

ALEKSANDER ZELIAŚ

## SZACOWANIE PARAMETRÓW MODELI AUTOREGRESYJNYCH Z UWZGLĘDNIENIEM OPÓŹNIEŃ ZMIENNEJ OBJAŚNIAJĄCEJ

Weźmy pod uwagę liniowy model typu autoregresyjnego o postaci

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + u_t \quad (t=1, \dots, n), \quad (1)$$

gdzie:

- $Y_t$  — zmienna endogeniczna wyjaśniana przez dane równanie,
- $X_t$  — nieopóźniona zmienna objaśniająca,
- $X_{t-j}$  — opóźnione zmienne objaśniające dla  $j=1, \dots, k$ ,
- $\alpha, \beta_0, \beta_1, \beta_2, \dots$  — parametry strukturalne, które estymuje się najczęściej za pomocą metody najmniejszych kwadratów,
- $u_t$  — składnik losowy.

Zauważmy, że w modelu (1) występuje nieskończony zbiór opóźnionych zmiennych objaśniających. Zbiór ten trzeba jednak ograniczyć, ponieważ liczność próby ( $n$ ) jest skończona. W ten sposób otrzymujemy

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_k X_{t-k} + u_t. \quad (2)$$

Zauważmy dalej, że liczba opóźnionych zmiennych objaśniających w modelu (2) może być dowolna<sup>1</sup> z tym jednak, że musi być  $K < n$ , gdzie  $K$  oznacza liczbę wszystkich parametrów występujących w modelu (2), a  $n$  jest liczbą obserwacji.

Przy założeniu, że spełnione są następujące warunki:

1) Zmienne objaśniające  $X_{t-j}$  ( $j=0, 1, \dots, k$ ) są wielkościami nielosowymi,

2)  $E(u_t) = 0,$

3)  $E[u_t - E(u_t)]^2 = E(u_t)^2 = \sigma_u^2,$

<sup>1</sup> Problem właściwego ustalenia systemu opóźnień czasowych w modelu (2) należy do trudniejszych problemów ekonometrii. Por. w tej sprawie pracę Z. Pawłowskiego, *Ekonometria*, Warszawa 1969, s. 60.

$$4) \text{cov}(u_t, u_s) = E\{[u_t - E(u_t)][u_s - E(u_s)]\} = 0 \quad \text{dla} \quad s \neq t,$$

$$5) \text{cov}(u_t, X_{t-j}) = \text{cov}(u_t, X_t) = 0 \quad (j=1, \dots, k),$$

co oznacza, że składnik losowy  $u_t$  jest nie skorelowany ze zmiennymi objaśniającymi modelu (2), nieznane parametry  $\alpha, \beta_0, \beta_j$  ( $j=1, \dots, k$ ) szacuje się najczęściej za pomocą klasycznej metody najmniejszych kwadratów.

Przy stosowaniu metody najmniejszych kwadratów parametry modelu (2) wyznacza się z równania

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \quad (3)$$

gdzie  $\hat{\beta} = \{\hat{\alpha} \hat{\beta}_0 \hat{\beta}_1 \dots \hat{\beta}_k\}$  oznacza wektor ( $K \times 1$ ) ocen parametrów strukturalnych.

Zauważmy jeszcze, że jeżeli spełnione są podane na wstępie pracy założenia 1)-5), to macierz wariancji i kowariancji wyznaczonych relacją (3) estymatorów parametrów  $\alpha, \beta_0, \beta_j$  ( $j=1, \dots, k$ ) dana jest wzorem

$$\text{var}(\hat{\beta}) \cong \hat{\sigma}_u^2 (\mathbf{X}'\mathbf{X})^{-1}, \quad (4)$$

gdzie  $\hat{\sigma}_u^2$  jest nieobciążoną oceną wariancji  $\sigma_u^2$  składnika losowego  $u_t$ .

Zakładać będziemy dalej, że w zbiorze zmiennych objaśniających występujących w modelu (2) nie pojawiają się zmienne, między którymi zachodzi dokładnie zależność liniowa (jest to wstępny warunek konieczny do estymacji). Oznacza to, że macierz  $\mathbf{X}$  zaobserwowanych wartości zmiennych objaśniających  $X_{t,j}$  ( $j=0, 1, \dots, k$ ) ma rząd<sup>2</sup>:  $r(\mathbf{X})=K < n$ .

