

DYNAMICZNE MODELE EKONOMETRYCZNE

IX Ogólnopolskie Seminarium Naukowe, 6–8 września 2005 w Toruniu
Katedra Ekonometrii i Statystyki, Uniwersytet Mikołaja Kopernika w Toruniu

Elżbieta Szulc

Uniwersytet Mikołaja Kopernika w Toruniu

Specyfikacja dynamicznego modelu z przestrzenną strukturą zależności

1. Wprowadzenie

Wiele zjawisk będących przedmiotem analiz ekonometrycznych traktuje się jako procesy przestrzenno-czasowe i próbuje się odkryć w nich, z jednej strony określoną dynamikę czasową, z drugiej zaś pewne powiązania przestrzenne. Takie analizy mogą dotyczyć, na przykład, bezrobocia. Przeprowadza się je na podstawie danych dotyczących wykorzystania zasobów pracy, które odnoszą się do jednostek przestrzennych wyznaczonych przez podziały administracyjne a więc, na przykład, do województw, powiatów czy gmin, w kolejnych jednostkach czasu, np. miesiącach.

Badania koncentrujące się na czasowych modelach określonej miary bezrobocia nie są wystarczające, nawet wtedy, gdy dotyczą oddzielnie każdej jednostki przestrzennej, jeśli nie rozwijają analizy w kierunku porównań wyników uzyskanych dla tych jednostek i nie stawiają pytań o ewentualne zależności przestrzenne.

Miary zasobów pracy lub bezrobocia oparte na miejscu zamieszkania ludzi mogą wykazywać przestrzenną zależność choćby z uwagi na pewną mobilność siły roboczej, która przekraczając granice rejonów zamieszkania poszukuje zatrudnienia w sąsiadujących strefach. Powiązania przestrzenne nie muszą być definiowane geograficznie, choć znaczenie odległości fizycznej jest ciągle duże w określaniu rozmiarów zależności przestrzennych. Jest oczywiste, że powiązania te do pewnego stopnia są ograniczone przez odległości fizyczne i sieci transportowe.

Wielkości bezrobocia odnoszone do określonych jednostek przestrzennych tworzą przestrzenne szablony tego zjawiska, które mogą wynikać, z jednej strony z przestrzennych szablonów popytu na siłę roboczą, związanych z określoną aktywnością ekonomiczną jednostek przestrzennych, z drugiej zaś mogą być związane z przestrzennymi rozkładami pewnych charakterystyk siły roboczej, takich jak kwalifikacje, poziom wykształcenia, wiek itp. Zatem charakterystyki ekonomiczne jednostek przestrzennych, zarówno z punktu widzenia popytu na siłę roboczą, jak i jej podaży, mogą być podstawą określania tzw. ekonomicznej odległości między nimi i definiowania powiązań w przekroju ekonomicznych sąsiadów różnych rzędów.

W wielu sytuacjach w badaniach tzw. autokorelacji przestrzennej pojęcie odległości ekonomicznej będzie ważnym uogólnieniem pojęcia odległości fizycznej. Wykorzystanie odległości ekonomicznej w badaniach zależności między obiektami może prowadzić do wyników odmiennych od tych uzyskanych na podstawie odległości czysto fizycznej.

Definiowanie odpowiedniej miary odległości jest utrudnione, gdy analizuje się tzw. autokorelację przestrzenno-czasową. Również w tym wypadku naturalna wydaje się potrzeba rozróżnienia odległości fizycznej i ekonomicznej. Jednak włączenie tej ostatniej do analiz przestrzenno-czasowych utrudnia fakt jej zmienności w czasie.

Celem referatu jest rozważenie pewnych możliwości modelowania autozależności w procesach przestrzenno-czasowych. W punkcie 2 określa się autozależności w terminach autokorelacji przestrzennej i autokorelacji przestrzenno-czasowej. Podkreśla się rolę, jaką w badaniu autokorelacji odgrywa odpowiednie zdefiniowanie odległości pomiędzy jednostkami badania. Kwestia ta jest nieco szerzej omawiana w punkcie 3. Punkt 4 przedstawia zarys modelowania, gdy próbuje się ominąć niektóre sygnalizowane wcześniej problemy, takie jak mała liczba obserwacji przestrzennych oraz trudna do zdefiniowania odległość przestrzenno-czasowa. W punkcie 5 omówiono wyniki badań empirycznych. Ze względu na ograniczone ramy publikacji nie umieszczono szczegółowych wyników obliczeń. Na koniec sformułowano wnioski ogólne oraz zapowiedź dalszych badań.

