REGIONAL INCOME DISPARITIES IN CENTRAL-EASTERN EUROPE – A MARKOV MODEL APPROACH

Anna Decewicz

Abstract

The paper concerns analysis of regional income disparities and their dynamics by use of Markov models including spatial effects. The main objective is to examine spatial interactions between regions by means of a non-homogeneous Markov model with transition probabilities conditioned by variables reflecting region's position in its neighbourhood. The attention focuses on disparities of GDP per capita related to European average to include differences in regional economies value and their levels of well-being. Data referring to GDP per capita on NUTS3 in 2000 – 2010 are taken from Eurostat regional database and national statistic offices. Analysis starts with studying cross-sectional distribution of regional GDP per capita by means of parametric and non-parametric estimation. Regional spatial effects are measured by Moran spatial statistics, using wage matrices of k-nearest neighbours and maximal distance. Several Markov models are estimated to examine dynamics of GDP and identify patterns of income distribution evolution and to explain how neighbourhood affects region's current and future position.

Key words: regional income, Markov chain, spatial effects

JEL Code: C23, R11

Introduction

The aim of this paper is to analyze disparities between regions of "new" members of European Union by use of Markov models. Markov models are present in literature (Fingleton, 1007, Magrini, 1999, Le Gallo, 2004) as an alternative to econometric modelling of convergence giving the possibility of predicting future income distribution, limit distribution and the speed of convergence (or perhaps divergence). The main purpose of the paper is an attempt to incorporate spatial effects between regions directly into a Markov model.

The attention is focused on disparities and convergence of Gross Domestic Product per capita to include differences in regional economies value and their levels of well-being. Data taken from Eurostat regional database and national statistic offices include annual GDP (as a percentage of UE–27 and percentage of up-regional average) at NUTS3 level for 2000–2010 period for 211 regions located in Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovenia and Slovakia.

The analysis begins with studying cross-sectional distribution of regional GDP. In order to study the dynamics of GDP distribution a non-homogeneous Markov model incorporating spatial dependence between regions is constructed. In the Markov model transition probabilities depend on variables describing region's current position in its geographic neighbourhood (i.e. the regional GDP is related to waged average of its nearest neighbours or up-regional average). This is a modification of previous approach called "spatial" Markov chain (Rey, 2001, Le Gallo, 2001) in which spatial effects and other factors impacting transition probabilities were taken into account by estimating several transition matrices for different groups of regions distinguished by the initial class to which their neighbours belong.

1 Modelling convergence by Markov models

1.1. Discrete Markov model

By definition (Bhat, 1984) Markov chain with state-space $S = \{1, 2, ..., r\}$ is a stochastic process $\{X_n\}_{n \geq 0}$ with memory-less property:

$$\forall i, j, i_{n-1}, ..., i_0 \in S \qquad (1)$$

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = P(X_{n+1} = j \mid X_n = i) = p_{ij}(n),$$

meaning that transition probabilities depend only on the last state observed. Homogeneous Markov chain is a process with time-stable transition probabilities, usually noted in transition matrix $\boldsymbol{P} = \lfloor p_{ij} \rfloor$. With known transition matrix it is possible to predict future distribution, according to formula

$$\boldsymbol{d}_{n+1} = \boldsymbol{d}_n \boldsymbol{P}, \qquad (2)$$

where $\boldsymbol{d}_n = [d_{n1} \quad ... \quad d_{nr}]$ denotes distribution at time n, $d_{ni} = P(X_n = i)$. A Markov chain is called ergodic if its limit (also called ergodic or stationary) distribution exists and does not depend on the initial distribution.

Applying Markov chains in modelling income convergence consists in defining a finite number of states referring to a set of classes distinguished basing on income distribution among regions. Transition probabilities of a homogeneous Markov chain are estimated from panel data with formula

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i}, \tag{3}$$

with n_{ij} denoting observed number of transitions $i \to j$ and n_i number of visits in state *i*. Convergence is evident when the probability mass in ergodic distribution is concentrated around one state, otherwise divergence or club convergence may be considered. Modifications of this simple model including spatial effects (Rey 2001, Le Gallo, 2004) consists in decomposition of transition matrix in a way enabling to extract transition probabilities conditioned by a class to which region's neighbours belong.