Zauważmy jednak, że w naszym przypadku występuje duże prawdopodobieństwo pojawienia się przybliżonych liniowych zależności w szeregach czasowych opóźnionych zmiennych objaśniających<sup>3</sup>, czyli tym samym macierz momentów rzędu drugiego  $\mathbf{X}'\mathbf{X}$ , do której mamy wyznaczyć macierz odwrotną, jest macierzą źle uwarunkowaną (wartość wyznacznika tej macierzy niewiele różni się od zera). W tym przypadku wyznaczenie wektora  $\hat{\beta}$  według wzoru (3) jest kłopotliwe z uwagi na trudności numeryczne z wyznaczeniem macierzy  $(\mathbf{X}'\mathbf{X})^{-1}$ . Okoliczność ta powoduje, że otrzymuje się macierz odwrotną  $(\mathbf{X}'\mathbf{X})^{-1}$  o bardzo dużych (co do modułu) elementach, a w konsekwencji duże elementy macierzy wariancji i kowariancji rozpatrywanych estymatorów. Tak więc istnienie silnej współliniowości wpływa w bardzo poważny sposób na efektywność estymatorów parametrów strukturalnych modelu (2), otrzymywanych za pomocą klasycznej metody najmniejszych kwadratów. Wpływ ten jest

<sup>2</sup> Zauważmy, że gdyby  $r(\mathbf{X})=K$ , to macierz  $\mathbf{X}'\mathbf{X}$  byłaby macierzą osobliwą i estymator liniowy  $\hat{\beta}$  wektora  $\beta$  otrzymany metodą najmniejszych kwadratów nie byłby określony jednoznacznie.

<sup>3</sup> Zjawisko to znane jest w literaturze naukowej pod nazwą współliniowości lub łącznej korelacji. Por. na ten temat pracę A. Zeliaś, *Z problematyki badania współliniowości w modelach ekonometrycznych*, Przegląd Statystyczny 1977, z. 2.

ujemny, to znaczy istnienie współliniowości prowadzi prawie zawsze do obniżenia efektywności estymatorów, to znaczy do wydatnego zwiększenia ich wariancji. Fakt ten zmniejsza szanse prawidłowego oszacowania parametrów strukturalnych modelu (2) (duże błędy standardowe estymatorów spowodują, iż będziemy niepewni co do rzeczywistych wartości tych parametrów). Tych spraw nie będziemy jednak tu szczegółowo omawiać, gdyż wykraczają poza ramy niniejszych rozważań. W tym miejscu jedynie zasygnalizujemy, że z formalnego punktu widzenia istnieje możliwość uniknięcia zjawiska występowania współliniowości w szeregach czasowych opóźnionych zmiennych objaśniających, a to przez wprowadzenie macierzy  $\mathbf{W}$ , która ma postać

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_K \end{bmatrix},$$

gdzie elementy diagonalne macierzy  $\mathbf{W}$  oznaczają dowolnie małe dodatnie liczby  $w_i > 0$  ( $i=1, \dots, K$ ).

Następnie obliczamy sumę macierzy  $\mathbf{X}'\mathbf{X}$  (o wymiarach  $K \times K$ ) i macierzy  $\mathbf{W}$  o wymiarach ( $K \times K$ ), co wyraźnie zmniejsza trudności numeryczne przy wyznaczaniu macierzy  $(\mathbf{X}'\mathbf{X})^{-1}$ . W rezultacie wektor  $\hat{\boldsymbol{\beta}}$  dany wzorem (2) możemy napisać w postaci<sup>4</sup>

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + \mathbf{W})^{-1} \mathbf{X}'\mathbf{y}. \quad (5)$$

Zauważmy, że przy takim postawieniu sprawy efektywność estymatorów parametrów strukturalnych modelu (2) wyraźnie rośnie, ale równocześnie estymatory te stają się obciążone.

Powstaje w związku z tym zagadnienie znalezienia kryterium, jakim należy się kierować przy optymalnym ustalaniu elementów diagonalnych macierzy  $\mathbf{W}$ . Odpowiednie procedury postępowania są szczegółowo opisane w cytowanej już pracy A. E. Hoerla i R. W. Kennarda<sup>5</sup>, a także w pracy D. K. Guilkey i J. L. Murphy<sup>6</sup>, które pozwalają na zminimalizowanie wartości obciążeń estymatorów parametrów strukturalnych modelu (2).

