

## UWAGI DOTYCZĄCE WYKORZYSTANIA I STOSOWANIA METOD EKONOMETRYCZNYCH W BADANIACH EKONOMICZNYCH

*Bolesław Borkowski\**, *Stanisław Stańko\*\**

\*Katedra Ekonometrii i Statystyki Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie

\*\*Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych  
Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie

\*Kierownik: dr hab. Zbigniew Binderman, prof. nadzw. SGGW

\*\*Kierownik: prof. dr hab. Henryk Manteuffel

Słowa kluczowe: ekonomia, ekonometria, model, korelacja, regresja

*Key words: economics, econometrics, econometric models, correlation coefficients, regression*

S y n o p s i s. W opracowaniu przedstawiono podstawowe zagadnienia dotyczące wykorzystania w badaniach ekonomicznych stacjonarnych i dynamicznych modeli ekonometrycznych. Omówiono etapy postępowania w procesie budowy modeli ekonometrycznych oraz kwestie interpretacji otrzymanych wyników. Szczegółowo zaprezentowano problemy dotyczące selekcji cech, wyboru postaci analitycznej modelu, oceny stopnia dopasowania modelu do danych empirycznych oraz merytorycznej i formalnej weryfikacji otrzymanych wyników. Przedstawiono również niektóre niebezpieczeństwa wynikające z mechanicznego stosowania tych metod i interpretacji uzyskanych wyników.

### WSTĘP

W otaczającej nas przyrodzie działają określone prawidłowości. Ludzie poznają ich charakter i na tej podstawie formułują słowne określenia ich istoty, nazywane prawami nauki. Część tych prawidłowości jest przedmiotem badań ekonomii. Dotyczą one podstawowego problemu ekonomicznego, czyli zagadnień wyboru: co wytwarzać, w jaki sposób wytwarzać, dla kogo wytwarzać. Wynika to z faktu, że zasoby są ograniczone w swej wielkości, a ludzkie potrzeby są nieograniczone. Ekonomia jest więc nauką badającą, jak ludzie dają sobie radę z rzadkością (brakiem nieograniczonej dostępności dóbr) oraz jak rozwiązują problem alokacji ograniczonych zasobów w celu zaspokojenia stale rosnących potrzeb konsumpcyjnych. Rzeczywistość ekonomiczna jest jednak niezmiernie złożona i skomplikowana. Ogromna liczba zdarzeń gospodarczych i czynników je kształtujących, tempo zmian warunków gospodarowania, różnorodność decyzji ekonomicznych powodują, że bezpośrednio formułowanie na tej podstawie prawidłowości jest utrudnione. W świecie, w którym żyjemy, panuje jednak pewien porządek polegający na tym, że zdarzenia powiązane są różnymi zależnościami, oraz na tym, że zależności te podlegają pewnym prawidłowościom.

Prawidłowości i zależności mogą być różnego typu: mieć charakter funkcyjny, przyczynowo-skutkowy, bezpośredni, pośredni, pierwotny, wtórny, główny, uboczny itp. Przedmiotem ekonomii jest wyodrębnienie obiektywnych prawidłowości występujących w realnej rzeczywistości gospodarczej. Określenie tego typu związków i zależności określa się jako prawa ekonomiczne.

Prawidłowości rządzące zdarzeniami nie tylko determinują ich konieczność, ale i pozwalają na pewną ich przypadkowość. Dotyczy to zwłaszcza działalności gospodarczej człowieka. Klasycznym przykładem tych związków są zależności między nakładami a wynikami produkcji, opisywane za pomocą funkcji produkcji. W produkcji roślinnej plony roślin w poszczególnych latach wynikają nie tylko z poziomu i struktury nakładów, ale i z działania czynników przyrodniczo-klimatycznych, które zakłócają ściśle zależności między nakładami a efektami.

Cechą praw ekonomicznych, podobnie jak praw przyrody, jest ich obiektywny charakter. Oznacza to, że działają one niezależnie od tego, czy ludzie zdają sobie sprawę z ich istnienia, czy nie. Drugą ich cechą jest to, że mają charakter praw stochastycznych, tzn. występują przy masowym powtarzaniu się danego rodzaju zdarzeń. Stochastyczny charakter zjawisk ekonomicznych powoduje niemożność ich pełnego opisu, dlatego by ułatwić zrozumienie złożonych sytuacji, tworzy się modele. *Konstruuje się je po to, by zrozumieć mechanizm danego zjawiska w krótszym czasie i mniejszym koszcie* [Neal, Shone 1982, s. 35]. Analiza modelowa jest najbardziej rozpowszechnionym i podstawowym narzędziem analizy ekonomicznej. Przez model rozumie się *schematyczne uproszczenie, pomijające nieistotne aspekty w celu wyjaśnienia wewnętrznego działania, formy lub konstrukcji bardziej skomplikowanego mechanizmu* [Klein 1982, s. 15]. Podobne definicje modeli możemy spotkać w wielu innych publikacjach, z których wynika, że model stanowi uproszczony opis rzeczywistości. Wśród modeli ekonomicznych wyróżnia się modele opisowe i matematyczno-cybernetyczne [Marciniak 2005, s. 32].

Celem opracowania jest przedstawienie niektórych uwag dotyczących stosowania metod ekonometrycznych w badaniach ekonomicznych.

#### ISTOTA BADAŃ EKONOMETRYCZNYCH

Na początku należy postawić pytanie, czy do analizy związków pomiędzy zjawiskami gospodarczymi i do weryfikacji praw ekonomicznych muszą być stosowane metody ekonometryczne.

Analizy ekonomiczne nie należą do badań łatwych. Złożoność tych zjawisk, brak możliwości przeprowadzenia eksperymentu kontrolowanego, nadmiar informacji i ich jakość (zbieranie danych przez osoby trzecie, krótkie okresy badawcze, zmienność danych i ich niedostępność, sezonowość, duży poziom agregacji, brak stacjonarności) mogą mieć wpływ na końcowy efekt. Spośród wielu informacji niezbędnych do sformułowania prawidłowości ekonomicznych potrzebne są zwłaszcza te, które pozwolą, najlepiej w sposób mierzalny, na wyodrębnienie tych najistotniejszych. Należy dodać także i to, że rzeczywistość ekonomiczna jest bardziej złożona od fizycznej – w jej opisie stosuje się zazwyczaj modele formalne (matematyczne) ujmujące powiązania o charakterze ilościowym. Związki ujęte za pomocą modeli stanowiących sformalizowany opis rzeczywistości spełniają trzy zasadnicze funkcje: analityczną, diagnostyczną i prognostyczną [Hozer 1993, s. 7].

W ostatnich latach szeroko rozwinęła się gałąź nauki o metodach badań ilościowych prawidłowości zachodzących między zjawiskami ekonomicznymi. Gałąź tej wiedzy to ekonometria, o której ogólnie możemy powiedzieć, że jest dziedziną ekonomii wykorzystującą modele i sposoby wnioskowania statystyki matematycznej do opisu i analizy zjawisk ekonomicznych [Osińska 2006]. Z definicji ekonometrii wynika, że jest ona połączeniem teorii ekonomii, ekonomii matematycznej, statystyki i matematyki. Ekonometria dosłownie oznacza „mierzenie w ekonomii”. W literaturze naukowej możemy spotkać wiele definicji. W pracy przytoczymy dwie z nich – naszym zdaniem najtrafniejsze. Według Gangadharrao S. Maddali [2006b, s. 31] ekonometria to: *zastosowanie metod statystycznych i matematycznych do analizy danych ekonomicznych w celu nadania teoriom ekonomicznym kontekstu empirycznego oraz ich potwierdzenia lub odrzucenia*. Według Jana Tinbergena [1951, s. 12-17] *pierwszą rzeczą w każdym konkretnym zastosowaniu jest podanie poprawnej analizy ekonomicznej badanej zależności. O dwóch rzeczach należy koniecznie pamiętać: niezbędne są, po pierwsze, świadomość tego, jaka relacja nas naprawdę interesuje, po drugie – znajomość czynników występujących w tej relacji*. Podstawowym narzędziem ekonometrii jest model ekonometryczny (punktem wyjścia do budowy modelu jest istniejąca teoria lub hipoteza ekonomiczna). Pojęcie modelu ekonometrycznego jest różnie definiowane [Pawłowski 1971, Welfe, Welfe 1996, Osińska 2006]. Klasycznie model ekonometryczny stanowi równanie lub system równań, które opisują relacje między ekonomicznymi i nieekonomicznymi zmiennymi losowymi.

