DYNAMIC ECONOMETRIC MODELS
Vol. 7 – Nicolaus Copernicus University – Toruń – 2006

Elżbieta Szulc
Nicolaus Copernicus University in Toruń

Specification of the Dynamic Model with the Spatial Structure of Connections

1. Introduction

A lot of phenomena which are a subject of econometric analyses one treats as space-time processes and tries to discover in them the established temporal dynamics – on the one hand – and some spatial connections – on the other hand. Such analyses may concern, e.g. unemployment. They are realized on the ground of the data relating to the use of labour resources, which refer to the spatial units delimited by the administrative division, i.e. to provinces, counties or districts, in the successive units of time, e.g. in months.

The investigations focused on the purely temporal models of a single measure of unemployment are not sufficient even then when they relate to each time unit separately, if they do not extend the analysis towards the comparison of the results obtained for these units and if they do not ask about possible spatial dependence.

The measures of labour resources or of unemployment based on the dwelling place may show the spatial dependence considering some labour mobility. People often crossing the borders of their regions find out jobs in the neighbouring zones. The spatial connections do not have to be geographically defined, but the importance of the physical distance is still great to determine the magnitude of the spatial dependence. It is obvious, that these connections are – to a certain degree – limited by physical distances and by transport nets.

The magnitudes of unemployment, which referred to the spatial units form spatial patterns of this phenomenon. The patterns may result – on the one hand – from the spatial patterns of labour demand connected with the economic activity of spatial units and – on the other hand – they may be connected with the spatial distribution of some labour characteristics, such as qualifications, a level of education, age etc. Thus, the economic characteristics of spatial units, from both the labour demand's and the labour supply's points of view, may be the basis of determining the so-called economic distance between them and of defining the connections across the economic neighbours of different orders.

In many situations in the investigation of the so-called spatial autocorrelation the notion of the economic distance will be the important generalization of the notion of the physical distance. The use of the economic distance in the investigations of the dependence between individuals may give the results which differ from these ones obtained on the basis of the purely physical distance.

Defining the appropriate measure of the distance is difficult, when the so-called space-time autocorrelation is analysed. In this case differing physical distance from economic distance is necessary as well. However, it is hard to include the economic distance into the spatio-temporal analyses with regard for its changeability in time.

The main purpose of the paper is to consider the possibilities of identifying the autodependence in the space-time processes. In section 2 the autodependence is determined in terms of the spatial and space-time autocorrelations. There is underlined the part whose the appropriate defining of the distance between the units of the investigation plays in the studies of the autocorrelation. This question is wider discussed in section 3. Section 4 presents the outline of the modelling, when some problems, pointed before, such as too little number of observations across space and a difficult for defining space-time distance, are attempted to be omitted. In section 5 the results of the empirical researches are discussed. Considering the limited volume of the paper, the detailed results of the calculations are not given here. Finally, the general conclusions are formulated and the further investigations are announced as well.

2. Spatial and Space-time Autocorrelation

In the literature the idea of the spatial autocorrelation came into existence soon after the conception of the time autocorrelation (see, e.g. Cliff & Ord, 1973, 1981). There is underlined a special character of the autocorrelation in space, pointing to some important characteristic features. First of all the lack of natural ordering on the line: *past – present – future* is showed. Thus multidirectionality of the connections and dependence is pointed out. From that some difficulties in the matter of the spatial autocorrelation measurement and of testing it come.

In practice, when the spatial autocorrelation is defined the following situations may be distinguished:

1) the positive spatial autocorrelation; it occurs, when the phenomena localized close by have similar properties;

2) the negative spatial autocorrelation, when the phenomena close in the space, differ from each other with regard to their properties;

3) the zero spatial autocorrelation, which means that the properties of the phenomena are independent of location.

Finding out the appropriate ways for discovering and modelling the spatial autodependence, investigators more and more frequently try to give to the spatial data the structure similar to that one whose the time series (namely the data indexed by time variable *t*) have. The basis of the investigation of the spatial autocorrelation is the assumption, that the autodependence grows small together with the distance.