2. Autokorelacja przestrzenna i przestrzenno-czasowa

Koncepcja autokorelacji przestrzennej pojawiła się w literaturze w ślad za pojęciem autokorelacji czasowej (zob. np. Cliff, Ord, 1981). Podkreśla się specjalny charakter autokorelacji w przestrzeni, wskazując na pewne ważne cechy charakterystyczne. Wskazuje się przede wszystkim na brak naturalnego uporządkowania na linii: *przeszłość* – *teraźniejszość* – *przyszłość*, a zatem na wielokierunkowość powiązań i wpływów. Rodzi to określone problemy pomiaru i oceny istotności tego zjawiska.

W praktyce, definiując autokorelację przestrzenną można wyróżnić następujące sytuacje:

- 1) występowanie dodatniej autokorelacji; zachodzi, gdy zjawiska zlokalizowane blisko siebie mają podobne właściwości;
- 2) występowanie ujemnej autokorelacji, gdy bliskie w przestrzeni zjawiska różnią się wyraźnie co do właściwości;
- 3) brak autokorelacji, co oznacza, że własności zjawiska są niezależne od lokalizacji.

Poszukując odpowiednich sposobów na wykrywanie i modelowanie przestrzennych auto-zależności, badacze coraz częściej próbują nadać danym przestrzennym strukturę podobną do tej, którą posiadają szeregi czasowe, a więc dane indeksowane za pomocą kolejnych, równo odległych jednostek czasu. Podstawą zaś badania autokorelacji przestrzennej jest przyjęcie założenia, że auto-zależność maleje wraz z odległością.

Innym ważnym stwierdzeniem badaczy, logicznie uzasadnionym i popartym obserwacjami przebiegu zjawisk jest to, że zjawiska w bliskich lokalizacjach oddziałują w czasie a nie natychmiast (patrz, np. B. Epperson, 2000). Innymi słowy, realizacja zależności w przestrzeni wymaga upływu czasu. W ten sposób bardziej naturalne niż pojęcie autokorelacji czysto przestrzennej staje się pojęcie autokorelacji przestrzenno-czasowej. Istotę autokorelacji przestrzenno-czasowej pomiędzy przestrzennymi jednostkami i oraz j można wyrazić za pomocą następującej ogólnej formuły

$$y_{it} = f(y_{it-1}, y_{jt-1}, \varepsilon_{it}). \quad (1)$$

Zależność y_{it} od $y_{it-1}, \dots, y_{jt-1}$ w równaniu (1) jest tzw. czysto czasowym składnikiem regresyjnym. Nie uwzględnia się natomiast czysto przestrzennego składnika regresyjnego. Jest to ważny problem praktyczny, który pominięto w omawianym dalej badaniu. Chodzi o to, że w praktyce odstęp czasu pomiędzy rejestracją danych w poszczególnych jednostkach przestrzennych jest zbyt długi w porównaniu z realizacją zależności. Wtedy mogą być obserwowane tzw. równoczesne zależności przestrzenne (autokorelacja czysto przestrzenna).

3. Miary odległości¹

Możliwe są różne miary odległości pomiędzy parami jednostek przestrzennych. Odległość nie musi oznaczać odległości fizycznej, a zatem przestrzeń, w której zlokalizowane są jednostki badania nie musi być przestrzenią geograficzną.

Określone cechy jednostek przestrzennych, jako argumenty funkcji opisujących przebieg zjawisk w tych jednostkach tworzą inną przestrzeń, którą można

¹ Na temat pomiaru odległości w kontekście badania przestrzennych powiązań w tempie bezrobocia, zob., np. Conley, Topa (2002).

nazwać przestrzenią ekonomiczną. Wówczas rozważa się relatywne położenie jednostek na podstawie odległości ekonomicznej pomiędzy nimi.

Ustalając relatywne położenie jednostek przestrzennych (rejonów) bierze się pod uwagę, na przykład, zwykłą odległość fizyczną pomiędzy centrami tych rejonów lub średnią odległość pomiędzy głównymi miejscowościami dla kojarzonych rejonów, koszty transportu kapitału ludzkiego, bądź też czas podróży. W sytuacji konstrukcji szerzej rozumianej odległości ekonomicznej zwykle oblicza się odległość euklidesową dla wektora charakterystyk (atrybutów) jednostek przestrzennych, tj.

$$d_{ij} = \sqrt{\sum_k (x_{ik} - x_{jk})^2}, \quad (2)$$

gdzie: x_{ik} , x_{jk} , k -te charakterystyki jednostki i oraz j .