The usual assumption in modelling convergence by means of Markov chain is models' homogeneity, however it seems obvious that transition mechanism may be time and cross-sectional varying. Heterogeneity may be included by means of treating transition probabilities as functions of some explanatory variables or studying parameters' stability. In order to include the impact of explanatory variables on transition probabilities into a discrete model each row of a transition matrix should be estimated by an ordered logit model with *r* classes corresponding to a chain's states.

1.2 Continuous Markov model – estimation from panel data

Continuous Markov process describes transition mechanism in continuous time. A movement from any state to another may take place at any moment, opposite to a discrete model in which transitions occur only in fixed time points, usually identical with moments of observation. The following estimation method assumes that movements take place any time but the process is observed in fixed time moments.

Markov property in continuous case takes form

$$\forall i, j, i_{n-1}, ..., i_0 \in S, \ \forall t_0 < t_1 < ... < t_n < t_{n+1} \tag{4}$$

$$P(X_{t_{n+1}} = j \mid X_{t_n} = i, X_{t_{n-1}} = i_{n-1}, ..., X_{t_0} = i_0) = P(X_{t_{n+1}} = j \mid X_{t_n} = i) = p_{ij}(t_{n+1} - t_n).$$

The process is called homogeneous if transition probabilities within period of the length *t* are constant and it is usually described by a transition intensities matrix

$$\boldsymbol{Q} = \begin{bmatrix} -q_1 & q_{12} & \cdots & q_{1r} \\ q_{21} & -q_2 & & q_{21} \\ \vdots & & \ddots & \vdots \\ q_{r1} & q_{r2} & \cdots & -q_r \end{bmatrix},$$

with transition intensities $q_{ij} = p_{ij}^{'}(0)$, $-q_i = p_i^{'}(0) = \sum_{j \in S} q_{ij}$. Relation between transition intensity matrix and transition probability matrix in time period from 0 to t takes form $\boldsymbol{P}(t) = e^{\boldsymbol{Q}t}$. Memory-less assumption is equivalent to stating that sojourn time in state i has exponential distribution with parameter q_i.

Estimating parameters of Markov process from panel data (Kalbfleisch&Lawless, 1985) consists in maximizing likelihood

$$L(Q) = \prod_{k,n} p_{s(t_{k,n})s(t_{k,n+1})}\left(t_{k,n+1} - t_{k,n}\right), \qquad (5)$$

subject to $\ln q_{ij}$, with $p_{s(t_{k,n})s(t_{k,n+1})}$ denoting probability that an individual k observed at the moment $t_{k,n}$ in state $s(t_{k,n})$ moves to state $s(t_{k,n+1})$ at the moment $t_{k,n+1}$. In order to identify impact of exogenous factors on transition intensities a Markov model with covariates might be applied. Transition intensities are defined as functions of variables vector $\boldsymbol{z}_{k,n}$

$$q_{ij}\left(\boldsymbol{z}_{k,n}\right) = q_{ij}^{(0)} \exp\left(\boldsymbol{\beta}_{ij}^{T} \boldsymbol{z}_{k,n}\right). \qquad (6)$$

The hazard ratio $\exp(\beta_{ij})$ is then interpreted as approximate rate of increasing transition intensity from state i to state j respect to variable $z_{k,n}$.

1.3 Spatial effects

Estimation method described in section 1.2 enables to analyze dynamics of income distribution by a Markov model including spatial dependencies between regions by means of variables reflecting the impact of neighbourhood on region's movement from state to state. To measure spatial effects the neighbourhood has to be established first. The distance matrix is constructed basing on physical distance between regions geographical centres and a k-near neighbours and maximal distance wage matrices are applied. The spatial autocorrelation is tested by Moran global statistics

$$I = \frac{1}{\sum_{l,m} w_{lm}} \cdot \frac{\sum_{l,m} w_{lm}\left(x_l - \bar{x}\right)\left(x_m - \bar{x}\right)}{\frac{1}{n}\sum_{l}\left(x_l - \bar{x}\right)^2}, \qquad (7)$$

