

Mariusz Doszyń

Zastosowanie metod ekonometrycznych do badania heterogeniczności obiektów

Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania 31/1, 83-93

2013

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Mariusz Doszyń

Uniwersytet Szczeciński

ZASTOSOWANIE METOD EKONOMETRYCZNYCH DO BADANIA HETEROGENICZNOŚCI OBIEKTÓW

Streszczenie

W artykule scharakteryzowano powiązania między parametrami trzech rodzajów modeli: dla danych przekrojowych, danych panelowych (model z ustalonymi efektami), danych w postaci szeregów czasowych. Omówiono ekonometryczne metody badania heterogeniczności obiektów. W przykładzie empirycznym proponowane w artykule procedury wykorzystano do oszacowania efektów indywidualnych charakteryzujących skłonność do konsumpcji ludności województw w Polsce w latach 1998–2005.

Słowa kluczowe: heterogeniczność obiektów, modele ekonometryczne dla danych przekrojowych, modele dla danych panelowych z ustalonymi efektami, modele szeregów czasowych.

Wprowadzenie

Dysponując danymi przekrojowymi lub danymi panelowymi, trzeba mieć na uwadze fakt, że analizowane obiekty różnią się od siebie w sposób, który często jest trudno zaobserwować bezpośrednio. Wiele zmiennych (cech) mających istotne znaczenie z punktu widzenia modelowanych procesów nie poddaje się bezpośredniej obserwacji. W przypadku osób do takich zmiennych można zaliczyć na przykład poziom inteligencji czy różnego rodzaju cechy o charakte-

rze psychologicznym. Heterogeniczność przedsiębiorstw może wynikać z różnic w gospodarności, skuteczniejszego zarządzania, większej innowacyjności czy lepszej motywacji pracowników.

Szczególną grupę przyczyn zróżnicowania obiektów tworzą skłonności ludzkie. Zgodnie z definicją prof. J. Hozera, skłonność to „nachylenie postawy” względem czegoś lub kogoś zwiększające prawdopodobieństwo pewnych zdarzeń [4; 2]. Takie skłonności, jak skłonność do konsumpcji, oszczędzania, inwestowania, ryzyka czy kooperacji, wpływają na różnice w zachowaniach, przyczyniając się tym samym do heterogeniczności obiektów, która powinna być uwzględniana w modelowaniu ekonometrycznym. Podobnie jak w przypadku szeregów czasowych uwzględnia się upływ czasu (przez wprowadzanie do modeli zmiennej czasowej lub badanie zintegrowania zmiennych), tak mając do czynienia z danymi przekrojowymi lub panelowymi, należy uwzględnić heterogeniczność obiektów.

1. Metodologia badania

Heterogeniczność obiektów uwidacznia się wtedy, gdy dysponuje się danymi przekrojowymi lub panelowymi. W przypadku danych panelowych obserwacje dotyczą tych samych obiektów analizowanych w określonym przedziale czasu. W modelach budowanych na podstawie szeregów czasowych wpływ czynników specyficznych dla danego obiektu jest zazwyczaj stały, a tym samym trudny do wyodrębnienia.

Jeżeli analizowanych jest n obiektów w określonym okresie (momencie) t , to model liniowy (dla danych przekrojowych) można zapisać następująco:

$$y_{it} = \alpha_t + \sum_{l=1}^k \beta_l x_{l,it} + c_{it} + u_{it} \quad (1)$$

gdzie:

- y_{it} – wartość zmiennej objaśnianej w i -tym obiekcie w badanym okresie t ,
- α_t, β_l – parametry modelu, $l = 1, \dots, k$, k – liczba zmiennych objaśniających, α_t – wyraz wolny dla rozważanego okresu,
- $x_{l,it}$ – wartość l -tej zmiennej objaśniającej w i -tym obiekcie w okresie t ,

- c_{it} – wpływ czynników specyficznych dla i -tego obiektu w danym okresie t ,
- u_{it} – składnik losowy odpowiadający obserwacji w i -tym obiekcie w okresie t .

