net regional migration per 1,000 population for the given cross-section unit \( r \) in time \( t \), if immigration is higher than emigration in the given region \( r \), the variable takes on values of \( >0 \) and vice versa; if the influx and departure of individuals was more balanced, this indicator would take on a value close to 0. The selection of this statistical indicator has been determined by the availability of data from the Polish Central Statistical Office.

In secondary data taken from the Polish Central Statistical Office, it is not possible to directly observe the direction of migration among regions, and therefore a method that could suitably approximate the spatial difference between regions has been proposed: for the purposes of this study, the pseudo-spatial difference will be labeled “pseudo_diff”, and can be mathematically expressed by the following formula:

\[
pseudo\_diff_{fr} = \left( \frac{v_r}{v_t} - 1 \right)
\]  

(3)

The symbol \( v \) represents the given cardinal variable, \( r \) is the index for the given voivodeship, \( s \) is the index for the state (national) level and \( t \) is the studied year. We see that by using the relationship from the equation, we combine time series with the use of their ratios; these modifications are fundamental ones for working with time lines and help solve problems with the non-stationarity of the time lines.

STATA statistical software was used for the estimate of the econometric model of Poland’s internal migration. In estimating the econometric model of internal migration, a classic procedure of estimates and tests was carried out in order to find out whether the FE or RE model was more suitable. First the suitability of the RE model was tested; based on an LM test (P-value = 0.0000), the RE model can be used; then tests were
carried to find whether use of the FE model was logical. According to a parametric test (P-value = 0.0000), it is possible to use the FE model, and therefore the Hausman test was used, which helped to determine the selection of either the FE or RE model. The results of the Hausman test (P-value = 0.0000) pointed towards a higher suitability of the FE model; nonetheless, diagnostic tests pointed to the fact that the estimate is burdened by autocorrelation and cross-sectional correlation, and therefore a series of GLS estimates were carried out in order to decrease the influences of autocorrelation and cross-sectional correlation.

2.2 Data and Statistical Indicators
Data were taken from the database of the Polish Central Statistical Office; in regard to the size of Poland and its territorial division, data on the NUTS2 level, i.e. on the level of voivodeships, were selected for the analysis. There is a total of 16 voivodeships in Poland (N = 16), the majority of time lines for individual variables were available at least from 2003 and the most recent available data were for the year 2017 (T = 15). The final panel data set contains 240 observations (NxT) for each variable.

In regard to the size of individual voivodeships, it can be assumed that the selected level of territorial units will not have a negative effect on the predicative ability of the results.

In the case of the variable expressing the unemployment rate (UNEMPL), the time line from 2004 was available; for the variable representing migration (MIG) and foreign migration (FOREIGNERS), data for 2015 were missing. So as not to have to shorten the researched time period and reduce the information for variables with missing values, values for 2003 were estimated using OLS in the context of unemployment (UNEMPL) and for 2015 in the case of variables of migration (MIG and FOREIGNERS).

As was mentioned previously, the variable (MIG) expresses net internal migration per 1,000 population. (MIG_LAG) is then the lagged variable of (MIG) in t-1; this control lagged variable has been incorporated into the model in order to follow whether the degree of migration is determined by the historical development of migration in a given region, while the second reason is a technical one – the delayed variable partially helps us with the problem of autocorrelation.

For the variable (UNEMPL), the Registered unemployment rate indicator was selected, which was available from 2004. In addition, data on the prices of apartment per m² were not available, and therefore the indicator of Average monthly expenditures per capita – housing, water, electricity, gas and other fuels was selected; for the sake of simplicity, we will call these expenses Housing expenses (HOUSING).

The variable (FOREIGNERS) enters into the econometric model as net foreign migration. The degree of industrialization (INDUS) is expressed as the percentage of employees working in the industrial sector. Human capital (HC) is expressed as a percentage of employees with tertiary education. The degree of urbanization (URB) represents the ration of urbanized towns in the whole number; this variable was again recoded to an ordinal variable, in which ratio values up to 0.1 take on a value of 0 and values between 0.1 and 0.2 equal 1 and values higher than 0.2 take on a value of 2. The last explanatory variable in model is the economic crisis (CRISIS), which was defined as a binary variable. In 2008–2011 the value is 1, in other years it is 0.