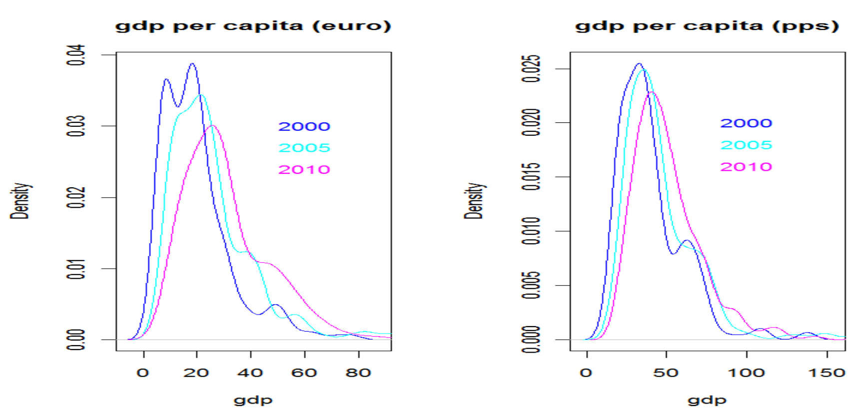
with n denoting number of regions, $\boldsymbol{W} = \left[w_{lm}\right]$ wage matrix.

Positive result of spatial autocorrelation tests proves that neighbour regions are more similar to each other then more distanced ones, i.e. poor regions have poor neighbours and rich regions have also rich neighbours. Negative autocorrelation consists in neighbourhood of regions which are not similar to each other. It is convenient to illustrate spatial autocorrelation on Moran scatterplot showing value of income in a region versus waged value in its neighbourhood. Regions lying below regression line on Moran scatterplot have bigger value of income then their neighbours, regions above the line are surrounded by richer neighbours.

2 Regional GDP per capita distribution

Figure 1 plots density function for GDP per capita in 2000, 2005 and 2010. The GDP per capita (in euro) distribution seems to be 3-modal in 2000 with the main mode around 18% of European average. The other modes seem to be vanishing in later years and are hardly seen in 2010. Table 1 gives the results of best fitted (according to BIC criteria) mixture of normal distribution for 2000, 2005 and 2010.

Fig. 1: GDP per capita in 2000, 2005, 2010



Source: Author's computation, R CRAN

Tab. 1: Mixture of normal distribution, GDP per capita (euro)

		Component 1	Component 2	Component 3
2000	Mixing proportion	0.251	0.5667	0.183
	Average	7.387	18.865	39.721
	Variance	1.470	32.349	230.945
2005	Mixing proportion	0.224	0.455	0.321
	Average	10.882	21.765	40.798
	Variance	3.389	23.037	308.576
2010	Mixing proportion		0.636	0.364
	Average		23.069	47.921
	Variance		58.417	489.284

Source: Author's computation, with spded package of R CRAN

To study spatial dependencies between regions wage matrix has been computed basing on *k*-near neighbours (k=5). Distance is measured between geographical centres of regions. Moran statistics for years 2000, 2005 and 2010 are presented in Table 2. Each case proves significant positive autocorrelation, however it seems to be decreasing in proceeding years. The results obtained for maximal distance (250 km) matrix show similar tendency.

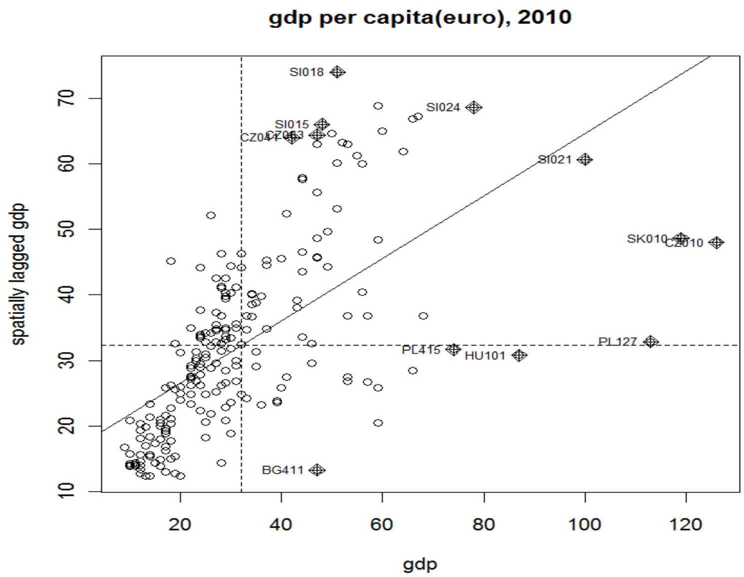
Tab. 2: Global Moran statistics - GDP per capita relative to EU average