Gdy przedmiotem badania jest bezrobocie charakterystyki we wzorze (2) mogą dotyczyć, z jednej strony, poziomu lub rodzaju wykształcenia ludności w poszczególnych rejonach, struktury wiekowej albo, z drugiej strony, wielkości produkcji przemysłowej, nakładów inwestycyjnych itp.

Nie jest jasne jak zdefiniować odległość przestrzenno-czasową. Dwuwymiarowa przestrzeń geograficzna jest wielostronna. Podobnie jest w wypadku k -wymiarowej przestrzeni ekonomicznej, która zresztą zwykle jest redukowana do przestrzeni dwuwymiarowej. Czas jest jednokierunkowy i w układzie *prze-strzeń – czas* nie jest on po prostu dodatkowym wymiarem. Ponadto, zwykle dane przestrzenno-czasowe odnoszą się do dużej liczby obserwacji „po czasie” i do niewielkiej liczby jednostek przestrzennych. Dlatego ciekawym rozwiązaniem jest propozycja odrębnego badania szeregów czasowych dotyczących poszczególnych jednostek przestrzennych a następnie włączenia do modelu struktury przestrzennej określonej przez odległość pomiędzy jednostkami.

4. Przestrzenno-czasowy model ekonometryczny

Zakłada się, że jednostka przestrzenna i zlokalizowana jest w punkcie s_i w przestrzeni euklidesowej R^n , w szczególności R^2 . Lokalizację i -tej jednostki w czasie t oznacza się jako (s_i, t) . Obserwuje się wartości zmiennej $Y_{s_i, t}$ w zbiorze lokalizacji $\{s_i, t\}_{i=1}^N$, dla $t=1, 2, \dots, T$. Obserwacje oznacza się przez $y_{s_i, t}$. Zmienna losowa $Y_{s_i, t}$ związana z przestrzenną pozycją s_i oraz punktem t nazywa się przestrzenno-czasowym polem losowym².

² Główne zastosowania pól losowych w ekonometrii dotyczą analiz przestrzennych i przestrzenno-czasowych. Zob., np. Conley (1999), Chen, Conley (2001), Conley, Topa (2002).

Dla każdej ustalonej lokalizacji s_i , otrzymuje się zmienną $Y_{\cdot,t}$, tj. proces stochastyczny. Stąd na $Y_{s_i,t}$ można patrzeć jak na wektor $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{Nt}\}$. Podobnie można rozważać zmienną $Y_{s_i,\cdot}$, tj. proces przestrzenny, otrzymując następnie $Y_{s_i} = \{Y_{s_i,1}, Y_{s_i,2}, \dots, Y_{s_i,N}\}$. $Y_{\cdot,t}$ można określić jako pole losowe warunkowe względem przestrzeni. Odpowiednio $Y_{s_i,\cdot}$ może być traktowane jako pole losowe warunkowe względem czasu.

Modele auto-zależności dla procesów stochastycznych (modele autoregresyjne) budowane są przy założeniu, że zależność maleje wraz ze wzrostem odległości między obserwacjami procesu (wartościami szeregu czasowego). Podobnie, w modelach auto-zależności dla procesów przestrzennych zasadą jest, że zależności te są charakteryzowane przez odległość pomiędzy przestrzennymi lokalizacjami obserwacji. Jeśli obserwacje i oraz j są „bliskie”, wtedy zmienne losowe Y_{s_i} oraz Y_{s_j} mogą być silnie skorelowane. Natomiast, gdy odległość pomiędzy s_i oraz s_j rośnie, wtedy Y_{s_i} oraz Y_{s_j} są coraz słabiej skorelowane lub niezależne. Ponieważ nie jest jasne jak mierzyć odległość przestrzenno-czasową (zob. punkt 3), określenie modelu auto-zależności przestrzenno-czasowych jest utrudnione.

Łączne modelowanie zjawisk przestrzenno-czasowych jako pól losowych wymaga rozwiązania również wielu innych problemów. Są to, na przykład, trudne do spełnienia w praktyce warunki: stacjonarności, jednorodności, izotropowości. Ponadto, zazwyczaj mała jest liczba jednostek przestrzennych, dla których dostępne są dane statystyczne.