W modelu (1) heterogeniczność obiektów przejawia się w bezpośrednio nieobserwowalnym parametrze c_{it} . Obrazuje on wpływ czynników specyficznych dla danego obiektu, innych niż te, które zostały uwzględnione bezpośrednio w modelu w postaci zmiennych objaśniających $x_{l,it}$.

Jednym ze sposobów pozwalających na uwzględnienie zróżnicowania obiektów jest wprowadzanie do zbioru zmiennych objaśniających zmiennej zero-jedynkowej d_z , która w danym równaniu ma wartość równą jedności dla i -tego obiektu oraz zera dla pozostałych:

$$d_z = \begin{cases} 1, & \text{gdy } z = i \\ 0, & \text{gdy } z \neq i \end{cases} \quad (2)$$

gdzie $i = 1, 2, \dots, n$.

Ponieważ przy wprowadzaniu kolejnych zmiennych d_z zmieniają się oceny parametrów, dla każdego okresu t można oszacować n modeli o postaci:

$$y_{it} = \alpha_{it} + \sum_{l=1}^k \beta_l x_{l,it} + \delta_{zt} d_z + u_{it} \quad (3)$$

gdzie $z = 1, 2, \dots, n$.

Dla każdego okresu t szacuje się n modeli, a zatem każdemu t odpowiada n ocen każdego z parametrów modelu (3). Na podstawie modelu (3) uzyskuje się więc n ocen parametru δ_{zt} dla danego okresu t . Oceny te odzwierciedlają wpływ czynników charakterystycznych dla badanych obiektów¹. Obliczone w ten sposób oceny parametrów przy zmiennych zero-jedynkowych to tak zwane reszty predyktywne, a wyznaczone dla tych ocen wartości statystyk t -Studenta to z kolei reszty studentyzowane.

Na marginesie warto dodać, że na podstawie modeli typu (3) można również bezpośrednio analizować wpływ różnego rodzaju skłonności na zjawiska gospodarcze. W takim przypadku zmienna $d_z = 1$, jeśli obiekt wykazuje daną

¹ Innym sposobem umożliwiającym uwzględnienie heterogeniczności obiektów w modelach dla danych przestrzennych jest wprowadzenie do modelu zmiennych zastępczych (*proxy variables*). Zob. [6].

skłonność, a zero, gdy obiekt nie ma skłonności. To, czy obiekt (osoba, grupa społeczna) cechuje się badaną skłonnością, można ustalić na przykład metodą ankietową. W tym podejściu potrzebna jest dodatkowa wiedza, zazwyczaj o charakterze psychologicznym, a zmienna d_z może być równa 1 w więcej niż jednym przypadku.

Wracając do modelu (3), uwzględnienie wszystkich n zmiennych zero-jedynkowych, których wartość wynosi 1 dla danego obiektu, w zbiorze zmiennych objaśniających w równaniu (3) jest niemożliwe dlatego, że liczba zmiennych objaśniających byłaby większa od liczby obserwacji ($n + k + 1 > n$). Tego typu problemy nie występują w wypadku danych panelowych, a więc odnoszących się do tych samych obiektów obserwowanych w przynajmniej kilku okresach (momentach) [3; 6].

Model dla danych panelowych z ustalonymi efektami można zapisać jako:

$$y_{it} = \sum_{l=1}^k \beta_l x_{l,it} + \alpha_i + \varepsilon_{it} \quad (4)$$

gdzie $\varepsilon_{it} \sim IN(0, \sigma^2)$.

W modelu (4) każdemu obiektowi odpowiada inny wyraz wolny α_i , który uwzględnia heterogeniczność obiektów. Parametr α_i odzwierciedla czynniki specyficzne dla danego obiektu. Przyjmuje się założenie, że czynniki te są stałe w czasie.