	GDP per capita (euro)	**GDP per capita (pps)**
2000	Moran I statistic standard deviation = 17.6262, p-value < 2.2e-16, alternative hypothesis: greater Moran I statistic Expectation Variance 0.694576117 -0.004761905 0.001574192	Moran I statistic standard deviate = 13.2785, p-value < 2.2e-16, alternative hypothesis: greater Moran I statistic Expectation Variance 0.520593621 -0.004761905 0.001565341
2005	Moran I statistic standard deviate = 14.7911, p-value < 2.2e-16 Moran I statistic Expectation Variance 0.580522285 -0.004761905 0.001565781	Moran I statistic standard deviate = 10.2016, p-value < 2.2e-16 Moran I statistic Expectation Variance 0.396688762 -0.004761905 0.001548556
2010	Moran I statistic standard deviate = 12.1757, p-value < 2.2e-16 Moran I statistic Expectation Variance 0.475239543 -0.004761905 0.001554163	Moran I statistic standard deviate = 6.969, p-value = 1.596e-12 Moran I statistic Expectation Variance 0.268671601 -0.004761905 0.001539429

Source: Author's computation, with spded package of R CRAN

Moran scatterplot presented on Figure 2 shows the most outstanding points on the plot lying below the regression line referring to regions which are significantly richer comparing to their neighbours: SI021 – Osrednjeslovenska (Slovenia), CZ010 – Hlavní mesto Praha (Czech Republic), PL127 – Miasto Warszawa (Poland), PL415 – Miasto Poznań (Poland), HU101 – Budapest (Hungary).

Fig. 2: Moran scatterplot – GDP per capita (euro)



Source: Author's computation, with spded package of R CRAN

3 Distribution dynamics – estimation of Markov models

To analyze dynamics of regional income distribution by means of Markov chain one has to remember that definition of states, i.e. the method of transforming a continuous variable to a discrete one, may influence the results. Distinguishing more classes should better reflect income distribution but results in increasing number of model's parameters and possible problems with quality of estimates.

Seven states of a Markov model have been distinguished basing on 15th, 30th, 45th, 60th, 75th and 90th percentiles of regional income distribution. As available data form a panel the analysis is conducted with methods described in section 1.2. The intensity matrix for homogeneous model has been estimated and one-year transition probability matrix, limit distribution and sojourn times in each state have been calculated (Table 3 and 4). Probabilities to stay in the same state in the next year are very high, particularly for state 1 and 7 (the poorest and the richest class). Expected time to next movement (up or down) is between 2.86 years for regions in state 4 (middle class) to 11.13 years in state 7. For states 2, 5 and 6 movements down within one year are more likely than ups, for states 3 and 4 the opposite. Process is ergodic and its limit distributions gives no signs of GDP convergence.

Tab. 3: Transition intensities, sojourn time and ergodic distribution for homogenous model

from \ to	1	2	3	4	5	6	7
1	-0.1283 (0.0237)	0.1283 (0.0237)					
2	0.1469 (0.0224)	-0.2365 (0.02830)	0.0895 (0.0171)				
3		0.0883 (0.0176)	-0.2027 (0.0269)	0.1143 (0.0203)			
4			0.1319 (0.0227)	-0.3501 (0.0385)	0.2182 (0.0308)		
5				0.1821 (0.0254)	-0.2601 (0.0298)	0.0781 (0.0155)	
6					0.0888 (0.0174)	-0.1530 (0.0229)	0.0642 (0.0148)
7						0.0899 (0.0207)	-0.0899 (0.0207)
sojourn time	7.79 (1.44)	4.23 (0.50)	4.94 (0.65)	2.86 (0.31)	3.84 (0.44)	6.53 (0.97)	11.13(2.56)
ergodic distribution	0.1716	0.1498	0.1518	0.1315	0.1577	0.1386	0.0990