3. Results
3.1 Analysis of Internal and Foreign Migration in Poland
First internal migration from the perspective of net migration was dealt with, i.e. regions that have a positive/negative net migration were identified using cartograms. Fig. 1 contains three cartograms in which the average overall net migration, the dominance of the Masovia Voivodeship (MZ) is evident. This dominance is visible also in the case of the division of migration into internal and foreign. In terms of net foreign migration in Poland, it is evident that almost all voivodeships (except for voivodeship MZ) have a negative average net foreign migration, i.e. foreign migration at bottom right.

According to the values of the average overall net migration, the dominance of the Masovia Voivodeship (MZ) is evident. This dominance is visible also in the case of the division of migration into internal and foreign. In terms of net foreign migration in Poland, it is evident that almost all voivodeships (except for voivodeship MZ) have a negative average net foreign migration, i.e. foreign migration had a negative effect in the majority of voivodeships on changes in population volume.
In the voivodeships MAŁOPOLSKIE (MP), WIELKOPOLSKIE (WP) and POMORSKIE (PM), the average net foreign migration in the absolute value is lower than net internal migration – for this reason, we can observe positive values of the overall net migration in the cartogram on the right for the aforementioned regions.

An overview of the flows of foreign migration according to individual states is listed in the following Fig. 2; on the left are countries from which the most citizens flow into Poland and on the right are countries where inhabitants of Poland most commonly migrate to. For a better comparison of the development of migration flows, their average was once again calculated for the 10-year period from 2006 to 2015 and the states are ordered according to values achieved in 2016. It is evident that between the years 2006 and 2015, the most migrants came to Poland from Ukraine (avg: 30%) and then, with a large gap, from Germany (avg: 6%) and Belarus (avg: 6%). If the influx of foreign migration is compared according to 2016, it is evident that Ukrainians held the dominant portion of immigration, then with a large gap China (4%) and Belarus (3%). The most desired countries for the emigration of the inhabitants of Poland is Germany (avg: 55%; 2016: 61%) and then with a large gap Great Britain (avg: 15%, 2016: 11%) and the Netherlands (avg: 6%, 2016: 9%).

The Webb diagram offers a comparison of net internal and foreign migration and their development over time. A total of four Webb diagrams have been created; first for the whole analyzed period of 2003–2017, then for equally long periods: 2003–2005; 2006–2011 and 2012–2017. The results are presented in the following Fig. 3.

The majority of voivodeships was located in the sector where a drop in the population due to internal migration was higher than the drop due to foreign migration (voivodeships: OP, WM, LB, SK, PL, PK, ZP, KP, ZP, LS and LD). From 2003 to 2007, only four voivodeships had positive internal migration (i.e. MZ, MP, WP and PM); in the following two periods, voivodeship DS joined the group; nonetheless, out of all five of these voivodeships, only the MZ voivodeship
Fig. 2: Overview of foreign migration flows in Poland (TOP 10, %)

Source: own based on OECD (2018)

Fig. 3: Webb diagrams

Source: own based on OECD (2018)
Economics

had a positive net foreign migration in all three observed periods; in the years 2008–2012, voivodeship MP also had positive foreign migration. Voivodeship DS saw a significant shift between sectors in time, when in 2003–2007 it was located in sector G; in the next period it shifted to sector E, and in the final period we can see a shift to sector D. In the case of voivodeship OP can we observe a decrease in the negative net foreign migration. A significant decrease in the negative net internal migration also took place; nonetheless, the given voivodeship is still located in sector G together with 10 other voivodeships. In voivodeship SK, we can see that the foreign balance shifted slightly into positive values in 2013–2017.