Przy właściwej konstrukcji modelu czynniki specyficzne mogą odzwierciedlać wpływ różnego rodzaju skłonności na badane zjawiska. Przykładowo, jeżeli model (4) jest funkcją oszczędności uwzględniającą w zbiorze zmiennych objaśniających $x_{l,it}$ wszystkie istotne czynniki obiektywne (dochód, stopę procentową itp.), to można postawić hipotezę, że efekty indywidualne α_i odzwierciedlają wpływ skłonności do oszczędzania.

Do precyzyjnego określenia wpływu skłonności do oszczędzania potrzebna jest wiedza (głównie o charakterze psychologicznym) na temat tego, czy dana osoba wykazuje skłonność do oszczędzania. Jeżeli tak, wartość zmiennej wynosi jeden, jeżeli nie – zero. Wprowadzając do modelu zmienne zero-jedynkowe można bezpośrednio określić wpływ analizowanych skłonności. Często jednak ekonometryk nie dysponuje tego typu wiedzą. W takim przypadku można oszacować efekty indywidualne, które przy właściwej specyfikacji modelu,

czyli po uwzględnieniu wszystkich istotnych zmiennych o charakterze obiektywnym, odzwierciedlają wpływ danej skłonności².

Model dla danych panelowych z ustalonymi efektami można przedstawić również w postaci macierzowej:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad (5)$$

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \dots \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \dots \\ \mathbf{X}_n \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \mathbf{i}_T & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{i}_T & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{i}_T \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \dots \\ \boldsymbol{\varepsilon}_n \end{bmatrix} \quad (6)$$

gdzie:

- \mathbf{y} – wektor o wymiarze $nT \times 1$, gdzie n – liczba obiektów, a T – liczba okresów,
- \mathbf{X} – macierz o wymiarze $nT \times k$, składająca się z obserwacji dokonanych na zmiennych $x_{l,it}$ dla kolejnych obiektów,
- \mathbf{D} – macierz blokowo-diagonalna o wymiarze $nT \times n$, w której na głównej przekątnej znajdują się wektory (\mathbf{i}_T) o wymiarze $T \times 1$, składające się z jedynek,
- $\boldsymbol{\alpha}$ – wektor o wymiarze $n \times 1$ składający się z wyrazów wolnych dla kolejnych obiektów,
- $\boldsymbol{\varepsilon}$ – wektor zakłóceń losowych o wymiarze $nT \times 1$.

Zmienne zero-jedynkowe zawarte w macierzy \mathbf{D} pozwalają na uwzględnienie wpływu czynników charakterystycznych dla danego obiektu. W omawianych modelach można również uwzględnić wpływ czasu, na przykład przez wprowadzenie do modelu zmiennej czasowej t .

Model (5) można przedstawić w dwojaki sposób. W postaci (6) obserwacje pogrupowano tak, że powstały szeregi czasowe dla kolejnych obiektów³. Dla i -tego obiektu model (6) ma postać:

² Za czynniki obiektywne uznaje się te, które bezpośrednio nie zależą od właściwości (struktury) obiektu (osoby, zbiorowości, przedsiębiorstwa *ect*). W przypadku ludzi są to czynniki wynikające z różnych cech osobowości, możliwości, wyznawanych norm, stylu życia *ect*, które często konstrytuują różnego rodzaju skłonności.

³ Jest to tzw. stos szeregów czasowych (*stacked time series*). Inną możliwością jest uporządkowanie obserwacji tak, że mamy do czynienia z danymi przekrojowymi dla kolejnych okresów. Jest to tzw. stos danych przekrojowych (*stacked cross sections*).

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{i}_T \alpha_i + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \dots, n \quad (7)$$

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} x_{1,i1} & x_{2,i1} & \dots & x_{k,i1} \\ x_{1,i2} & x_{2,i2} & \dots & x_{k,i2} \\ \dots & \dots & \dots & \dots \\ x_{1,iT} & x_{2,iT} & \dots & x_{k,iT} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \alpha_i + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \dots \\ \varepsilon_{iT} \end{bmatrix} \quad (8)$$

W modelach (7) i (8) występują szeregi czasowe zmiennych dla i -tego obiektu. Takie przedstawienie zmiennych kolejno dla wszystkich obiektów daje w rezultacie model (6). Czynniki specyficzne przejawiają się w wyrazie wolnym α_i . Można zauważyć, że w modelach dla danych czasowych stałe w czasie czynniki specyficzne dla danego i -tego obiektu przejawiają się w wyrazie wolnym α_i . W modelach szeregów czasowych jest jeden „efekt indywidualny” w postaci wyrazu wolnego, z założeniem, że wpływ czynników specyficznych jest stały w czasie.