W sytuacji, gdy mało jest danych zebranych w przestrzeni ale dużo w czasie, można badać szeregi czasowe oddzielnie dla każdej jednostki przestrzennej. Modelowanie szeregów czasowych oddzielnie dla każdego punktu w przestrzeni jest uzasadnione tylko w wypadku niezależnych czasowych pól losowych. W odpowiednich modelach bierze się pod uwagę zależności czasowe, ale nie uwzględnia się zależności przestrzennych.

W literaturze pojawiła się ciekawa propozycja włączenia do tego typu analiz zależności przestrzennych w drodze konstrukcji modeli z przestrzenną strukturą (zob. de Luna, Genton, 2004). Rozważa się wektor $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{Nt}\}$ w rozumieniu, jak wyżej. Buduje się modele niezależnie dla każdej wyodrębnionej jednostki przestrzennej, biorąc pod uwagę przestrzenno-czasowe uporządkowanie jednostek, po to aby ustalić hierarchię tych jednostek, według której wprowadza się elementy wektora Y_t do modelu. Podstawą uporządkowania przestrzennego jest odległość pomiędzy jednostkami określona, na przykład, przez rozważenie geograficznych lokalizacji tych jednostek. Ustala się naturalny porządek „sąsiedztwa” jednostek przestrzennych dla danej jednostki. Ostatecznie można mówić o macierzy „odległości” o wymiarach $N \times N$, określającej taki porządek dla wszystkich jednostek (zob., np. tabela 2).

Zmienne $Y_{s_i,t}$ wprowadza się do modelu sekwencyjnie, według naturalnego porządku przestrzenno-czasowego, tj. najpierw $Y_{s_i,t-l}$, w ustalonym porządku przestrzennym, dla $l=1$, następnie, znowu w ustalonym porządku przestrzennym, dla $l=2$, itd.

Powiązanych przestrzennie i czasowo „sąsiadów” danej jednostki ustala się na podstawie wartości korelacji cząstkowej, to znaczy nowe zmienne wprowadza się o modelu tak długo, jak długo wartości te istotnie różnią się od zera. Przestrzenno-czasowa korelacja cząstkowa oznacza tutaj korelację pomiędzy określonym elementem wektora Y_t a innym jego elementem w odstępnie czasowym l , przy wyłączeniu wpływu uwzględnionych wcześniej zmiennych.

5. Analiza danych empirycznych

Przeprowadzone badanie dotyczyło bezrobocia w Polsce według województw w okresie: styczeń 1999 – kwiecień 2005. Dane statystyczne pochodzą z Biuletynów statystycznych GUS, *Bezrobocie rejestrowane, Informacje i opracowania statystyczne*, Warszawa.

Jak argumentowano wcześniej, dane te mogą być przestrzennie skorelowane. Oznacza to, że poziom bezrobocia w danym województwie może być skorelowane z poziomem bezrobocia w województwach sąsiadujących. Próba statystyczna liczy 76 obserwacji czasowych dla każdej z 16 jednostek przestrzennych, tj. łącznie 1216 obserwacji. Te 16 szeregów czasowych tworzy realizacje procesów wchodzących w skład wektora definiującego proces przestrzenno-czasowy.

Najpierw zbadano strukturę trendowo-sezonową i autoregresyjną osobno dla każdego szeregu czasowego. Tabela 1 przedstawia wyniki badania w tym zakresie.

Tabela 1. Trend, sezonowość, autoregresja badanych szeregów czasowych

Województwo	Dln.	K.-p.	Lbl.	Lbs.	Łdz.	Młp.	Mzw.	Opl.
<i>Stopień trendu</i>	3	3	2	4	3	2	3	3
<i>Sezonowość</i>	+	+	+	+	+	+	+	+
<i>Rząd autoregresji</i>	1	8	1	1	1	8	1	1
Województwo	Pdk.	Pdl.	Pmr.	Śls.	Śwt.	W.-m.	Wlk.	Zch.
<i>Stopień trendu</i>	2	4	3	2	2	4	3	3
<i>Sezonowość</i>	+	+	+	+	+	+	+	+
<i>Rząd autoregresji</i>	9	10	1	12	8	12	1	11

Źródło: obliczenia własne.

„+” oznacza obecność składnika sezonowego.

Skróty nazw województw utworzono przez podanie pierwszych trzech spółgłosek z wyjątkiem nazw dwuczłonowych. Dla nich przyjęto: K.-p. (Kujawsko-pomorskie), W.-m. (Warmińsko-mazurskie).