3.2 Basic Descriptive Statistics of Variables

Before the actual estimate of the econometric model, a basic statistical description of variables was carried out; a graphic depiction of the development of individual statistical indicators is depicted in Fig. 4. In terms of the statistical indicator of gross monthly wage, we can observe a permanent growth over the course of the observed period; nominal wages grew in all voivodeships. The highest average wage was in voivodeship MZ (4.2 thousand PLN); from 2000 to 2017 it grew by almost 1.5 thousand PLN (from 3 to 5.5 thousand). It surpassed the border of 3 thousand PLN in nine voivodeships: LD (3 thousand PLN), MP (3.2 thousand CZK), SL (3.5 thousand PLN), LB (3 thousand PLN), WP (3.1 thousand PLN), ZP (3.1 thousand PLN), DS (3.4 thousand PLN), OP (3.1 thousand PLN) and PM (3.4 thousand PLN). The lowest average wage was in voivodeships KP and LS (2.9 thousand PLN). The highest percentage growth was recorded in voivodeship DS, where the average wage grew by 108% (from 2.2 to 4.7 thousand PLN). Voivodeship DS is among regions which have observed a positive net foreign migration in previous years. The growth of nominal wages by more than 100% was seen in three voivodeships: MP, LD and LB.

At the beginning of the observed period, Polish regions had a relatively high degree of registered unemployment, when in 2004 in all voivodeships the degree of registered unemployment reached its maximum. Up to 2008, we can observe an annual decrease in the unemployment rate; in 2009 it grew sharply, but it was far from reaching values such as in 2004, e.g. in voivodeship WM, when in 2004 unemployment was the highest of all voivodeships and reached 29.2%; in 2009, it grew from 16.8% to 20.7%. In all regions, there was a sharp growth in unemployment in 2009, but from 2013 unemployment decreases and inn almost all regions (aside from WM), it reaches values lower than 10%. The lowest values were reached in voivodeship WP (3.7%).

The highest number of flats appeared in voivodeship MZ (an average of 4.5 per thousand inhabitants annually), and also in voivodeship DS (an average of 4 per thousand inhabitants annually). In 13 regions, supply reached its maximum in 2013; in 9 voivodeships it reached its maximum in 2017. The smallest number of flats appeared in voivodeship PK (an average of 1.9 per thousand inhabitants annually); nonetheless, this region recorded the highest percentage growth from 2003 to 2017, by a whole 325%.

Between the years 2003 and 2017, costs of housing grew in all regions; the development between individual regions varies among regions. The highest average costs were in MZ (216 PLN) and SL (216 PLN); voivodeship DS (201 PLN) also surpassed the value of 200 PLN in average costs of housing. Year-to-year fluctuations take place in all regions, but it is not possible to say with certainty that these fluctuations are caused a priori by the global economic crisis; this can also be caused by the fact that the indicator is created not only from rent prices, but also contains other components such as payments for water, electricity, etc. and these elements may have different pricing.

In terms of the development of the (INDUS) indicator, the most industrialized voivodeships are SL, WP, DS and OP; contrarily, the least industrialized were voivodeships MZ, LB and PL. According to the HC indicator, voivodeships with the highest concentration of HC are: LS, WP and PL; on the contrary, voivodeships with the lowest concentration of human capital are: SK and ZP. Regions with the greatest degree of urbanization are SL and PL (URB = 2); the least urbanized areas (URB = 0) are six voivodeships: MP, LB, SK, WP, ZP and OP.

In addition, before carrying out an econometric estimate, linear dependency was calculated between the degree of migration and absolute values of selected statistical indicators, i.e. statistic indicators do not yet take
into consideration the pseudo-spatial difference between individual voivodeships; the calculated association serves to broaden the information on the development of statistical indicators in time between one another. The calculations of Pearson’s correlation coefficient together with P-value are presented in Tab. 1.

All calculated values of correlation coefficients proved to be statistically significant (5% level); in terms of the strength of dependencies, we have found the relatively strong linear dependency between variables WAGE–FLAT ($p_{wcorr} = 0.7018$). Furthermore, a strong dependence can be identified between the

---

**Source:** own based on Polish Central Statistical Office (2019)

**Note:** (1) LD: ŁÓDZKIE; (2) MZ: MAZOWIECKIE; (3) MP: MAŁOPOLSKIE; (4) SL: ŚLĄSKIE; (5) LB: LUBELSKIE; (6) PK: PODKARPACKIE; (7) PL: PODLASKIE; (8) SK: ŚWIĘTOKRZYSKIE; (9) LS: LUBUSKIE; (10) WP: WIELKOPOLSKIE; (11) ZP: ZACHODNIOPOMORSKIE; (12) DS: DOLNOŚLĄSKIE; (13) OP: OPOLSKIE; (14) KP: KUJAWSKO-POMORSKIE; (15) PM: POMORSKIE; (16) WM: WARMiŃSKO-MAZuRSKIE.
WAGE–HOUSING variables (pwcorr = 0.8886); there is also a relatively strong dependence between FLAT–HOUSING (pwcorr = 0.7332). This indicates that voivodeships with relatively high wages also have a relatively fast-growing supply of housing. Costs of housing grow along with this supply, and thus it is possible to assume that the price of flats in m² has also risen.