The another important statement, which is logically justified and supported by the observations of the course of phenomena, is that the phenomena in the near locations interact over time, and not „instantaneously" (see, e.g. Epperson, 2000). In other words, the realization of the dependence in space needs a lapse of time. In this way the notion of the space-time autocorrelation becomes more natural than the notion of the purely spatial autocorrelation. The essence of the space-time autocorrelation between the spatial units *i* and *j* may be expressed with the general formula as follows:

$$y_{it} = f\left(y_{it-l}, y_{jt-l}, \varepsilon_{it}\right). \tag{1}$$

The dependence y_{it} on y_{it-1}, …, y_{it-l} in equation (1) expresses the so-called purely temporal regressive component. On the contrary, the purely spatial regressive component is not considered. Although it is an important practical problem, it was not taken into account in this investigation. The matter is that – in practice – the time lag between the registrations of data in the separate spatial units is too long in comparison with the realization of the dependence. Then the so-called simultaneous spatial dependence (the purely spatial autocorrelation) may be observed.

3. Distance Metrics[1]

There are possible various measures of the distance between the pairs of the spatial units. The distance does not have to denote the physical distance, thus the space in which the units of the investigation are located does not have to be geographical space.

Specified features of the spatial units, as the arguments of the functions describing the course of the phenomena in these units, create another space, which may be called the economic space. Then the relative position of the units, on the basis of the economic distance between them, is considered.

[1] On the measurement of the distance in the context of the investigation of the spatial connections in the unemployment rates, see, e.g. Conley, Topa (2002).

To establish the relative position of the spatial units (regions) there are taken into consideration, e.g.: the usual physical distance between the centres of these regions or the mean distance between the main places of the connected regions, the transportation costs of human capital or the travelling time as well. To construct the economic distance in a wide sense the Euclidian distance is usually calculated for the vector of the attributes of the spatial units, i.e.:

$$d_{ij} = \sqrt{\sum_k \left(x_{ik} - x_{jk}\right)^2} \qquad (2)$$

where: x_{ik}, x_{jk}, k-th attribute of the unit i and j.

When the subject of the investigation is unemployment the characteristics in formula (2) may relate to a level or a kind of education in the separated regions and to the age structure, on the one hand, or to a level of the industrial production, to the investment etc., on the other hand.

It is not clear how to define the space-time distance. The two-dimensional geographical space is multilateral. Similarly, the k-dimensional economic space, which, after all, is usually reduced to the two-dimensional space, has multilateral character as well. Time is unidirectional and in the system: *space – time* it is not simply the additional dimension. Moreover, the space-time data usually refer to a large number of the observations „over time" and to a small number of the spatial units. Therefore, an interesting solution is the suggestion to investigate the time series, which deals with the individual spatial units separately, and then to include into the model the spatial structure determined by the distance between the units.

4. Space-time Econometric Model. Theoretical Background

It is established, that the spatial unit i is located at the point s_i in the Euclidian space R^n, in particular in R^2. The locations of the i-th unit at the time t are denoted as (s_i, t). The values of the random variable $Y_{s_i,t}$ at the collection of locations $\{s_i, t\}_{i=1}^{N}$, for $t=1, 2, ..., T$, are observed. The observations are designated by $y_{s_i,t}$. The random variable $Y_{s_i,t}$ associated with the spatial position s_i and with the time point t is called the spatio-temporal random field[2].