Analizując model dla danych przekrojowych (3) oraz model dla danych panelowych (6), można wykazać, że w modelu (6) efekty indywidualne dla danego obiektu są średnią z sumy wyrazów wolnych α_{it} oraz parametrów δ_{zt} (z założeniem, że wpływ zmiennych objaśniających $x_{i,zt}$ w każdym rozważanym modelu jest identyczny):

$$\alpha_i = \frac{\sum_{t=1}^T (\alpha_{it} + \delta_{zt})}{T} = \bar{\alpha}_i + \bar{\delta}_z \quad (9)$$

gdzie $i = 1, 2, \dots, n$ oraz $z = i$.

Jak łatwo zauważyć, $\bar{\alpha}_i = 1/T \sum_{t=1}^T \alpha_{it}$ to średni wyraz wolny w modelach dla danych przekrojowych dla i -tego obiektu w analizowanym okresie, a $\bar{\delta}_z = 1/T \sum_{t=1}^T \delta_{zt}$ – średnia reszta predyktywna dla danego obiektu. Reszty predyktywne odpowiadają parametrom przy zmiennych zero-jedynkowych $d_{z,t}$.

Jak to wcześniej przedstawiono, stosując modele dla danych przekrojowych uzyskuje się n ocen parametru dla każdego okresu t . Tym samym dla każdego obiektu otrzymuje się T ocen parametru α_{it} oraz δ_{zt} . Suma uśrednionych (względem czasu) ocen parametrów daje efekty indywidualne w modelu dla

danych panelowych (6). Należy przy tym pamiętać, że zależność (9) jest spełniona wówczas, gdy wpływ zmiennych objaśniających $x_{i,t}$ w modelach dla danych przestrzennych i modelu dla danych panelowych jest taki sam.

Model (6) to model szeregów czasowych dla n obiektów. Można więc oszacować oddzielnie dla każdego obiektu model o postaci (7). Przyjmując, że wpływ zmiennych objaśniających w modelu (6) i modelach typu (7) dla kolejnych obiektów jest identyczny, efekty indywidualne w modelu dla danych panelowych będą odpowiadały wyrazom wolnym w modelach szeregów czasowych dla poszczególnych obiektów.

Reasumując, można zauważyć, że efekty indywidualne w poszczególnych rodzajach modeli są ze sobą ściśle powiązane. Pojawia się pytanie: czy wyznaczać je na podstawie modeli dla danych przekrojowych (dla kolejnych okresów), modeli w postaci szeregów czasowych (dla kolejnych obiektów) czy na podstawie modeli panelowych *sensu stricto*?

Przewaga modeli panelowych polega na tym, że są one szacowane z wykorzystaniem znacznie większej liczby obserwacji, których jest nT . Jeżeli estymatory parametrów są zgodne, to szacunki parametrów mają większą precyzję. W modelach przekrojowych liczba obserwacji dla każdego okresu t jest równa n , z kolei w modelach szeregów czasowych występuje T obserwacji dla każdego obiektu i .

Szacowanie modeli dla danych panelowych jest uzasadnione wtedy, gdy oceny parametrów przy zmiennych objaśniających nie zmieniają się istotnie z okresu na okres (w modelach dla danych przestrzennych) lub przy szacowaniu modeli dla kolejnych obiektów (w modelach szeregów czasowych).