Drugim niebezpieczeństwem związanym z estymacją modelu (2) jest utrata informacji statystycznej. Przyjmijmy, że dysponujemy danymi statystycznymi o charakterze szeregów czasowych<sup>7</sup>, potrzebnymi do esty-

<sup>4</sup> Przykład takiego postępowania podany jest w pracy A. E. Hoerla i R. W. Kennarda, *Ridge Regression: Based Estimation for Nonorthogonal Problems*, Technometrics 1970, vol. 12.

<sup>5</sup> Ibidem.

<sup>6</sup> D. K. Guilkey, J. L. Murphy, *Directed Ridge Regression Techniques in Cases of Multicollinearity*, Journal of the American Statistical Association 1975. vol. 70.

<sup>7</sup> Tego rodzaju dane charakteryzują się tym, że liczba ich nie jest duża, tzn. odpowiednie szeregi czasowe liczą po kilkanaście elementów.

macji modelu (2). Weźmy jako przykład zależność występującą między popytem na dobra konsumpcyjne trwałego użytku ( $Y_t$ ) i dochodem ( $X_t$ ). Chcemy teraz model (2) oszacować na podstawie danych za lata 1960 - 1975, przy czym  $t=1$  dla 1960 i  $t=n=16$  dla 1975 r. Jednocześnie zakładamy, że liczba  $k$  opóźnień w czasie wartości zmiennej objaśniającej  $X_t$  jest równa np. pięć<sup>8</sup> ( $k=5$ ). Dane te przedstawione są w tabeli 1.

Tabela 1

$Y_t$	$X_t$	$X_{t-1}$	$X_{t-2}$	...	$X_{t-k}$
$y_1$	$x_1$	—	—	...	—
$y_2$	$x_2$	$x_1$	—	...	—
$y_3$	$x_3$	$x_2$	$x_1$	...	—
$y_4$	$x_4$	$x_3$	$x_2$	...	—
$y_5$	$x_5$	$x_4$	$x_3$	...	—
$y_6$	$x_6$	$x_5$	$x_4$	...	$x_1$
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
$y_n$	$x_n$	$x_{n-1}$	$x_{n-2}$	...	$x_{n-k}$

Jak widać, ocena modelu o równaniu

$$Y_t = \alpha + \sum_{j=0}^5 \beta_j X_{t-j}$$

oparta będzie na  $n-k=11$  obserwacji, a więc nastąpiła poważna utrata informacji, co w sposób negatywny musi rzutować na wiarygodność końcowych wniosków. Wydłużanie w badaniach odcinka czasowego ( $t_1, t_n$ ) jest zwykle w praktyce niemożliwe, gdyż dane najczęściej są niejednorodne na skutek zachodzących w gospodarce zmian strukturalnych, instytucjonalnych itd. Przejście na mniejsze jednostki czasu (miesiące lub kwartały) nie rozwiązuje trudności, bowiem pojawić się mogą efekty sezonowe wymagające uwzględnienia dodatkowych zmiennych objaśniających.

Należy dodać, że w przypadkach gdy liczba uwzględnionych opóźnień w modelu jest duża ( $k > 5$ ), estymacja modelu (2) napotyka na poważne trudności, a czasem jest wręcz niemożliwa do zrealizowania. Stwarza to konieczność stosowania określonego postępowania, polegającego na wprowadzeniu odpowiednich ograniczeń na jego parametry. Przyjmuje się zwykle, że parametry  $\beta_1, \dots, \beta_k$  przy kolejnych opóźnionych w czasie zmiennych objaśniających systematycznie maleją wraz z przechodzeniem

<sup>8</sup> Ustalenie optymalnej wartości parametru  $k$  nie jest proste. Decyzje co do wyboru wartości tego parametru muszą być oparte na metodzie prób i błędów, a także na pewnym wyczuciu subiektywnym, na intuicji ze strony ekonometryka.

do coraz odleglejszych okresów czasu. Szczególnie często korzysta się tu z założenia sformułowanego i rozwiniętego przez L. M. Koycka<sup>9</sup>, który przyjmuje a priori, że parametry  $\beta_j$  ( $j=1, \dots, k$ ) maleją w postępie geometrycznym. Mamy zatem

$$\beta_j = \beta_0 \lambda^j \quad (j=1, \dots, k), \quad (6)$$

gdzie  $\lambda$  jest stałą, przy czym  $0 < \lambda < 1$ .