Związki pomiędzy zjawiskami ekonomicznymi są zazwyczaj bardzo skomplikowane i wielostronne. Na jakieś zjawisko wpływa wiele innych zjawisk i czynników zarówno o charakterze ekonomicznym, jak i pozaekonomicznym. W wielu przypadkach spotkać się można z wielostronnymi powiązaniem przyczynowymi szeregu zmiennych, przy czym odpowiednie relacje przyczynowe nie występują zwykle w postaci czystej, lecz są zakłócane przez oddziaływanie czynników ubocznych, a także nieistotnych i przypadkowych. Praktycznie nie jest możliwe uwzględnienie wszystkich powiązań, w związku z tym reprezentacja jakiegokolwiek zjawiska przy pomocy modelu powinna być taka by:

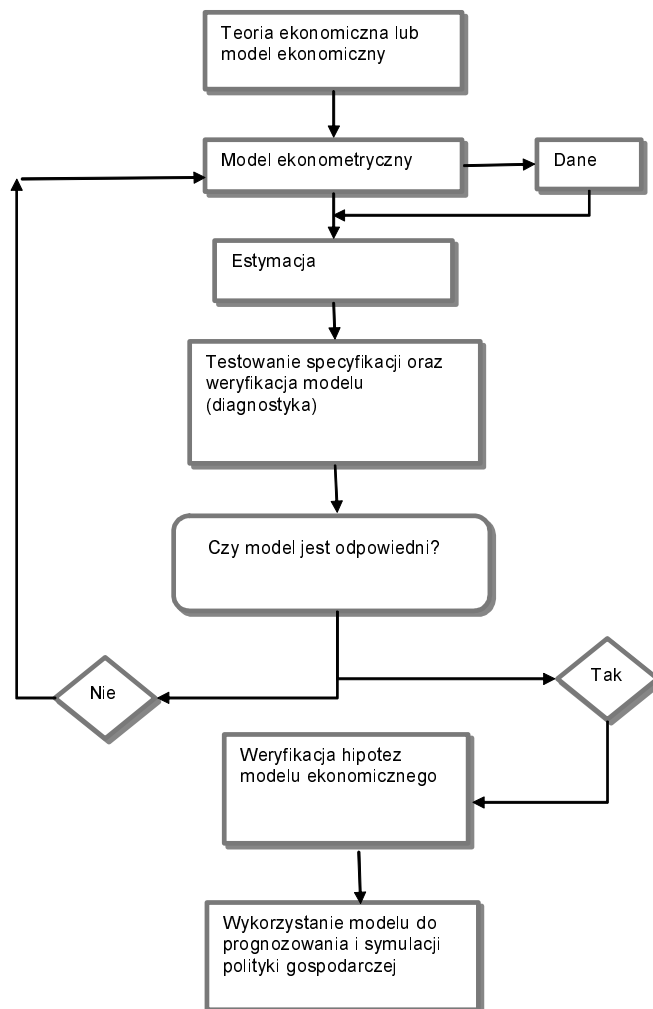
- w sposób realistyczny uwzględniać główne elementy reprezentowanego zjawiska,
- można było wyciągnąć wnioski o reprezentowanym zjawisku,
- model dostatecznie dokładnie uwzględniał związki między elementami.

Budowa modeli ekonometrycznych oznacza więc formułowanie modeli ekonomicznych w postaci nadającej się do weryfikacji empirycznej. Czy należy dążyć do budowy prostych modeli ekonometrycznych, czy bardziej złożonych? Wśród ekonometryków zdania są podzielone. Łatwość zrozumienia i weryfikacji empirycznej sugeruje budowę prostych modeli [Friedman 1980, s. 14-15]. Jednak pogląd ten jest krytykowany ze względu na nadmierne uproszczenie rzeczywistości i konieczność przyjęcia nierealistycznych założeń [Koopmans 1957, Klein 1962].

Sposób postępowania w procesie badawczym, w którym konstruujemy i wykorzystujemy model ekonometryczny nie jest dowolny. Przedstawiamy go za Maddalą [2006b] na rys. 1. Schemat ten nie powinien być interpretowany dosłownie, spełnia on jedynie rolę ilustracyjną.

W rozwoju ekonometrii jako dyscypliny naukowej w dziedzinie nauk ekonomicznych można wyróżnić kilka znaczących okresów:

1. Okres badań Frischa Regnara – prawdopodobny twórca pojęcia ekonometria (modele pomiaru elastyczności popytu i badania cykli koniunkturalnych około 1940 roku),



Rysunek 1. Schemat postępowania w ekonometrycznej analizie modeli ekonomicznych  
 Źródło: Maddala G. S. [2006b, s. 32-39].

2. Lata 30-50. XX w. – rozwój tradycyjnej ekonometrii. Komisja Cowlesa ds. Badań w Ekonomii – podjęcie prac nad określaniem cen giełdowych. W skład tej komisji wchodził m.in. Haavelmo, Hurwicz, Klein, Koopmans i Wald z USA oraz Tinbergen i Frish z Europy. Punkt ciężkości przesunięto na testowanie teorii oraz ustalono wskazówki dotyczące budowy modelu i jego testowania. Nastąpił rozwój dużych modeli makroekonomicznych (model Kleina i Goldbergera – 20 równań, Tinbergena – 24 równania, itp.),
3. Lata 50-70. XX w. – kryzys ekonometrii. Rozwój techniki komputerowej sprzyjał tworzeniu nowych algorytmów obliczeniowych i komplikacji postaci modeli ekonometrycznych. Tworzono nowe teorie efektywnej estymacji danego modelu w poszukiwaniu wysokiego  $R^2$ . Możliwości predykcyjne wielkich modeli makroekonomicznych były kiepskie. W większości przypadków naiwne prognozy były lepsze od otrzymanych na podstawie modeli ekonometrycznych.

4. Koniec XX w. – rozwój nowej ekonometrii. Przesunięcie akcentu na testowanie specyfikacji, diagnostykę (weryfikację) i porównywanie modeli. Rozwój ekonometrii szeregów czasowych i modeli opartych na danych panelowych oraz ekonometrii przestrzennej. Pojawiło się wiele nowych modeli, tj. modele racjonalnych oczekiwań, autoregresji wektorowej, ARCH, GARCH, i innych. Nowe testy autokorelacji, pierwiastki jednostkowe, kointegracja, przyczynowość itp.

NAJCZĘSTSZE PROBLEMY BADAWCZE Z ZAKRESU BADAŃ ILOŚCIOWYCH  
W EKONOMII

W pierwszej kolejności napotykamy na problem wstępnej selekcji cech. W badaniach ekonomicznych prawie zawsze mamy do czynienia z nadmiarem informacji, najczęściej dotyczących krótkich okresów, rzadko porównywalnych ze sobą, o różnych mianach czy charakteryzujących się różnorodnymi właściwościami. Selekcje cech możemy przeprowadzać różnymi metodami. Do najbardziej znanych zaliczane są metody wzorcowe i bezwzorcowe, wybór zmiennej syntetycznej, standaryzacja lub unitaryzacja itp. W literaturze naukowej do grupowania i porządkowania obiektów poleca się wiele metod wielowymiarowej analizy danych [Cieślak 1978, Kukuła 2000, Malina 2004, Zeliaś 2000 i inni]. Pominiemy szczegółową analizę tych metod, a czytelnika odsyłamy do literatury z zakresu taksonomii lub analizy danych wielowymiarowych.

Kolejnym problemem jest wybór postaci analitycznej modelu ekonometrycznego. W badaniach ekonomicznych często przyjmuje się, że model ekonometryczny opisujący relacje pomiędzy zmiennymi jest z góry znany, a główny problem polega na uzyskaniu dobrych oszacowań parametrów tego modelu. W rzeczywistości jednak niemal zawsze musimy dokonać wyboru właściwej postaci analitycznej funkcji. Ustalenie modelu matematycznego powinno poprzedzać określenie teorii lub wstępnej hipotezy badawczej. Następnym etapem badań jest określenie właściwej funkcji zależności (stochastycznej) o ogólnej postaci:

$$y_t = f(x_1, x_2, x_3, \dots, x_k, \varepsilon_t)$$

Jedno- i wielorównaniowe modele ekonometryczne ogólnie można podzielić na:

- a) liniowe,  
b) nieliniowe:
- względem zmiennych objaśniających,
  - względem zarówno zmiennych objaśniających, jak i parametrów, przy czym możliwa jest transformacja do modeli liniowych względem parametrów, np.

$$y_t = \beta_0 x_{1t}^{\beta_1} x_{2t}^{\beta_2} \varepsilon_t,$$

- ściśle nieliniowe, np.  $y_t = \beta_0 x_t^{\beta_1} \beta_2^{x_t} e^{\varepsilon_t}$ ,

modele szeregów czasowych: modele tendencji rozwojowej, np. model logistyczny o postaci:

$$y_t = \frac{\beta_0}{1 + \beta e^{-\gamma t}} + \varepsilon_t$$

oraz modele z rozkładem opóźnień, np. modele ARMAX o postaci:

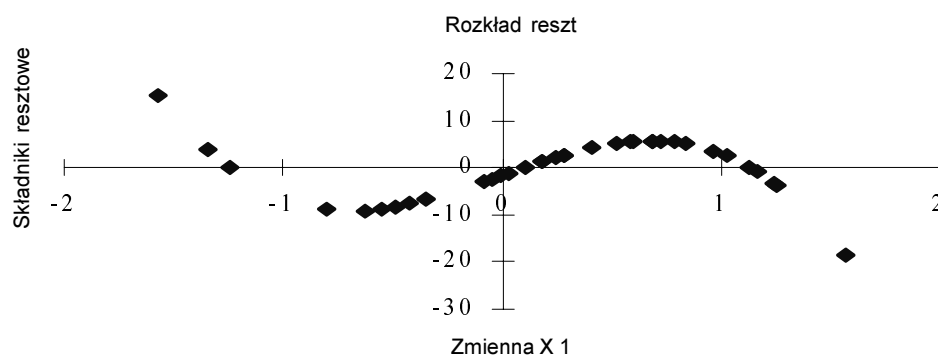
$$A(L)y_t = B(L)x_t + C(L)\varepsilon_t$$

Polecamy, aby dobór właściwej postaci funkcyjnej oprzeć na wiedzy merytorycznej lub/i analizie graficznej reszt, lub/i wykorzystać do weryfikacji liczne testy statystyczne, tj. test Serii, test Reset, test  $P_E$  Davidsona-McKinona, modele segmentowe (test stabilności Chowa), kryteria informacyjne, tj.: Akaike'a, Schwarz, Boxa i Pierca, Ljunga-Boxa. Na danych fikcyjnych prezentujemy oszacowanie modelu liniowego wraz z niektórymi miarami dopasowania.

$$\hat{y}_i = 455,2 + 58,88x_i \quad \bar{R}^2 = 0,9805$$

$$[113,0][14,3] \quad D-W = 1,98$$

Wszystkie parametry modelu są statystycznie istotne, a miary dopasowania świadczą o dobrej zgodności danych empirycznych do modelu. Jednak wykres reszt tego modelu (rys. 2.) wykazał zły dobór postaci analitycznej funkcji.



Rysunek 2. Wykres reszt dla oszacowanego modelu liniowego  
Źródło: opracowanie własne.

Bartosiewicz [1990] przy wyborze postaci analitycznej funkcji zaleca sporządzanie k-wykresów dla pojedynczych zmiennych objaśniających ze zmienną objaśnianą. Na ich podstawie przyjmuje się postacie analityczne funkcji dla poszczególnych zmiennych objaśniających. Postać modelu ekonometrycznego powstaje przez sumowanie tych części szczegółowych funkcji, które są przy zmiennych objaśniających, łącznie z uwzględnieniem wyrazu wolnego. Takie podejście budzi jednak pewne wątpliwości, ponieważ nie bierze pod uwagę współzależności pomiędzy tymi zmiennymi. Metodę tę polecamy przede wszystkim do doboru typu funkcji w przypadku, gdy rozpatrujemy model pomiędzy dwiema zmiennymi [Borkowski i in. 2003].

W literaturze naukowej jest polecanych kilka miar dopasowania modelu do danych empirycznych. Najpopularniejszą miarą jest współczynnik determinacji (współczynnik R-kwadrat). W jego magiczną rolę wszyscy wierzymy i dążymy do maksymalizacji jego wartości, czy słusznie? Celem badań jest ustalenie związków między zmiennymi (zjawiskami ekonomicznymi) a nie ustalenie wysokiego R-kwadrat. Dostępne pakiety statystyczne liczą (niezależnie od postaci modelu) współczynnik korelacji wielorakiej (R), współczynnik determinacji wielorakiej (R-kwadrat) oraz skorygowany współczynnik determinacji wielorakiej ( $\bar{R}^2$ ). Istnieje jednak wiele sytuacji, w których interpretacja R-kwadrat budzi pewne wątpliwości. Występują one gdy:

1.  $n = k+1$ , wtedy R-kwadrat równa się 1. Generalnie im  $n$  bliższe  $k+1$ , tym R-kwadrat bliższy jedności.

2. Liczba  $n$  jest niewiele większa od  $k+1$ . Raczej stosować skorygowany R-kwadrat o postaci:

$$\bar{R}^2 = 1 - \frac{SSE / (n - k - 1)}{SST / (n - 1)}.$$

3. Model nie jest liniowy lub model nieliniowy sprowadzony do liniowego (np. przez linearyzację). Wówczas R-kwadrat informuje o dopasowaniu zmiennych przekształconych (ich logarytmów), a nie wartości rzeczywistych.  
 4. Stosowano inną metodę estymacji niż metodę najmniejszych kwadratów (MNK). Wówczas R-kwadrat nie ma własności  $R^2 \in < 0, 1 >$ .  
 5. Model nie ma wyrazu wolnego. Stosuje się wówczas inną miarę zgodności – niescentrowany R-kwadrat

$$R_N^2 = \frac{\sum_{i=1}^n \hat{Y}_i^2}{\sum_{i=1}^n Y_i^2}.$$

6. Występuje efekt katalizy (zagadnienie współliniowości). Efekt katalizy – najogólniej rzecz ujmując – to wysoka wartość R-kwadrat determinowana silnym skorelowaniem zmiennych objaśniających między sobą. Wzajemne skorelowanie zmiennych objaśniających nazywa się współliniowością. Duża współliniowość powoduje trudności we wnioskowaniu o poszczególnych parametrach (średnie błędy parametrów są bardzo duże lub ilorazy  $t$  są bardzo małe; ma to bezpośredni wpływ na przedziały ufności, które są wówczas bardzo szerokie). Z drugiej strony w badaniach relacji pomiędzy zjawiskami ekonomicznymi musimy liczyć się z wzajemnym skorelowaniem tych zjawisk. W praktyce musimy więc stwierdzić, jak duże skorelowanie jest dopuszczalne, aby nie wystąpiły trudności we wnioskowaniu o poszczególnych parametrach. W literaturze naukowej proponowanych jest kilka rozwiązań tego problemu [Gujarati 2003, Welfe i in. 2003].  
 7. Wzrost skorelowania  $X$ -ów między sobą możemy zmierzyć wykorzystując tzw. parę korelacyjną i R-kwadrat [Gruszczynski, Podgórska 2003] o postaci:

$$R^2 = r^T R^{-1} r, \text{ gdzie: macierz } R = [r_{ij}], \text{ przy } r_{ij} = Corr(X_i, X_j),$$

wektor  $r$  oznacza  $[k \times 1]$  wektor korelacji  $r_i = Corr(X_i, Y)$ . Ze wzoru korelacyjnego na R-kwadrat możemy zobaczyć, jak zmienia się R-kwadrat w miarę skorelowania  $X$ -ów między sobą. Gdy zmienne objaśniające są wzajemnie ortogonalne (tzn. współczynniki korelacji dla wszystkich par zmiennych równają się zero) wówczas R-kwadrat jest sumą kwadratów współczynnika korelacji zmiennej  $Y$  z  $X$ -ami.

8. W przypadku wprowadzania dodatkowej zmiennej objaśniającej do modelu wartość współczynnika  $R^2$  nigdy nie maleje. Green [2000] udowodnił, że jeżeli zestaw zmiennych objaśniających  $x_1, x_2, \dots, x_k$  uzupełnimy zmienną  $z$ , to zachodzi zależność:

$$R_z^2 = R^2 + (1 - R^2) r_{yz, x_1, x_2, \dots, x_k}^2, \text{ gdzie}$$

$R^2$  – współczynnik determinacji przy zestawie zmiennych objaśniających  $x_1, x_2, \dots, x_k$ ,  
 $R_z^2$  – współczynnik determinacji przy zestawie zmiennych objaśniających  $x_1, x_2, \dots, x_k, z$ ,  
 $r_{yz, x_1, x_2, \dots, x_k}$  – współczynnik korelacji cząstkowej<sup>1</sup> między zmienną  $y$  a zmienną  $z$ , po wyłączeniu wpływu zmiennych  $x_1, x_2, \dots, x_k$ .

<sup>1</sup> Współczynnik korelacji cząstkowej jest „zwykłym” współczynnikiem korelacji między resztami z modeli, w których zmiennymi objaśnianymi są zmienne  $y$  i  $z$  a zmiennymi objaśniającymi zmienne  $x_1, x_2, \dots, x_k$ .