Przyjmując następujące hipotezy zerowe:

$$H_0 : \beta_{.1} = \beta_{.2} = \dots = \beta_{.T} = \beta,$$

$$H_0 : \beta_{1.} = \beta_{2.} = \dots = \beta_{n.} = \beta,$$

gdzie:

$\beta_{.1}, \beta_{.2}, \dots, \beta_{.T}$ – wektory parametrów przy zmiennych objaśniających w modelach dla danych przestrzennych dla $t = 1, 2, \dots, T$,

$\beta_{1.}, \beta_{2.}, \dots, \beta_{n.}$ – wektory parametrów w modelach szeregów czasowych dla $i = 1, 2, \dots, n$,

β – wektor parametrów w modelu dla danych panelowych.

Jeżeli nie ma podstaw do odrzucenia hipotez zerowych, to lepsze własności mają estymatory parametrów w modelu dla danych panelowych. Jeśli hipotezę zerową należy odrzucić, bardziej zasadne jest szacowanie modeli dla danych przekrojowych lub modeli szeregów czasowych⁴.

2. Przykład empiryczny

W przykładzie empirycznym podjęto próbę rozstrzygnięcia, czy lepszym wyjściem jest oszacowanie efektów indywidualnych na podstawie modelu z ustalonymi efektami, czy na podstawie modeli szeregów czasowych estymowanych oddzielnie dla poszczególnych obiektów. Wykorzystano dane dotyczące kształtowania się wydatków na towary i usługi konsumpcyjne w zł na osobę (y_{it}) oraz dochodów do dyspozycji ludności w zł na osobę (x_{it}) w poszczególnych województwach w Polsce w latach 1998–2005. Zmienne urealniono i wyrażono w cenach na poziomie z 2005 roku. Dane pochodzą z Rocznika statystycznego województw GUS.

Model z ustalonymi efektami po oszacowaniu ma następującą postać⁵:

$$\hat{y}_{it} = \hat{\alpha}_{pi} + 0,482 x_{it}, S_e = 21,570, \bar{R}^2 = 0,883 \quad (10)$$

(7,163)

Oceny efektów indywidualnych ($\hat{\alpha}_{pi}$) modelu (10) podano w pierwszej kolumnie tabeli 1. W kolejnych dwóch kolumnach zamieszczono oceny wyrazów wolnych ($\hat{\alpha}_i$) oraz oceny parametrów przy zmiennej objaśniającej ($\hat{\beta}_i$) otrzymane po oszacowaniu modeli szeregów czasowych oddzielnie dla każdego województwa.

W celu rozstrzygnięcia, czy do oceny efektów indywidualnych lepiej zastosować model dla danych panelowych z ustalonymi efektami, czy modele szeregów czasowych, zweryfikowano hipotezy, zakładające, że oceny parametrów przy zmiennej objaśniającej w modelach szeregów czasowych dla poszczególnych województw ($\hat{\beta}_i$) nie różnią się istotnie statystycznie od oceny parametru $\hat{\beta} = 0,482$ przy analogicznej zmiennej w modelu z ustalonymi efek-

⁴ Do zweryfikowania powyższych hipotez można zastosować na przykład test bazujący na rozkładzie t-Studenta lub rozkładzie F .

⁵ W nawiasie pod oceną parametru znajduje się wartość statystyki t-Studenta.

tami (10). W tym celu posłużono się testem t-Studenta. Empiryczne poziomy istotności ujęto w ostatniej kolumnie tabeli 1. Przy poziomie istotności $\alpha = 0,05$ we wszystkich przypadkach nie ma podstaw do odrzucenia hipotezy zerowej. Zastosowanie modelu dla danych panelowych z ustalonymi efektami jest zatem lepszym sposobem szacowania efektów indywidualnych. Wynika to z tego, że liczba obserwacji jest większa, a tym samym – dokładniejsze oceny parametrów.

Przyjmując roboczą hipotezę, że w modelu (10) uwzględniono główne, istotne czynniki kształtujące wydatki konsumpcyjne województw, można stwierdzić, że efekty indywidualne są powiązane ze skłonnością do konsumpcji. Jak widać, największą skłonnością do konsumpcji cechowali się mieszkańcy województwa mazowieckiego, a najmniejszą – województwa warmińsko-mazurskiego.