3.3 Estimate of the Econometric Model of Internal Migration in Poland

Before the estimate itself, it was necessary to carry out important steps that included finding and measuring the significance of multicollinearity and the testing of the stationarity of time lines. The presence of multicollinearity was first indicated with the help of a correlation matrix; furthermore, the measurement of the significance of multicollinearity was carried out via the VIF method. For all associations, a correlation coefficient was found in an absolute value lower than 0.55, which indicates the presence of weak or almost non-existent multicollinearity between independent variables, which is confirmed also by VIF < 10 results. The highest VIF did not exceed value 4, so it can be assumed that the model does not suffer from multicollinearity.

Subsequently, tests for the presence of a unit root in time lines were carried out; in regard to the relatively low N, the application of Levin Lin Chu test (LLC test) was sufficient. It was found that, in regard to the HC variable, the zero hypothesis of the absence of a unit root of a time line cannot be refuted on the 5% level; it can be assumed that the time series are not stationary, and therefore the transformation of values via a logarithm was selected. Because the pseudo_diff for takes on negative values, the transformation was carried out in the following manner:

\[
\text{if } HC_r \neq \min[HC] \rightarrow logHC_r = \log(HC_r + \text{abs}(\min[HC]))
\]

\[
\text{if } HC_r = \min[HC] \rightarrow logHC_r = \min[HC] + \min[\log(HC_r)]
\]

After the logarithmic transformation of variable HC to InHC, the stationarity of time lines were tested again. Results show that the logarithmic transformation helped solve the problem of the existence of the unit root and now we can assume that all time lines are stationary.

The results of GLS estimates are presented in Tab. 2 (Model 1–5). The analysis of Poland’s internal migration has shown that in some voivodeships in periods 2008–2012 and 2013–2017, changes took place in net internal and foreign migration (SK, DS; see Webb diagrams), which may have been caused by the economic crisis. Testing the impact of the economic cycle on migration was carried out using a test of the significance of adding
a dummy variable. A dummy variable was added to the econometric model (CRISIS), where this variable took on a value of 1 for the period of recession (2008–2011); in other years, it took on the value of 0. Furthermore, the analysis of internal migration pointed to one more powerful voivodeship, MZ; therefore, an estimate was carried out in which this region was omitted (Tab. 2 – Model 5). The reason was to verify whether the model provides consistent estimates of regression coefficients. Results of the series of GLS estimates are listed in Tab. 2.

According to the statistical significance of regression coefficients, Model 4 proves to be the most suitable; nonetheless, in terms of the direction of dependence (signs of regression coefficients), Models 1–4 provide coherent results for the majority of coefficients apart from the InHC variable. After omitting relationships of the strong region (MZ), results of estimated coefficients (signs and statistical significance) are in agreement with Model 4, i.e. the estimates can be considered to be consistent.

The first estimated regression coefficient for the independent variable MIG_lag turns out positive; migration in past years heightened present migration MIG. Estimates of coefficients of key variables WAGE and UNEMPL are in agreement with the assumptions of the neoclassical theory. In the case of the degree of industrialization INDUS, the coefficient turns out positive; the growing degree of industrialization had a positive effect on MIG, as inhabitants prefer to move to industrial voivodeships. The estimated coefficients for foreign migration FOREIGNERS turns out positive for all Models 1–5 and are statistically significant (Models 4–5). Thus, it can be assumed that the supply of the foreign labor force heightened MIG, as the accumulation of foreign human capital took place, which can be understood as a determinant of economic development, which attracts local workers from other regions. The influence of the accumulation of human capital in the form of a qualified local work force InHC is also positive; nonetheless, it is statistically insignificant. In contrast to our expectations, the growing supply of flats (FLAT) had a negative effect on MIG; this finding is in agreement with the estimated negative sign