Trudno zatem na podstawie  $R^2$  porównywać modele z taką samą zmienną objaśniającą i różną liczbą zmiennych objaśniających. Z tego względu czasami zaleca się w takiej sytuacji zastosowanie tzw. skorygowanego współczynnika determinacji (ang. *adjusted R-squared*):

$$\bar{R}^2 = 1 - \frac{\left( \sum_{i=1}^n e_i^2 \right) / (n - k - 1)}{\left( \sum_{i=1}^n (y_i - \bar{y})^2 \right) / (n - 1)}$$

9. Inną miarą korelacji są współczynniki (stosunki) korelacyjne Pearsona ( $e_{yx}$  oraz  $e_{xy}$ ). Wskaźniki korelacyjne Pearsona nie posiadają ograniczeń liniowych, co oznacza, że mogą być stosowane do pomiaru siły korelacji zmiennych niezależnie od tego, czy powiązania między nimi są liniowe, czy też krzywoliniowe. Im większy stopień krzywoliniowości związku korelacyjnego pomiędzy zmiennymi, tym coraz mniejsza użyteczność współczynnika korelacji liniowej w ocenie współzależności. Wyraźne preferencje uzyskują wtedy wskaźniki (stosunki) korelacyjne Pearsona [Jóźwiak, Podgórski 1992, Luszczewicz, Słaby 1997].
10. Teresa Marszałkiewicz [1986] przedstawiła interesującą propozycję obliczania współczynników determinacji krzywoliniowości regresji, które określają, jaką część wariancji zmiennej zależnej, niewyjaśnioną przez zastosowanie pewnego modelu regresji, została wyjaśniona przez zastosowanie innego modelu regresji, przy czym w obu modelach występują te same zmienne, ale w innych potęgach.

Przy budowie modeli ekonometrycznych zawsze mamy problem z właściwym doбором zmiennych objaśniających. W latach sześćdziesiątych zaproponowano szereg metod regresji krokowej [Draper, Smith 1981]. Wśród nich najpopularniejszymi procedurami są *backward* i *forward selection*. Są one oparte na doborze zmiennych pod względem kryterium istotności parametrów modeli. Te procedury mechanicznego wyboru zmiennych tracą popularność wśród badaczy. W literaturze zaproponowano wiele innych kryteriów doboru zmiennych; zestawiono je w tabeli 1.

Tabela 1. Wybrane kryteria wyboru między modelami regresji

Kryterium	Minimalizowane wyrażenie	Kryterium wyboru na podstawie wartości statystyki F
$\bar{R}^2$ Theila (1961 r.)	$RSS_j / (n - k)$	$F < 1$
$S_p$ Hockinga (1976 r.)	$RSS_j / [(n - k_j)(n - k_j - 1)]$	$F < 2$
$C_p$ Mallowsa (1973 r.)	$RSS_j + 2k_j \hat{\sigma}_m^2$	$F < \frac{2n}{n + k_1}$
PC Amemiyi (1980 r.)	$RSS_j / (n + k_j)(n - k_j)$	$F < 2 + \frac{k_2 + 1}{n - k - 1}$
AIC Akaike'a (1987 r.)	$RSS_j \exp[2(k_j + 1) / n]$	$F < \frac{n - k}{n + k_1}$

Źródło: Maddala [2006a, s. 548-554], gdzie:  $RSS_j$  oznacza resztową sumę kwadratów  $j$ -tego modelu z  $k_j$  zmiennymi objaśniającymi.



Czy zmienna w modelu musi być istotna? Jest to wskazane, bowiem pośrednio świadczy o braku przypadkowości. Jednak stosowanie testu statystycznego t-Studenta nie daje pewności, że dana zmienna jest w modelu faktycznie istotna. Klasycznym zakłóceniem wartości statystyki t jest np. omawiana wyżej współliniowość zmiennych objaśniających. W praktyce wynik testu oznaczający brak dowodu na istotność zmiennej nie może być jedynym powodem usunięcia danej zmiennej z modelu. Chcemy mocno podkreślić, że badania istotności parametrów modeli za pomocą testu t-Studenta czy F-Snedecora wymagają spełnienia założenia o normalności rozkładu składnika losowego (schemat Gaussa-Markowa).

W ostatnim okresie ekonometrycy szczególną uwagę poświęcali weryfikacji modeli na podstawie testów diagnostycznych i testowaniu specyfikacji oraz konstruowaniu metod estymacji opartych na słabszych założeniach na temat rozkładu składnika losowego, czyli estymacji semiparametrycznej. Testy diagnostyczne są oparte na resztach klasycznej metody najmniejszych kwadratów (KMNK). Z resztami KMNK wiążą się dwa problemy.

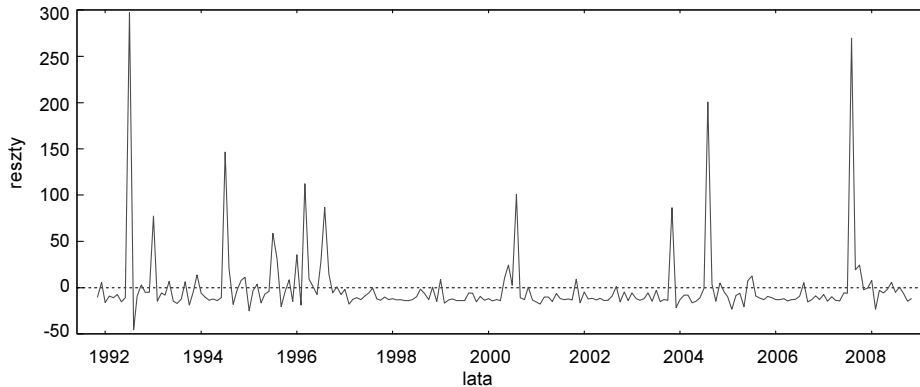
Reszty te cechują się heteroskedastycznością i autokorelacją, nawet jeśli w rzeczywistości składnik losowy nie podlega autokorelacji i ma stałą wariancję. W celu rozwiązania tych problemów zaproponowano alternatywne rodzaje reszt. Są to reszty BLUS i reszty rekurencyjne (pakiet SAS). W modelach ekonometrycznych niepewność związana z działalnością gospodarczą znajduje swoje odzwierciedlenie w wariancji  $\sigma^2$  składnika losowego  $\varepsilon_t$ . Wiele badań wykazało, że modelowanie wariancji jest bardzo ważne, szczególnie na rynkach finansowych [Doman, Doman 2009]. Do najbardziej popularnych metod jej modelowania należą model ARCH (*autoregressive conditionally heteroskedastic model*) zaproponowany przez Engle'a [1982] oraz model GARCH (*generalized autoregressive conditionally heteroskedastic model*) zaproponowany przez Bollerslewa w 1986 roku [Engel, Granger 1987]. W modelach tych wariancja bezwarunkowa jest stała a zmienna jest wariancja warunkowa.

$$\text{Wariancja bezwarunkowa: } \text{Var}(a_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i} \text{ dla } \sum_{i=1}^q \alpha_i < 1$$

$$\text{Wariancja warunkowa: } \sigma_t^2 = \text{Var}(r_t / F_{t-1}) = E[(r_t - u_t)^2 / F_{t-1}],$$

gdzie:  $u_t = E(r_t / F_{t-1})$ ,  $F_{t-1}$  – zbiór informacji dostępnych w momencie  $t-1$ .

Wyjątkowo wysoka wariancja w jednym okresie prowadzi do zwiększania niepewności w następnym. Wariancja ma tendencję do skupiania się, szczególnie własność tę zaobserwowano przy analizach cen giełdowych. Nasze badania wykazały występowanie statystycznie istotnych efektów ARCH również w modelach ekonometrycznych opisujących tendencje zmian cen produktów rolnych (rys. 3.).



Rysunek 3. Wykres wartości wariancji warunkowej dla realnych cen skupu pszenicy w Polsce w latach 1992-2009  
Źródło: opracowanie własne.