For each established location s_i, the random variable $Y_{\cdot,t}$, i.e. the stochastic process is obtained. Thus, $Y_{s_i,t}$ may be seen as the vector $Y_t = \{Y_{1t}, Y_{2t}, ..., Y_{Nt}\}$. Likewise, the random variable $Y_{s_i,\cdot}$, i.e. the spatial process may be considered,

[2] The main use of the random fields in econometrics refers to the spatial and spatio-temporal analyses, see, e.g.: Conley (1999), Chen, Conley (2001), Conley, Topa (2002).

and then $Y_{s_i} = \{Y_{s_i,1}, Y_{s_i,2}, ..., Y_{s_i,N}\}$ is obtained. $Y_{\cdot,t}$ may be named the conditional random field with regard to space. Respectively, $Y_{s_i,\cdot}$ may be treated as the conditional random field with regard to time.

The autodependence models for the stochastic processes (autoregressive models) are constructed assuming that the dependence grows small as the distance between the observations of the process (the values of the time series) grows large. Likewise, in the models of autodependence for the spatial processes, there is the principle that the dependence is characterized by the distance between the spatial locations of the observations. If the observations i and j are „close", then the random variables Y_{s_i} and Y_{s_j} may be strongly correlated. On the contrary, as the distance between s_i and s_j grows large, the random variables Y_{s_i} and Y_{s_j} become closer to being independent. Since it is not clear how to measure the space-time distance (see, section 3), the definition of the space-time autodependence model is difficult.

The modelling of economic phenomena as the spatio-temporal random fields needs many other problems to be solved. There are conditions practically difficult to be satisfied, e.g. the conditions of: stationarity, homogeneity, isotropy. Moreover, the number of the spatial units, for which the statistical data is accessible, is usually small.

In the situation, when there is the data collected at few spatial regions but at many regular time intervals, the time series may be investigated for each spatial unit separately. The modelling of the time series for each spatial point or area separately is reasonable only in the case of the independent temporal random fields. In the appropriate models the time dependence is taken into account, but the spatial dependence is not considered.

In the literature there was appeared an interesting suggestion to include into such analyses the spatial dependence by constructing models with spatial structure (see de Luna & Genton, 2004)[3]. The idea of taking into consideration the structure of the spatial connections was adapted in the presented study as well. The vector $Y_t = \{Y_{1t}, Y_{2t}, ..., Y_{Nt}\}$ in the meaning mentioned above is considered. The models are constructed independently for each separated spatial unit, with regard to a spatio-temporal ordering of the units to establish a hierarchy in which the components of the vector Y_t are introduced into the model. The basis of the spatial ordering is the distance between the units, determined, e.g. by considering the geographical locations of these units. The natural order of the

[3] The model, which de Luna and Genton consider, is a vector autoregressive (VAR) model, similar to the models commonly used in the multivariate time series analysis. However, it is essentially different, because it refers to the single variable (i.e. unemployment) observed at several spatial regions of a lattice at the same time, whereas the others are applied in the situations where at the same time several variables are observed.

„neighbourhood" of the spatial units is established for the given unit. Finally, it is possible to say about the $N \times N$ matrix of „distance" designating such orders for each of the units (see, e.g. Table 2).

The variables $Y_{s_i,t}$ are introduced into the model sequentially, according to the natural spatio-temporal order, i.e. firstly $Y_{s_i,t-l}$, in the spatially defined order, for l=1, secondly, in the spatially defined order again, for l=2, etc.

The „neighbours" of the given unit connected in space and time are established on the basis of the partial correlations. It means that the new variables are introduced into the model so long as the correlations are significantly different from zero. The spatio-temporal partial correlation denotes here the correlation between the established element of the vector Y_t and its another element at the time lag l, with including the impacts of the variables introduced before.

5. Analysis of Empirical Data

The investigation concerned the unemployment in Poland, across the provinces over the period: January 1999 – December 2005. The data comes from the statistical bulletins by the Central Statistical Office: *Registered unemployment. Information and statistic papers*, Warsaw and from: (http://www.stat.gov.pl).

As it was argued above (see, section 2) the data may be spatially correlated. It means, that the unemployment rate in a given province may be correlated with the rates in neighbouring provinces. The statistical sample consists of 84 time observations for each from 16 spatial units, i.e. of 1344 observations together. These 16 time series create the realizations of the processes, which enter into the vector defining the space-time process.