Notes: Standard errors in parenthesis
Source: Author's computation, with msm package of R CRAN

Tab. 4: One-year transition probabilities for homogeneous model

from \ to	1	2	3	4	5	6	7
1	0.8876	0.1074	0.0048	0.0002			
2	0.1231	0.8003	0.0724	0.0040	0.0002		
3	0.0054	0.0715	0.8256	0.0877	0.0096	0.0002	
4	0.0003	0.0045	0.1012	0.7247	0.1626	0.0066	0.0001
5		0.0003	0.0092	0.1356	0.7887	0.0640	0.002
6			0.0003	0.0063	0.0729	0.8635	0.0570
7				0.0002	0.0034	0.0798	0.9166

Notes: Standard errors in parenthesis
Source: Author's computation, with msm package of R CRAN

Next part of analysis is an attempt to include spatial relations between regions directly into a model. The first nonhomogeneous Markov model (Model 1) with covariates is the one with variable *N* defined as waged average of GDP per capita level in region's neighbourhood defined by *k*-near neighbours matrix **W** computed before. Such definition is supposed to regard impact of neighbours' current position on transition intensities and transition probabilities. In the second model (Model 2) covariate *U* refers to regions position in its neighbourhood measured by percentage of NUTS2 level average of the area to which it belongs. The third model (Model 3) has time covariate. Table 5 contains hazard ratios $\exp(\beta_{ij})$ for each pair of states with non-zero transition intensities for three nonhomogeneous models. The likelihood ratio test indicates models 1 and 2 to perform better than the homogeneous model, model 3 has been rejected.

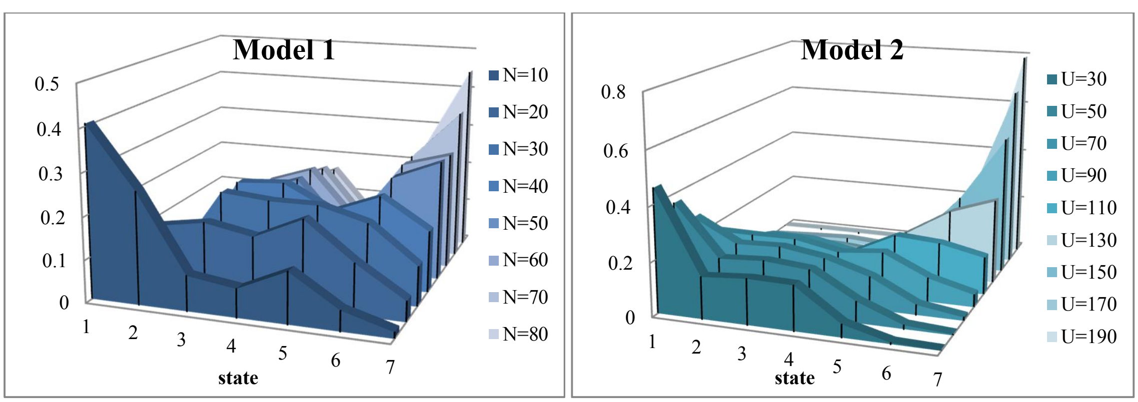
Tab. 5: Hazard ratio, nonhomogeneous models