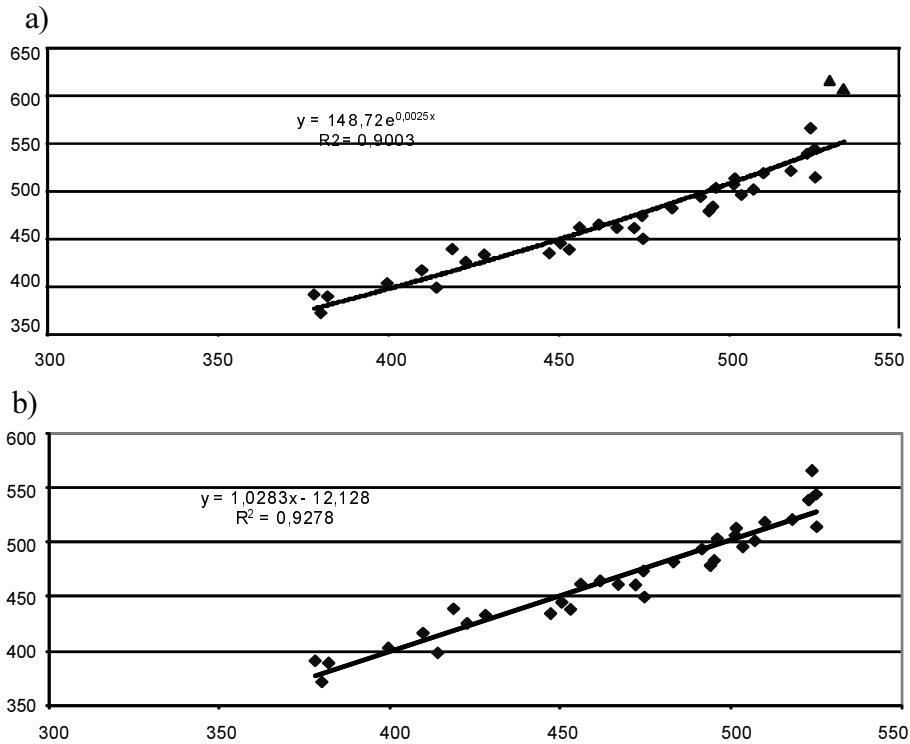
W badaniach ekonomicznych często występują obserwacje nietypowe (skrajne – *outliers*). Łatwo je zaobserwować podczas badania reszt oszacowanego modelu. Często stosowaną metodą rozwiązania tego problemu jest usunięcie obserwacji o dużych resztach i ponowne oszacowanie modelu. Jednak reszty MNK nie tylko nie mają jednakowej wariancji, lecz także nie dostarczają żadnej wskazówki, na ile ważny jest wpływ konkretnej obserwacji na końcowy wynik estymacji. Ostatnio zaproponowano nowe kryterium wykrywania obserwacji nietypowych metodą DFFITS (*difference of fits*), która mierzy zmianę teoretycznej wartości  $Y$ -teoretycznej zmiennej  $Y$  wynikającą z pominięcia konkretnej obserwacji. Nie eliminujemy jej, sprawiamy, że wpływ obserwacji jest minimalizowany (tzw. metoda estymacji ograniczonego wpływu – pakiet ekonometryczny SAS). Na rysunkach 4. a) i b) prezentujemy efekt wpływu obserwacji nietypowych na wartości estymacji parametrów modelu.

W podsumowaniu tej części artykułu należy stwierdzić, że budowa, specyfikacja i weryfikacja modeli ekonometrycznych wymagają obszernej wiedzy, umiejętności i dużej pokory. Szczególnie dyskusyjny jest sposób weryfikacji zbudowanych modeli ekonometrycznych. Powszechnie uważa się, że proces modelowania jest sztuką i w związku z tym nie poddaje się ścisłym regułom [Gruszczyński, Mierzejewski 1998]. Na podstawie tych samych danych empirycznych i tych samych ogólnych założeń możemy otrzymać różne modele. Każdy badacz ma prawo być przekonany o słuszności swojego modelu, jednak w procesie weryfikacji modelu każda konkluzja staje się w pewnej mierze dyskusyjna, bowiem opiera się na jakichś założeniach. Mimo to nie polecamy omijania tego etapu. Weryfikacja to sprawdzenie, czy oszacowany model spełnia postulaty poprawności merytorycznej i formalnej. Do danych statystycznych zawsze można dopasować jakiś model. Chodzi jednak o to, by model realnie opisywał zjawisko ekonomiczne. Z tego powodu postuluje się by:

- parametry modelu były merytorycznie sensowne,
- dopasowanie modelu było wystarczająco wysokie,
- zmienne objaśniające modelu były istotne [Guzik i in. 2004].

Uważamy bowiem, że wszechstronna ocena modelu pozwala określić stopień zaufania, jakim możemy model obdarzyć w dalszych zastosowaniach.

Przedstawione dotychczas zagadnienia i uwagi dotyczyły wykorzystania w badaniach ekonomicznych modeli statycznych. Ważne znaczenie w analizach ekonomicznych mają zagadnienia dotyczące zależności w czasie, zachodzące pomiędzy kolejnymi obserwacjami zmiennej.



a) z obserwacjami nietypowymi  
b) z usuniętymi obserwacjami nietypowymi (metodą DFFITS)

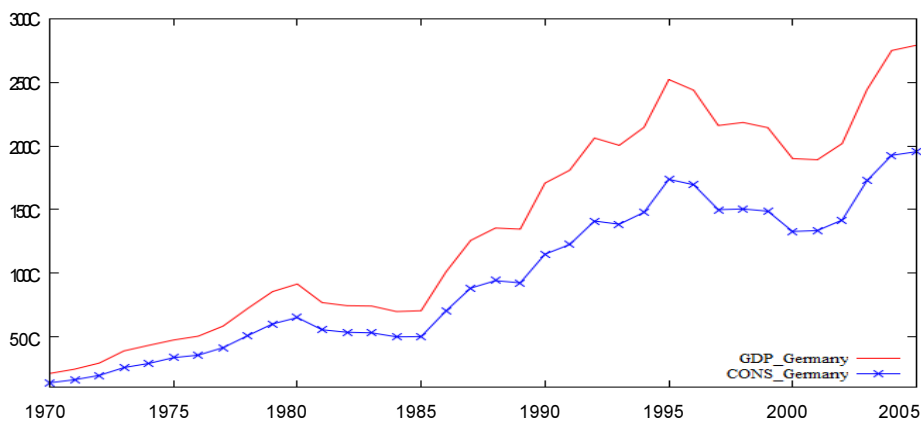
Rysunek 4. Estymacja KMNK

Źródło: obliczenia własne przy użyciu pakietu SAS.

#### DYNAMICZNE MODELE EKONOMETRYCZNE

Cechą charakterystyczną modeli dynamicznych jest występowanie czynnika czasu (rozważania nasze ograniczymy do zmian w czasie bez analizy częstości występowania tych zjawisk). Analiza dynamiki zmian procesów gospodarczych zazwyczaj ogranicza się do analizy czterech składników: trendu, wahań okresowych związanych z sezonowością, wahań koniunkturalnych i wahań nieregularnych. Modele dynamiczne możemy podzielić na modele tendencji rozwojowej (modele trendu) oraz modele, w których występują opóźnione wartości zmiennych: objaśnianej, objaśniającej lub obu tych kategorii jednocześnie. Pomijamy tu omawianie pierwszej grupy modeli, gdyż są one powszechnie znane, a koncentrujemy się na wybranych modelach z rozkładem opóźnień i autoregresyjnym. Wybrane problemy związane z analizą takich modeli zobrazujemy na przykładzie empirycznym. Zgodnie z

teorią ekonomii, model keynesowski wskazuje, że wartość konsumpcji zależy od poziomu dochodu do dyspozycji, a zależność ta jest określona przez krańcową skłonność do konsumpcji (KSK). Czy jednak wzrost konsumpcji w reakcji na zmiany dochodów jest natychmiastowy, czy raczej rozłożony w czasie? Podstawą do przedstawienia niektórych uwag dotyczących wykorzystania dynamicznych modeli ekonometrycznych były dane empiryczne charakteryzujące: wartości CONS – roczne wydatki na konsumpcję w USD/osobę (zmienna Y), GDP – roczny PKB w USD/osobę (zmienna X) w cenach z 2000 roku w Niemczech w okresie od 1970 do 2005 roku. Graficznie kształtowanie się PKB/osobę i wydatków na konsumpcję przedstawiono na rys. 5.