---

**Tab. 2: Results of the estimates of the econometric model**

<table>
<thead>
<tr>
<th></th>
<th>Model 1 GLS PANELS</th>
<th>Model 2 PANELS_AR1</th>
<th>Model 3 GLS PANELS_PSAR1</th>
<th>Model 4 GLS corr_PSAR1</th>
<th>Model 5 GLS corr_PSAR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>0.214</td>
<td>0.225</td>
<td>0.291**</td>
<td>0.312***</td>
<td>0.297***</td>
</tr>
<tr>
<td>MIG_lag</td>
<td>0.757***</td>
<td>0.731***</td>
<td>0.604***</td>
<td>0.588***</td>
<td>0.582***</td>
</tr>
<tr>
<td>WAGE</td>
<td>2.487***</td>
<td>2.714***</td>
<td>4.044***</td>
<td>3.968***</td>
<td>3.952***</td>
</tr>
<tr>
<td>UNEMPL</td>
<td>-0.155</td>
<td>-0.182</td>
<td>-0.312**</td>
<td>-0.396***</td>
<td>-0.401***</td>
</tr>
<tr>
<td>INDUS</td>
<td>0.998***</td>
<td>1.078***</td>
<td>1.429***</td>
<td>1.411***</td>
<td>1.450***</td>
</tr>
<tr>
<td>FOREIGNERS</td>
<td>0.563***</td>
<td>0.614***</td>
<td>0.648***</td>
<td>0.600***</td>
<td>0.603***</td>
</tr>
<tr>
<td>InHC</td>
<td>-0.014</td>
<td>-0.025</td>
<td>-0.017</td>
<td>0.115</td>
<td>0.020</td>
</tr>
<tr>
<td>FLAT</td>
<td>-0.007</td>
<td>-0.008</td>
<td>-0.016</td>
<td>0.023</td>
<td>-0.014***</td>
</tr>
<tr>
<td>HOUSING</td>
<td>-0.205</td>
<td>-0.198</td>
<td>-0.215</td>
<td>0.288</td>
<td>-0.225***</td>
</tr>
<tr>
<td>URB</td>
<td>-0.038</td>
<td>-0.042</td>
<td>-0.048*</td>
<td>0.029</td>
<td>-0.028**</td>
</tr>
<tr>
<td>CRISIS</td>
<td>0.036</td>
<td>0.042</td>
<td>0.047</td>
<td>0.036</td>
<td>0.092***</td>
</tr>
<tr>
<td>N</td>
<td>224</td>
<td>224</td>
<td>224</td>
<td>224</td>
<td>210</td>
</tr>
<tr>
<td>MZ</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>omitted</td>
</tr>
<tr>
<td>chi2</td>
<td>5,019.26</td>
<td>4,506.17</td>
<td>5,176.89</td>
<td>17,638.91</td>
<td>17,485.336</td>
</tr>
<tr>
<td>P–value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: own

Note: .01 – ***; .05 – **; .1 – *. 
Conclusions

The aim of this paper was to carry out an empirical verification of the theoretical model of internal migration and evaluate the influence of selected determinants on internal migration in Poland. In order to achieve this goal, an analysis of internal migration in Poland and an econometric analysis were carried out. The analysis was carried out on data from 16 voivodeships (i.e. on the NUTS2 level); the analyzed period was from 2003 to 2017.

For the analysis of internal migration in Poland’s voivodeships, a comparison of internal and foreign migration was first carried out using simple cartograms, which showed that from 2003 to 2017, 11 of 16 voivodeships had a negative net internal migration and that almost all voivodeships were seeing a decrease in the local population thanks to foreign migration. The Polish labor market is attractive primarily for migrants from Ukraine; Polish inhabitants primarily leave to other states: Germany, Great Britain, and the Netherlands. In addition, Webb diagrams were used to record the development of net foreign and internal migration; just like the cartograms, these diagrams have pointed out one strong voivodeship, which is the Masovia Voivodeship.