		Model 1			Model 2			Model 3		
from	**to**	**HR**	**L**	**U**	**HR**	**L**	**U**	**HR**	**L**	**U**
1	**2**	1.0102	0.9182	1.1115	1.0079	0.9851	1.0313	0.9005	0.7863	1.0312
2	**1**	0.9552	0.9039	1.0094	0.9901	0.9750	1.0055	0.8665	0.7743	0.9698
2	**3**	1.0659	1.0260	1.1074	0.9976	0.9798	1.0158	1.0861	0.9558	1.2342
3	**2**	0.9388	0.8888	0.9916	0.9988	0.9810	1.0169	1.0709	0.9288	1.2348
3	**4**	1.0110	0.9656	1.0584	1.0080	0.9923	1.0240	0.9797	0.8681	1.1057
4	**3**	1.0027	0.9581	1.0493	1.0110	0.9950	1.0272	1.0479	0.9367	1.1724
4	**5**	0.9677	0.9307	1.0062	1.0011	0.9870	1.0154	0.9640	0.8733	1.0641
5	**4**	0.9962	0.9598	1.0340	0.9788	0.9637	0.9942	0.9485	0.8620	1.0437
5	**6**	0.9922	0.9403	1.0469	0.9965	0.9761	1.0174	0.9242	0.8089	1.0559
6	**5**	0.9544	0.9189	0.9913	0.9692	0.9526	0.9861	0.9693	0.8490	1.1067
6	**7**	1.0044	0.9699	1.0401	1.0182	1.0004	1.0362	1.0157	0.8741	1.1804
7	**6**	0.9783	0.9517	1.0057	0.9964	0.9823	1.0108	1.0664	0.9115	1.2475

Notes: L and U are lower and upper limit of 95% confidence interval.
Source: Author's computation, with msm package of R CRAN

To see how neighbourhood impacts limit distribution several intensities matrices for the whole range of possible values of covariates resulting from their distributions have been calculated for Model 1 and Model 2. Limit distribution of GDP per capita in region surrounded by poor neighbours concentrates in state 1, for regions surrounded by rich neighbours mass of probability moves to states 4, 6 and 7. Similarly, the mass of probability in limit distribution moves to higher states as region's position compared to up-regional average increases (Figure 3).

Fig. 3: Limit distributions for models with covariates



Source: Author's computation, with msm package of R CRAN

Conclusions

The purpose of this paper was to show how Markov models with covariate can be applied to analyze income distribution dynamics. This approach is a modification of classic application of Markov chains in modelling convergence giving possibility to extract impact of particular factors (of time or spatial nature) on transition probabilities.

Markov models applied for GDP per capita distribution dynamics in regions of new member countries of UE from 2000 to 2010 show strong tendency to stay in the same class of GDP level in succeeding years. Models with spatial effects prove that region is more probable to be in higher states (referring to higher level of GDP per capita) if it has rich neighbours.

References

Anselin L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer, Dordrecht.

Bhat, N. (1984). *Elements of applied stochastic processes*. Wiley&Sons.

Fingleton, B. (1997). Specification and testing of Markov chain models: An application to convergence in the European Union. *Oxford Bulletin of Economics and Statistics*, *59*(3), 358-403. doi: 10.1111/1468-0084.00072

Kalbfleisch, J., & Lawless, J. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, *80*(392), 863-871. doi: 10.2307/2288545

Kopczewska, K. (2006). *Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN*. CeDeWu.

Le Gallo, J. (2004). Space-time analysis GDP disparities cross European regions: A Markov chain approach. *International Regional Science Review 27*, 138-163. doi: 10.1177/0160017603262402

Magrini, S. (1999). The evolution of income disparities among the regions of the European Union. *Regional Science and Urban Economics, 29*(2), 257-281. doi: 10.1016/S0166-0462(98)00039-8

Quah, D.T. (1996). Empirics for Economic Growth and Convergence. *European Economic Review, 40*(6), 1353-1375. doi: 10.1016/0014-2921(95)00051-8

Quah, D.T. (1996). Regional Convergence Clusters across Europe. *European Economic Review 40* (3-5), 951-958. doi: 10.1016/0014-2921(95)00105-0

Rey, S.J. & Montouri, B.D. (1999). US Regional Income Convergence: A Spatial Econometric Perspective. *Regional Studies 33*(2), 143-156.

Rey, S.J. (2001). Spatial empirics for regional economic growth and convergence. *Geographical Analysis 33*(3), 195-214.

Contact

Anna Decewicz
Institute of Econometrics, Warsaw School of Economics
Al. Niepodległości 162, 02-554 Warszawa
adecew@sgh.waw.pl