Rysunek 5. Wartość produktu krajowego brutto i konsumpcji indywidualnej w Niemczech w latach 1970-2005

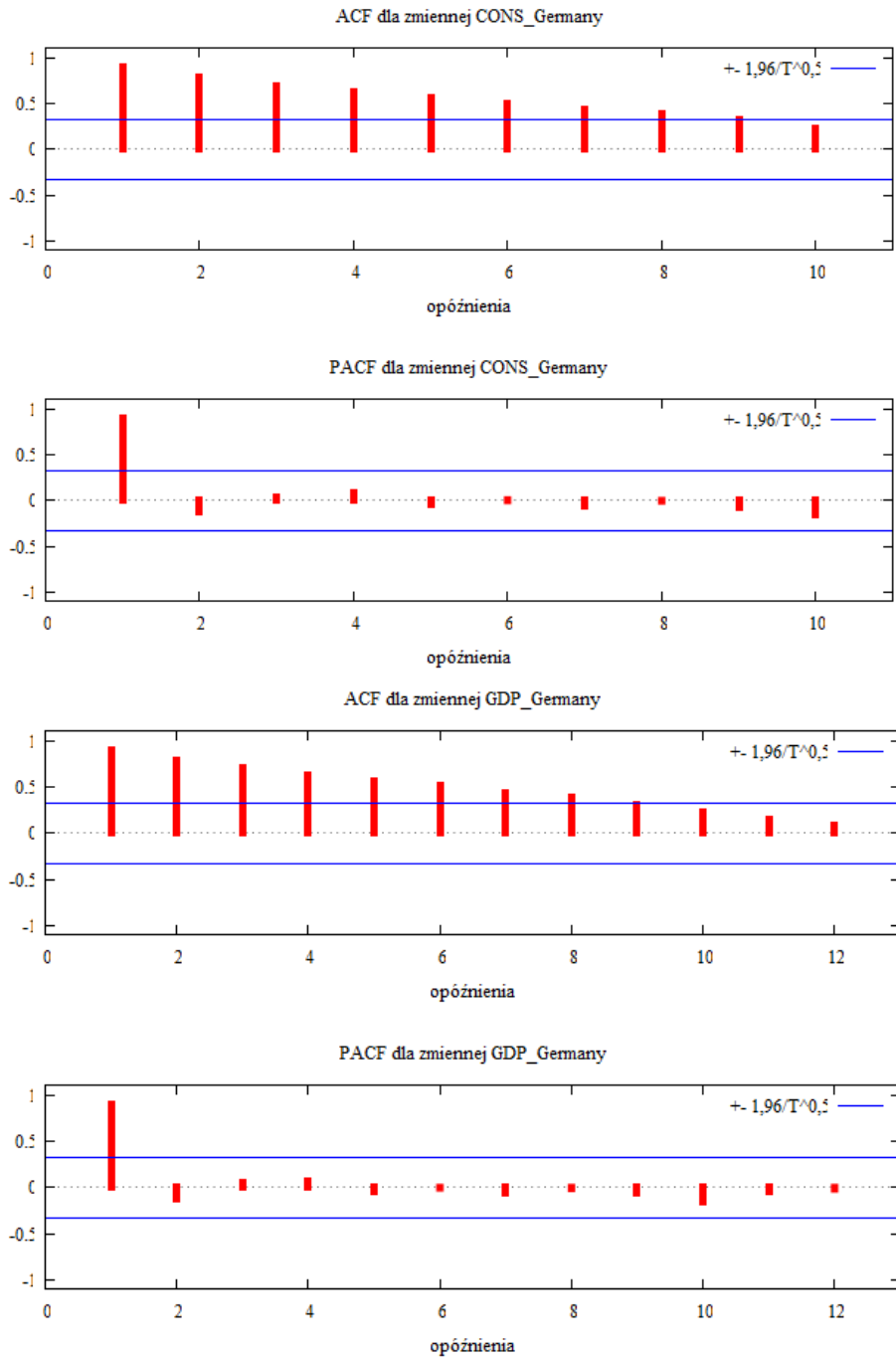
Źródło: obliczenia własne przy użyciu pakietu GRETL.

Analiza graficzna długookresowej tendencji zmian dwóch badanych czynników wskazała na występowanie tendencji, wahań okresowych oraz braku stacjonarności badanych cech. Wartości funkcji autokorelacji i autokorelacji cząstkowej potwierdzają brak stacjonarności badanych szeregów czasowych (rys. 6.).

Procesem stochastycznym  $\{Y_t\}$  nazywamy rodzinę zmiennych losowych określonych w przestrzeni probabilistycznej  $(\Omega, F, P)$ , indeksowanych indeksem  $t \in T$ , gdzie  $T$  jest pewnym podzbiorem indeksów.

Pojęcie szeregu czasowego  $Y_t$  można zdefiniować jako realizację procesu stochastycznego  $\{Y_t\}$ . Proces  $\{Y_t\}$  jest ściśle stacjonarny, jeśli dla dowolnych wartości  $t_1, t_2, \dots, t_n$  i dowolnego  $k$  łączny rozkład gęstości prawdopodobieństwa dla  $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}$  jest identyczny z łącznym rozkładem gęstości prawdopodobieństwa dla  $\{Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}\}$ . Proces stochastyczny  $\{Y_t\}$  jest procesem słabo stacjonarnym, czyli stacjonarnym drugiego rzędu, jeżeli średnia  $m$  i wariancja  $s^2$  nie zmieniają się w czasie, a kowariancje  $g_s$  mają wartości zależne od odstępów  $s$  pomiędzy momentami obserwacji, a nie zależą od  $t$ .

- średnia  $E(Y_t) = \mu_t$ ,
- wariancja  $E[(Y_t - \mu_t)^2] = \sigma_t^2$ ,
- kowariancje  $E[(Y_t - \mu_t)(Y_{t-s} - \mu_{t-s})] = \gamma_{t,s}$ .



Rysunek 6. Korelogramy (funkcje autokorelacji i autokorelacji cząstkowej)  
 Źródło: obliczenia własne przy użyciu pakietu GRETL.

Niestacjonarność szeregów czasowych jest zawsze problemem w analizie ekonometrycznej. W wielu pracach teoretycznych wykazano, że własności parametrów modeli ekonometrycznych otrzymanych dla niestacjonarnych szeregów czasowych są na ogół wątpliwe [Charemza, Deadman 1997]. W przypadku braku stacjonarności możemy popełniać błędy predykcji dotyczące:

- weryfikacji istotności parametrów (testy istotności oparte są na założeniu stacjonarności),
- znacznego, nieuzasadnionego wzrostu  $R^2$ ,
- występowania regresji pozornej (otrzymujemy model z obiecującymi wynikami testów diagnostycznych nawet wtedy, kiedy regresja nie ma sensu).

Dobrym przykładem regresji pozornej jest bezsensowna relacja pomiędzy liczbą wypadków drogowych w Polsce spowodowana przez nietrzeźwych kierowców a ilością wypić piwa w Grecji [Gruszczyński, Podgórska 2003, s. 188].

$$\begin{aligned} \hat{W}_t &= 6,02 + 0,12P_t & R^2 &= 0,98 \\ [1,11] [0,01] & & D-W &= 0,84 \end{aligned}$$

gdzie:

$\hat{W}_t$  – liczba wypadków drogowych w Polsce w latach 1980-1995 w tys. spowodowana przez nietrzeźwych kierowców,

$P_t$  – spożycie piwa w Grecji w l na mieszkańca.

Otrzymano następujące oceny parametrów, które mogą świadczyć o związku przyczynowo-skutkowym. Jedyne na podstawie wielkości testu D-W i porównaniu go do wielkości współczynnika determinacji można wnioskować o występowaniu regresji pozornej.

Stacjonarność szeregu czasowego możemy badać, wykorzystując formalne testy statystyczne, funkcje autokorelacji i autokorelacji cząstkowej lub też wykresy tendencji zmian badanych cech. Istnieje wiele testów statystycznych, służących do badania stacjonarności [Haremza, Deadman 1997, Osińska 2006]. Do najpopularniejszych należą testy Dickeya-Fullera (DF i ADF), integracyjna statystyka Durбина-Watsona (IDW), test Dickeya, Haszy i Fullera (DHF – do testowania sezonowej integracji), testy DFSI i ADFSI, test Hyllerberga, Engle'a, Grangera i Yoo (HEGY), test Phillipsa i Phillipsa-Perrona oraz test KPSS (Kwiatkowskiego, Phillipsa, Schmidta i Shina).

Do najbardziej popularnych testów pozwalających zweryfikować hipotezę o stacjonarności badanych szeregów (ocenę stopnia integracji szeregu) należą tzw. testy pierwiastka jednostkowego (DF i ADF). Modelem wyjściowym w tych testach jest oszacowanie parametrów metodą najmniejszych kwadratów jednego z następujących równań:

$$\begin{aligned} \Delta Y_t &= \alpha_1 Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= \alpha_0 + \delta_1 t + \alpha_1 Y_{t-1} + \varepsilon_t \end{aligned}$$

gdzie:

$t$  – zmienna czasowa,

$\alpha_1, \alpha_0, \delta_1$  – parametry modelu,

$Y_{t-1}$  – wartości szeregu obserwacji,

$$\Delta Y_t = Y_t - Y_{t-1}.$$

Szeregi czasowe, szczególnie w badaniach ekonomicznych, charakteryzują się autokorelacją, która wpływa bezpośrednio na autokorelację składnika losowego. Test DF nie uwzględnia autokorelacji składnika losowego. Jeżeli składnik losowy ( $\varepsilon_t$ ) nie jest białym szumem, to estymacja klasyczną metodą najmniejszych kwadratów (KMNK) daje nieefektywne parametry. W celu usunięcia tej wady zaproponowano modyfikację testu DF w postaci rozszerzonego testu Dickeya-Fullera (ADF). Równania wówczas są postaci:

$$\Delta Y_t = \alpha_1 Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \varepsilon_t$$

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \varepsilon_t$$

$$\Delta Y_t = \alpha_0 + \delta_1 t + \alpha_1 Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \varepsilon_t$$

Do weryfikacji hipotezy zerowej stosuje się wartości krytyczne testu DF (hipoteza zerowa zakłada niestacjonarność szeregu). Jeżeli nie ma podstaw do odrzucenia hipotezy zerowej, wnioskuje się, że szereg jest szeregiem niestacjonarnym i procedurę testowania należy powtórzyć dla szeregu pierwszych różnic (za zmienną  $Y_t$  przyjmuje się  $\Delta Y_t$  a za zmienną  $\Delta Y_{t-1}$  drugie różnice, tj.  $\Delta \Delta Y_{t-1} = \Delta Y_{t-1} - \Delta Y_{t-2}$ ).