Based on selected theoretical concepts of migration and present studies on internal migration, the theoretical model of internal migration was specified and subsequently an econometric equation was created. Significant attention was paid to the description of the implementation of the factor of space in the econometric model with the help of a so-called “pseudo_diff”. A basic statistical description, measurement of multicollinearity, and tests for the presence of a unit root were carried out. Use of the Hausman test showed that it would be suitable to use the FE model econometric equations for the estimate; nonetheless, diagnostic tests have shown that the estimate was burdened by autocorrelation and cross-sectional correlation, and therefore a series of GLS estimates were carried out, which allowed us to eliminate these discrepancies. A total of five models were estimated, initially four for all voivodeships (Models 1–4); subsequently a GLS estimate without the so-called MZ “strength region” (Model 5) was carried out. According to the statistical significance of regression coefficient, Model 4 seems to be the best. In addition, in terms of dependence (of signs of regression coefficients), Models 1–4 provide coherent results for the majority of coefficients with the exception of the lnHC variable. After omitting relationships of the strength region (MZ), the results of the estimated coefficients are in agreement with the results of Model 4; the estimates can be considered to be consistent.

Results of the econometric analysis show that wage differences determined internal migration, which is also in agreement with contemporary empirical research (see An et al., 2017; Carslen, 2013; Laamanen, 2014). Contrary to studies by authors Pop Silaghi and Ghatak (2011) and Kureková (2015), regional differences in the unemployment rate in Poland proved to be a significant factor of migration in Poland. These results are in line with assumptions of the neo-classical theory. A positive relationship was indicated between the degree of migration and the delayed independent variable expressing the degree of migration in the past year (MIG_lag), which indicates the validity of the assumption of the theory of cumulative causation (Myrdal, 1957); the influx of migration to a region spurs another wave of migration, which would point to the fact that internal migration contributes to additional disparities among regions. In terms of the degree of industrialization, a positive effect on the degree of migration was observed; inhabitants preferred moving to industrial regions, and the same direction of dependence was recorded in the study by Kureková (2015). This direct relationship shows that in regions in which the demand for work in the industrial sector of the economy persists, a heightened degree of migration can expected. According to estimated positive coefficients for the variable expressing foreign migration in a given region, we can assume that the displacement of internal migration (i.e. of local inhabitants) was not taking place; the accumulation of foreign human capital in regions can be understood...
as an indicator of economic development, which attracts local workers to a given region. The study by Mitze and Reinkowski (2014) and Maza and Villaverde (2004) did not consider the human capital variable to be very useful for modeling migration; our results have pointed to the positive influence of the variable of domestic human capital, but this relationship was statistically insignificant, just as in Kureková’s study (2015), which supports conclusions that this variable was not relevant for internal migration in Poland. Contrary to our expectations, the growing supply of flats in Polish voivodeships had a negative influence on the degree of migration; however, this finding is in agreement with the estimated negative sign for the URB variable. Inhabitants of Poland preferred to move to less urbanized regions. It has been shown that the growing costs of housing then had a negative effect on migration, which is in agreement with our expectations that interest in a given region decreases as housing costs grow. The test of the dummy variable showed that the degree of migration increased in the period of crisis. From the perspective of push and pull factors, pull factors in the host region were the following determinants: wage, demand for work in the industrial sector of the economy, and foreign human capital; push factors were the increasing unemployment rate, housing costs and the degree of urbanization. However, our data set about migration seems to be imperfect. In data appears only registered migration, it means that migrants who did not change the registered place of residence are invisible for our statistics. This fact may underestimate the real statistics of regional migration, this could mean that the determinants of migration are even stronger than our results revealed.

Our study has attempted to describe in detail the process of an econometric model estimate including the computation of used data and the testing of assumptions in the case of the analysis of panel data. Such a description may become a guiding point for a potential elaboration upon the research or its replication using data from other states. The model of internal migration could then be applied to so-called “selective migration”, making it possible to observe the impact of selected factors on migration of inhabitants with a specific profession (e.g. health care personnel). Furthermore, presented results and results of wider research in future may have practical implication for policymaking in the field of migration management.

Acknowledgement: This paper was created within the project SGS-2017-004 ‘Finance and Sustainable Development from the Perspective of Theory and Practice’ at the University of West Bohemia, Faculty of Economics.

References


Economics