Tabela 1. Oceny efektów indywidualnych otrzymane na podstawie modelu dla danych panelowych ($\hat{\alpha}_{pi}$), oceny parametrów modeli szeregów czasowych dla poszczególnych województw oraz empiryczne poziomy istotności w teście na brak różnic między parametrami modeli szeregów czasowych (β_i) a parametrem przy zmiennej objaśniającej w modelu z ustalonymi efektami

Województwo	$\hat{\alpha}_{pi}$	$\hat{\alpha}_i$	$\hat{\beta}_i$	P_{emp}
Dolnośląskie	337,631	586,253	0,132	0,485
Kujawsko-pomorskie	288,451	133,211	0,724	0,515
Lubelskie	292,596	177,698	0,674	0,177
Lubuskie	347,660	522,770	0,218	0,772
Łódzkie	344,734	253,869	0,611	0,681
Małopolskie	321,284	362,118	0,421	0,680
Mazowieckie	375,102	182,453	0,707	0,194
Opolskie	343,285	265,291	0,597	0,668
Podkarpackie	287,679	183,747	0,665	0,671
Podlaskie	305,163	437,532	0,266	0,096
Pomorskie	319,641	276,219	0,541	0,833
Śląskie	345,221	589,060	0,148	0,619
Świętokrzyskie	287,419	408,454	0,276	0,601
Warmińsko-mazurskie	281,736	452,408	0,207	0,161
Wielkopolskie	299,968	271,040	0,524	0,880
Zachodniopomorskie	335,419	430,025	0,347	0,710

Źródło: obliczenia własne.

Podsumowanie

Reasumując, w artykule przedstawiono powiązania między trzema rodzajami modeli: modelami dla danych przekrojowych, modelami dla danych panelowych z ustalonymi efektami oraz modelami szeregów czasowych. Wykazano, że efekty indywidualne w modelach dla danych panelowych są jednocześnie wyrazami wolnymi w modelach szeregów czasowych (z założeniem, że wpływ zmiennych objaśniających w każdym omawianym modelu jest taki sam). Indywidualne efekty w modelach dla danych panelowych to średnie wartości sumy ocen wyrazów wolnych oraz reszt predyktywnych w modelach dla danych przestrzennych, gdzie wartości te uśredniają się w czasie (również z założeniem, że wpływ zmiennych objaśniających jest identyczny).

W przykładzie empirycznym zastosowano zaproponowaną w artykule procedurę do oszacowania efektów indywidualnych charakteryzujących specyfikę (skłonność do konsumpcji) województw w Polsce w latach 1998–2005.

Literatura

1. Doszyń M., *Analiza wpływu skłonności na zjawiska ekonomiczne na podstawie modeli ekonometrycznych dla danych panelowych*, Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania nr 11, Szczecin 2008.
2. Doszyń M., *Statystyczno-ekonometryczna analiza skłonności ludzkich*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2008.
3. Greene W., *Econometric analysis*, fifth edit., Prentice Hall, New York 2003.
4. Hozer J., Doszyń M., *Ekonometria skłonności*, PWE, Warszawa 2004.
5. Hozer J., Zawadzki J., *Zmienna czasowa i jej rola w badaniach ekonometrycznych*, PWN, Warszawa 1990.
6. Wooldridge J.M., *Econometric Analysis of Cross Section and Panel Data*, MIT, Massachusetts 2002.

APPLICATION OF ECONOMETRIC METHODS IN ANALYZING OBJECTS' HETEROGENEITY

Summary

In the article connections between three types of models were introduced: models for cross section data, models for panel data and models of time series. Possibilities of analyzing objects' heterogeneity by means of presented models were discussed. In the empirical example procedure proposed in the article was used to estimate individual effects describing propensity to consume in voivodships in Poland in years 1998–2005.

Keywords: object's heterogeneity, models for cross section data, panel data models with fixed effects.

Translated by Mariusz Doszyń