First the trend-seasonal and autoregressive structure was investigated, separately for each time series[4]. Table 1 presents the results of the investigation in this domain.

Using the least squares method the parameters of the appropriate general models were estimated. Then the residuals obtained from the fitted models were investigated. The question was whether they created the space-time white noise process. The residual analysis was carried out to investigate the relevance of the fitted models. The correlations for the time series of the residuals were calculated. Most of them appeared significantly different from zero at the 5 percent level. Next, the correlations of the residuals at time lag 1 were calculated. In this case some of them appeared significant as well. The correlations at time lags larger than 1 were low and insignificant.

The conclusion from the analysis is following. Investigating the separated time series was not sufficient for discovering the full structure of the unem-

[4] For this purpose the standard techniques, such as: t-statistics, F-tests and autocorrelation and partial autocorrelation functions were used.

ployment process. There is the spatio-temporal dependence, which should be taken into consideration in an appropriate model.

Table 1. Trend, seasonality, autoregression of investigated time series

Province	Dln.	K.-p.	Lbl.	Lbs.	Łdz.	Młp.	Mzw.	Opl.
Degree of trend	3	2	2	3	2	2	3	3
Seasonality	+	+	+	+	+	+	+	+
Order of autoregression	1	1*	1	1	1	1	1	1
Province	Pdk.	Pdl.	Pmr.	Śls.	Świt.	W.-m.	Wlk.	Zch.
Degree of trend	2	2	3	2	2	2	3	3
Seasonality	+	+	+	+	+	+	+	+
Order of autoregression	1	1	1	1*	1	1	1	1

Source: own calculations.
„+" denotes a presence of the seasonal component.
„*" denotes that sporadically the significant partial correlations at remote time lags appeared. However, they were treated as spurious.
The abbreviations of Polish names of the provinces (which are not compounds) were created with the first three consonants of each of the name, i.e.:
Dln.=Dolnośląskie
Lbl.=Lubelskie
Lbs.=Lubuskie
Łdz.=Łódzkie
Młp.=Małopolskie
Mzw.=Mazowieckie
Opl.=Opolskie
Pdk.=Podkarpackie
Pdl.=Podlaskie
Pmr.=Pomorskie
Śls.=Śląskie
Śwt.=Świętokrzyskie
Wlk.=Wielkopolskie
Zch.=Zachodniopomorskie
For the compounds names they are:
K.-p.=Kujawsko-pomorskie
W.-m.=Warmińsko-mazurskie

In order to consider the spatio-temporal dependence in the investigated unemployment process the information on the spatial structure of the connections between the provinces was used. The structure is given in table 2. Each row of table 2 presents the spatial ordering (from 0 to 15) of the provinces. The spatial order of the „neighbourhood" for each province was fixed on the basis of the mean physical distance between the main cities of the provinces.

Again the detrended and deseasonalized time series of the unemployment rates (with the purpose to obtain stationarity) was investigated. With assumption on the given ordering of the spatial connections for each province the matrices of correlations of time series for time lags from 1 to 12 were computed. The correlations grew small quickly as the time distances grew large. The analogous conformity for the spatial „lags" was less clear. Next, the partial correlations were used to choose for each province the spatial „neighbours", significantly connected with regard to the unemployment rate. The settlements are given in Table 3. For example, the province Dln. connects with the provinces: Dln., Opl., Wlk., Lbs., Śls., Łdz., K.-p., Śwt., Młp., Zch., Mzw., at time lag 1, with 8 provinces in sequence (see, Table 2) at time lag 2 and with none of the provinces at time lag 3, 4 and at the greater ones.