Stosując klasyczną metodę najmniejszych kwadratów oszacowano parametry odpowiednich modeli podstawowych. Następnie zbadano reszty tych modeli pytając, czy tworzą one przestrzenno-czasowy proces białoszumowy? Obliczono współczynniki korelacji dla reszt szeregów czasowych. Większość z nich okazała się istotna na 5% poziomie istotności. Następnie obliczono współczynniki korelacji dla szeregów reszt z czasowym odstępem jeden. Również w tym wypadku niektóre z nich okazały się istotne. Współczynniki korelacji reszt z odstępami czasowymi większymi od jeden były już nieistotne.

Wniosek z przeprowadzonej analizy jest następujący. Badanie pojedynczych szeregów czasowych nie wystarczyło dla odkrycia pełnej struktury procesu bezrobocia. Istnieją powiązania przestrzenno-czasowe, które należałoby uwzględnić w odpowiednim modelu.

W celu uwzględnienia zależności przestrzenno-czasowych w badanym procesie bezrobocia wykorzystano informację o przestrzennej strukturze powiązań między województwami. Struktura ta została przedstawiona w tabeli 2. Każdy wiersz tabeli 2 przedstawia przestrzenne uporządkowanie (od 0 do 15) województw. Ten przestrzenny porządek „sąsiedztwa” dla każdego województwa został ustalony na podstawie średniej odległości fizycznej (w km) pomiędzy głównymi miastami województw.

Tabela 2. Przestrzenna struktura „odległości” pomiędzy województwami

	Dln	K-p	Lbl	Lbs	Ldz	Mjp	Mzw	Opl	Pdk	Pdl	Pmr	Sls	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
Ldz	9	3	8	13	0	7	2	6	10	12	14	5	1	11	4	15
Mjp	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
Sls	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

Źródło: obliczenia własne.

Badaniu poddano ponownie szeregi czasowe stóp bezrobocia po wyeliminowaniu trendu i sezonowości (dla zachowania stacjonarności). Przy założeniu ustalonego porządku powiązań przestrzennych dla każdego województwa policzono odpowiednie macierze współczynników korelacji szeregów czasowych dla odstępów czasowych od 1 do 12. Współczynniki szybko malały przy wzroście odstępów czasowych. Analogiczna prawidłowość dla „odstępów” przestrzennych była dużo mniej wyraźna. W dalszej kolejności wykorzystano współczynniki korelacji cząstkowej do wyboru dla każdego województwa przestrzennych „sąsiadów” istotnie powiązanych pod względem poziomu bezrobocia. Tabela 3 przedstawia ustalenia w tym zakresie.

Tabela 3. Struktura zależności przestrzenno-czasowych

odstęp	Dln	K-p	Lbl	Lbs	Ldz	Mlp	Mzw	Opl	Pdk	Pdl	Pmr	Śls	Śwt	Wm	Wlk	Zch
1	7	7	15	8	13	7	6	9	4	13	14	3	14	14	3	15
2	0	8	9	0	4	11	13	0	0	9	0	0	0	0	0	0
3	0	0	14	0	15	0	0	0	0	0	7	0	15	0	0	0

Źródło: obliczenia własne.

6. Podsumowanie

Celem badania było sprawdzenie czy przyjęty arbitralnie porządek przestrzenno-czasowych powiązań w układzie województw pozwoli wyspecyfikować logiczną, przekonującą strukturę przestrzennych i czasowych odstępów dla odpowiedniego ekonometrycznego modelu zmian bezrobocia.

Bardzo ważnym elementem proponowanej koncepcji metodologicznej jest prawidłowe określenie odległości pomiędzy jednostkami badania. W przeprowadzonej analizie ograniczono się do niedoskonałej w istocie miary odległości fizycznej. W efekcie uzyskane wyniki jedynie bardzo ogólnie potwierdzają wstępne hipotezy o naturze prawidłowości występujących w badanym zjawisku.

W dalszej kolejności zostanie przeprowadzone badanie z wykorzystaniem odpowiednio zdefiniowanej odległości ekonomicznej.

Literatura

- Cliff, A.D., Ord, J.K. (1981), *Spatial Processes*, 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, w: Lesage, J.P., Pace, R.K. (eds.), *Spatial and Spatiotemporal Econometrics, Advances in Econometrics*, vol. 17, Elsevier Ltd., Oxford, 279–294.