W przypadku, gdy kolejne różnicowanie nie prowadzi do uzyskania szeregów stacjonarnych, należy uwzględnić pewne modyfikacje statystyki testującej [Enders 1995].

Weryfikacji stacjonarności badanych cech dokonaliśmy wykorzystując zmodyfikowany test Dickeya-Fullera (ADF). W równaniu nie uwzględnialiśmy występowania wahań sezonowych. Weryfikowana była hipoteza zerowa, zakładająca, że dany szereg jest zintegrowany w stopniu pierwszym  $H_0: \alpha_1 = 0$ . Wyniki obliczeń przedstawia tabela 2.

Tabela 2. Wyniki testu ADF dla badanych realnych wartości cech

Zmienna	Statystyka testu tau	Asymptotyczna wartość p	Decyzja o stopniu integracji
Test z wyrazem wolnym			
CONS w USD/osobę	-0,1936	0,937	
d_CONS w USD/osobę	-3,3911	0,0113	I(1)
GDP w USD/osobę	-0,2440	0,9305	
d_GDP w USD/osobę	-3,3957	0,0111	I(1)
Test z wyrazem wolnym i trendem liniowym			
CONS w USD/osobę	-2,1972	0,4906	
d_CONS w USD/osobę	-3,3014	0,0660	
GDP w USD/osobę	-2,3031	0,4317	
d_GDP w USD/osobę	-3,3067	0,0651	

(d „nazwa zmiennej” oznacza pierwsze różnice zmienne, rząd opóźnień p=6)

Źródło: opracowanie własne.

Przeprowadzone badania wykazały, że szeregi czasowe opisujące roczne realne wydatki na konsumpcję i wielkość produktu krajowego w USD/osobę są szeregami niestacjonarnymi (zintegrowane w stopniu pierwszym). Niestacjonarne szeregi czasowe, które można sprowadzić do postaci szeregów stacjonarnych przez ich pierwsze różnicowanie określone są jako szeregi przyrostostacjonarne, zaś występujący trend w takich szeregach jest trendem stochastycznym. Natomiast szeregi trendostacjonarne stanowią sumę trendu determi-

nistycznego oraz szeregu obserwacji stacjonarnej zmiennej losowej. W praktyce rozróżnienie tych szeregów nie jest rzeczą łatwą, a jest to bardzo istotne przy wyborze metody usunięcia trendu z szeregu czasowego [Gruszczyński i in. 2009].

Powróćmy do naszego problemu badawczego. W makroekonomicznym modelu równowagi zakłada się, że konsumpcja ( $CONS_t$ ) jest funkcją zarówno bieżącej wartości dochodów (w naszym przypadku PKB<sub>t</sub>) jak i opóźnionych wartości tej zmiennej. Możemy to zapisać:

$$CONS_t = \alpha + \beta_0 PKB_t + \beta_1 PKB_{t-1} + \beta_2 PKB_{t-2} + \dots + \varepsilon_t$$

Jest to model z rozkładem opóźnień (*Distributed Lag Model*) o następującej postaci:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_k x_{t-k} + \varepsilon_t \quad \text{lub} \quad y_t = b(L)x_t + \varepsilon_t \quad \text{przy wykorzystaniu}$$

operatora opóźnień ( $Ly_t = y_{t-1}$ ).

Dla szeregów niestacjonarnych model ADL ma postać:

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + \dots + \beta_k \Delta x_{t-k} + \varepsilon_t$$

Do estymacji tych modeli można wykorzystać KMNK. Problemem jest tu ustalenie wielkości  $k$  (opóźnień zmiennych objaśniających). Wzrost ich powoduje zmniejszanie liczby stopni swobody, szczególnie w przypadkach krótkich szeregów czasowych. Dodatkowo opóźnienia zmiennej charakteryzują się współliniowością. Alternatywnym sposobem estymacji jest transformacja Koycka. Autor przyjął pewne założenia dotyczące parametrów modelu  $\beta_k$  dla  $k = 0, 1, 2, \dots$ , a mianowicie, wszystkie parametry są tego samego znaku i maleją geometrycznie, tzn.:

$$\beta_k = \beta_0 \lambda^k \quad k = 0, 1, 2, \dots \quad \text{gdzie } 0 < \lambda < 1 \quad \text{nazywa się stopą zaniku rozkładu opóźnień.}$$

Koyck wychodząc z modelu ADL, po kilku prostych przekształceniach algebraicznych [Gujarati 2003] otrzymał model zwany w literaturze modelem Koycka:

$$y_t = \delta_0 + \lambda y_{t-1} + \beta_0 x_t + v_t, \quad \text{gdzie } \delta_0 = \alpha(1 - \lambda) \quad v_t = \varepsilon_t - \lambda \varepsilon_{t-1}$$

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 (1 + \lambda + \lambda^2 + \lambda^3 + \dots) = \beta_0 \left( \frac{1}{1 - \lambda} \right) \quad \text{jest mnożnikiem długookresowym.}$$

Wyniki estymacji MNK dla naszych danych empirycznych są następujące:

$$CONS_t^{\hat{}} = 3,0696 + 0,6681GDP_t + 0,1376CONS_{t-1}$$

Parametry postaci wyjściowej modelu są następujące:  $\hat{\beta}_0 = 0,6681$ ,  $\lambda = 0,1376$ . Po dalszych podstawieniach oraz obliczeniu  $\hat{\alpha}$  z zależności  $\delta_0 = \alpha(1 - \lambda)$  otrzymaliśmy:

$$CONS_t^{\hat{}} = 3,5594 + 0,6681GDP_t + 0,0919GDP_{t-1} + 0,01227GDP_{t-2} + \dots$$

Mnożnik długookresowy obliczony według formuły  $\beta_k = \beta_0 \lambda^k$  wynosił 0,7747. Oznacza to, że wzrost produktu krajowego brutto o 1 mln USD w Niemczech powodował wzrost wydatków na konsumpcję (po uwzględnieniu wszystkich opóźnień) w wysokości 0,77 mln USD. Mnożnik krótkookresowy był niższy i wynosił 0,6681, co oznacza, że konsumpcja charakteryzuje się pewną inercją względem GDP.

Wadą modelu Koycka jest występująca autokorelacja składnika losowego, poza przypadkiem, gdy  $\lambda = 0$ , a więc nie są spełnione założenia Gaussa-Markowa dla metody najmniejszych kwadratów, pewnym rozwiązaniem może być wykorzystanie metody zmiennych instrumentalnych jako alternatywnej metody estymacji. W modelu tym występuje opóźnio-



na zmienna objaśniająca, więc do testowania autokorelacji składnika losowego nie powinno się stosować testu Durбина-Watsona, a raczej test h D-W (w naszym przypadku statystyka Durбина h wynosiła 4,7696).

Sprawdźmy, czy relacja pomiędzy wydatkami na konsumpcję i wielkością produktu krajowego brutto nie jest regresją pozorną, czy też relacja ta jest relacją trwałą, stabilną, czyli relacją kointegrującą. Engle i Granger [1987] zaproponowali metodę pozwalającą na odróżnienie relacji kointegrującej od regresji pozornej. Zgodnie z definicją zależność między zmiennymi niestacjonarnymi, zintegrowanymi w stopniu pierwszym:

$$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t, \text{ gdzie } y_t \sim I(1) \text{ oraz } x_t \sim I(1)$$

jest relacją kointegrującą, jeżeli składnik losowy jest stacjonarny, czyli  $\varepsilon_t \sim I(0)$ . W przypadku przeciwnym, czyli gdy składnik losowy jest niestacjonarny, czyli  $\varepsilon_t \sim I(1)$  możemy spodziewać się regresji pozornej. Testowanie kointegracji szeregów czasowych może przebiegać w następujący sposób:

1. Testowanie stopnia integracji empirycznych szeregów czasowych ( $y_t, x_t$ ) za pomocą testów, np. DF lub ADF,
2. Szacowanie równania kointegrującego (relacji długookresowej) o postaci:

$$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t$$

3. Wyznaczanie reszt równania kointegrującego i testowanie stopnia integracji reszt za pomocą testów statystycznych:

a) test DF:  $\Delta e_t = \gamma e_{t-1} + v_t$

b) w przypadku wyraźnej autokorelacji reszt: test ADF o postaci:

$$\Delta e_t = \gamma e_{t-1} + \sum_{i=1}^k \Delta e_{t-i} + v_t$$

c) test kointegracyjny Durбина-Watsona (CIDW).