Table 2. Spatial structure of „distance" between provinces

	Dln	K-p	Lbl	Lbs	Łdz	Młp	Mzw	Opl	Pdk	Pdl	Pmr	Śls	Śwt	Wm	Wlk	Zch
Dln	0	6	14	3	5	8	10	1	13	15	11	4	7	12	2	9
K-p	10	0	13	6	2	14	4	8	15	11	5	12	7	3	1	9
Lbl	12	8	0	14	4	6	3	10	1	5	13	7	2	9	11	15
Lbs	2	4	14	0	7	12	10	5	15	13	6	8	11	9	1	3
Łdz	9	3	8	13	0	7	2	6	10	12	14	5	1	11	4	15
Młp	8	10	6	12	5	0	7	4	2	11	14	3	1	13	9	15
Mzw	13	3	5	14	1	10	0	12	9	2	11	8	4	6	7	15
Opl	2	9	11	8	4	6	10	0	7	15	14	1	5	12	3	13
Pdk	11	10	3	13	6	1	7	5	0	8	14	4	2	12	9	15
Pdl	15	5	3	13	4	11	1	12	9	0	7	10	6	2	8	14
Pmr	9	2	13	5	7	14	6	10	15	8	0	12	11	1	4	3
Śls	6	9	10	11	4	2	8	1	5	12	14	0	3	13	7	15
Śwt	10	9	5	13	1	2	6	7	4	11	14	3	0	12	8	15
W-m	12	2	10	8	5	14	3	11	15	4	1	13	9	0	6	7
Wlk	3	1	13	2	4	12	7	5	15	14	11	6	8	10	0	9
Zch	6	3	14	1	7	13	9	8	15	12	2	10	11	5	4	0

Source: own calculations.

Table 3. Structure of space-time dependence

lag	Dln	K-p	Lbl	Lbs	Łdz	Młp	Mzw	Opl	Pdk	Pdl	Pmr	Śls	Śwt	Wm	Wlk	Zch
1	11	15	13	13	15	9	12	11	15	14	15	9	15	15	13	14
2	8	8	5	0	2	6	0	0	7	0	0	5	2	0	9	0
3	0	0	0	0	0	0	10	0	8	4	12	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Source: own calculations.

6. Conclusions

The purpose of the investigation was to verify whether the arbitrarily taken order of the connections in time and in space across the provinces allows to specify the logical, convincing structure of the spatial and time lags for an appropriate econometric model of unemployment. There is little evidence of spatial dependence according to physical distance. The regions appear to be connected with the others, regardless of their geographical proximity.

In the proposed methodological conception the correct defining of the distance between the units of the investigation is important. However, the analysis was limited to the imperfect measure of the physical distance. Therefore, the obtained results only in general confirm the hypotheses of the nature of the dependence in the investigated phenomenon, i.e. that the dependence exists between adjacent spatial units and it disappears for more distant zones. On the contrary, sometimes the dependence appears only for the remote units.

The further investigation will be undertaken with the use of the properly defined economic distance.

References

Cliff, A.D., Ord, J.K. (1973), *Spatial Autocorrelation*, Pion, London.

Cliff, A.D., Ord, J.K. (1981), *Spatial Processes: Theory and Applications*, Pion, London.

Chen, X., Conley, T.G. (2001), A new semiparametric spatial model for panel time series, *Journal of Econometrics*, 105, 59–83.

Conley, T.G. (1999), GMM estimation with cross sectional dependence, *Journal of Econometrics*, 92, 1–45.

Conley, T.G., Topa, G. (2002), Socio-economic distance and spatial patterns in unemployment, *Journal of Applied Econometrics*, 17, 303–327.

Epperson, B.K. (2000), Spatial and space-time correlations in ecological models, *Ecological Modelling*, 132, 63–76.

de Luna, X., Genton, M.G. (2004), Spatio-temporal autoregressive models for U.S. unemployment rate, in: Lesage, J.P., Pace, R.K. (eds.), *Spatial and Spatiotemporal Econometrics*, *Advances in Econometrics*, vol. 17, Elsevier Ltd., Oxford, 279–294.