Przeprowadzone badania stopnia integracji zmiennych wykazały, że badane cechy są zintegrowane w stopniu pierwszym, tzn.  $CONS_t \sim I(1)$  oraz  $GDP_t \sim I(1)$ . Relacja długookresowa wynosi:

$$CONS_t = 1,4092 + 0,6938GDP_t \quad R^2 = 0,8997$$

[0,6212] [0,03299]

Autokorelacja reszt pierwszego rzędu - 0,8121

Statystyka testu Durбина-Watsona - 0,3780

Logarytm wiarygodności -148,6879,

Kryterium informacyjne Akaike'a - 301,38

Kryterium bayesowskie Schwarz - 304,54,

Kryterium Hannana-Quinna - 302,48

Stopnie swobody - 36

Ze względu na autokorelacje składnika losowego do testowania stopnia integracji reszt równania kointegrującego wykorzystaliśmy test ADF i otrzymaliśmy następujące równanie:

$$\Delta e_t = -0,3582e_{t-1} + 0,2394\Delta e_{t-1} + 0,2306\Delta e_{t-2} - 0,0507\Delta e_{t-3} + 0,4039e_{t-4}$$

Rozszerzony test Dickeya-Fullera ADF = - 2,398 a empiryczny poziom istotności p= 0,024. Oznacza to, że reszty te są stacjonarne, a relacja pomiędzy wydatkami na konsumpcję i wielkością produktu krajowego brutto jest kointegrująca. Wielkość  $k$  opóźnień dla przyrostów zmiennej objaśniającej ustaliliśmy na podstawie wielkości kryteriów informacyjnych Akaike'a i Hannana-Quinna.

## ZAMIAST PODSUMOWANIA

Powróćmy do pytania postawionego we wstępie. Zdajemy sobie sprawę, że po lekturze tego artykułu odpowiedź może być różna. Nie było naszym celem wykazanie wyższości badań ekonomicznych z wykorzystaniem najnowszych metod ekonometrycznych nad opisaną analizą zjawisk ekonomicznych. Chcieliśmy raczej pokazać pewne niebezpieczeństwa wynikające z mechanicznego stosowania tych metod i interpretacji uzyskanych wyników. Ze względu na ograniczenia wydawnicze skoncentrowaliśmy się tylko na nielicznych problemach metodycznych. Pamiętajmy, że modele stabilne w krótkim czasie nie zachowują stabilności w długim okresie. Zakończmy cytatem Davida F. Hendry'ego [1985] *Trzy złote zasady ekonometrii to: testować, testować, testować.*

## LITERATURA

- Bartosiewicz S. 1990: *Ekonometria z elementami programowania matematycznego i analizy porównawczej*, Wyd. AE we Wrocławiu, Wrocław.
- Borkowski B., Dudek H., Szczesny W. 2003: *Ekonometria. Wybrane zagadnienia*, PWN, Warszawa.
- Charemza W. W., Deadman D. F. 1997: *Nowa ekonometria*, PWE, Warszawa.
- Cieślak M. 1978: *Dobór cech prognostycznych metodą heurystyczną*, „Prace Naukowe Akademii Ekonomicznej we Wrocławiu”, nr 123, s. 47-59.
- Cieślak M. 2004: *Prognozowanie gospodarcze. Metody i zastosowania*, PWN, Warszawa.
- Doman M., Doman R. 2009: *Modelowanie zmienności i ryzyka. Metody ekonometrii finansowej*, Oficyna Wolters Kluwer business, Kraków.
- Draper N. R., Smith H. 1981: *Applied Regression Analysis*, John Wiley & Sons, New York.
- Enders A. 1995: *Applied Econometric Time Series*, John Wiley & Sons, New York.
- Engle R., Granger C. 1987: *Cointegration and Error Correction: Representation, Estimation and Testing*, „Econometrica”, vol. 55, s. 251-276.
- Friedman B. M. 1980: *The Methodology of Positive Economics*, University of Chicago Press, Chicago.
- Green W. H. 2000: *Econometric Analysis*, Prentice Hall, New Jersey.
- Gruszczyński M., Mierzejewski P. 1998: *Wstęp do ekonometrii w stu oknach*, Wyd. SGH, Warszawa.
- Gruszczyński M., Podgórska M. 2003: *Ekonometria*, Wyd. SGH, Warszawa.
- Gruszczyński M., Kuszewski T., Podgórska M. 2009: *Ekonometria i badania operacyjne. Podręcznik dla studiów licencyjnych*, PWN, Warszawa.
- Gujarati Domador N. 2003: *Basic Econometrics*, Mc Graw Hill, New York.
- Guzik B., Appenzeller D., Jurek W. 2004: *Prognozowanie i symulacje. Wybrane zagadnienia*, Wyd. AE w Poznaniu, Poznań.
- Hendry D. 1985: *Econometric Methodology, paper presented to the Econometric Society Fifth World Congress*.
- Hozer J. 1993: *Mikroekonometria. Analizy, diagnozy prognozy*, PWE, Warszawa.
- Józwiak J., Podgórski J. 1992: *Statystyka od podstaw*, PWE, Warszawa.
- Klein L. R. 1962: *An Introduction to Econometrics*, Macmillan, New York.
- Klein L. R. 1982: *Wykłady z ekonometrii*, PWE, Warszawa.
- Koopmans T. C. 1957: *Three Essays on the State of Economics Science*, McGraw Hill, New York.
- Kukuła K. 2000: *Metoda unitaryzacji zerowanej*, PWN, Warszawa.
- Luszniewicz A., Słaby T. 1997: *Statystyka stosowana*, PWE, Warszawa.
- Malina A. 2004: *Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw*, AE, Kraków.
- Maddala G. S. 2006a: *Introduction to Econometrics*, John Wiley & Sons, New York.
- Maddala G. S. 2006b: *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa.
- Marciniak S. (red.) 2005: *Makro- i mikroekonomia. Podstawowe problemy*, PWN, Warszawa.
- Marszałkiewicz T. 1986: *Metody statystyki opisowej w badaniach ekonomiczno-rolniczych*, wyd. IV poprawione. Wyd. SGGW-AR, Warszawa.
- Neal F., Shone R., 1982: *Proces budowy modeli ekonomicznych*, PWN, Warszawa.

- Osińska M. 2006: *Ekonometria finansowa*, PWE, Warszawa.  
 Osińska M. (red.) 2007: *Ekonometria współczesna*, Towarzystwo Naukowe Organizacji i Kierownictwa, Toruń.  
 Pawłowski Z. 1971: *Modele ekonometryczne równań opisowych*, PWN, Warszawa.  
 Tinbergen J. 1951: *Econometrics*, George Allen & Unwin, London.  
 Welfe W., Welfe A. 1996: *Ekonometria stosowana*, PWE, Warszawa.  
 Welfe A. 2003: *Ekonometria. Metody i ich zastosowanie*, PWE, Warszawa.  
 Zeliaś A. 2000: *Teoria prognozy*, PWN, Warszawa.

*Bolesław Borkowski, Stanisław Stańko*

REMARKS CONCERNING THE APPLICATION AND APPLYING ECONOMETRIC  
 METHODS IN THE ECONOMIC RESEARCH

Summary

In the study essential issues concerning the application of econometric methods in the economic research were presented. Stages of the proceeding in the construction of econometric models and interpretation issues of received results for stationary and dynamic econometric models were discussed. In detail problems concerning selection of variables, problems of choice of the model analytical form, the assessment of the degree of fitting the model to empirical data as well as logical and formal verification of received results were shown. Some threats resulting from mechanical applying these methods and interpretation of achieved results were presented as well.

Adres do korespondencji:

prof. dr hab. Bolesław Borkowski  
 Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
 Katedra Ekonometrii i Statystyki  
 ul. Nowoursynowska 166  
 02-787 Warszawa  
 e-mail: boleslaw\_borkowski@sggw.pl

dr hab. Stanisław Stańko, prof. SGGW  
 Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
 Katedra Ekonomiki, Rolnictwa i Międzynarodowych Stosunków Gospodarczych  
 ul. Nowoursynowska 166  
 02-787 Warszawa  
 e-mail: stanislaw\_stanko@sggw.pl