Podstawiając (6) do zależności (2) otrzymujemy

$$Y_t = \alpha + \beta_0 X_t + (\beta_0 \lambda) X_{t-1} + (\beta_0 \lambda^2) X_{t-2} + \dots + (\beta_0 \lambda^k) X_{t-k} + u_t. \quad (7)$$

Jeżeli teraz opóźnimy zmienne w (7) o przyjętą jednostkę czasu (np. rok) i pomnożymy przez  $\lambda$ , to otrzymamy

$$\begin{aligned} \lambda Y_{t-1} = \lambda \alpha + (\lambda \beta_0) X_{t-1} + (\lambda^2 \beta_0) X_{t-2} + (\lambda^3 \beta_0) X_{t-3} + \dots + \\ + (\lambda^{k+1} \beta_0) X_{t-k-1} + \lambda u_{t-1}, \end{aligned} \quad (8)$$

co po odjęciu od (7) daje po uporządkowaniu

$$Y_t = (\alpha - \lambda \alpha) + \beta_0 X_t + \lambda Y_{t-1} - (\lambda^{k+1} \beta_0) X_{t-k-1} + (u_t - \lambda u_{t-1}). \quad (9)$$

Zakładamy teraz, że liczba  $k$  uwzględnionych opóźnień w modelu (2) jest duża ( $k > 5$ ). W takim razie ostatni składnik występujący we wzorze (9) staje się liczbą relatywnie małą i maleje do zera, gdy  $k \rightarrow \infty$ . Wobec tego w zastosowaniach składnik ten można pominąć i wtedy wystarczającym przybliżeniem zależności (9) jest

$$Y_t = \alpha^* + \beta_0 X_t + \lambda Y_{t-1} + v_t, \quad (10)$$

gdzie  $\alpha^* = (\alpha - \lambda \alpha)$  oraz  $v_t = (u_t - \lambda u_{t-1})$ ,  $0 < \lambda < 1$ .

Zauważmy zatem, że przy założeniu (6) zależność o postaci (2) może być zredukowana do zależności (10), w której występują tylko dwie zmienne objaśniające (z których jedna ( $Y_{t-1}$ ) jest opóźnioną wartością zmiennej endogenicznej) i tylko trzy parametry  $\alpha^*$ ,  $\beta_0$  i  $\lambda$  podlegają estymacji<sup>10</sup>, w tym przypadku na podstawie  $n-1$  obserwacji.

Rozważmy następnie problem estymacji zależności (10). W praktyce badawczej dla równania (10) często przyjmuje się założenia:

- 1)  $E(v_t) = 0$ ,
- 2)  $E(v_t)^2 = \sigma_v^2$ ,
- 3)  $\text{cov}(v_t, v_{t-j}) = 0 \quad (j=1, \dots, k)$ ,
- 4)  $\text{cov}(v_t, X_t) = \text{cov}(v_t, Y_{t-1}) = 0$ ,

a następnie wyznacza wartości estymatorów  $\hat{\alpha}^*$ ,  $\hat{\beta}$ ,  $\hat{\lambda}$  klasyczną metodą

<sup>9</sup> L. M. Koyck, *Distributed Lags and Investment Analysis*, Amsterdam 1954.

<sup>10</sup> Warto zauważyć, iż w przypadku ogólnym, model (1) może być także zredukowany do zależności (9).

najmniejszych kwadratów. Wektor ocen parametrów strukturalnych równania (10) dany jest wzorem

$$\begin{bmatrix} \hat{\alpha}^* \\ \hat{\beta}_0 \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} (n-1) & \sum_{t=2}^n x_t & \sum_{t=2}^n y_{t-1} \\ \sum_{t=2}^n x_t & \sum_{t=2}^n x_t^2 & \sum_{t=2}^n (x_t y_{t-1}) \\ \sum_{t=2}^n y_{t-1} & \sum_{t=2}^n (x_t y_{t-1}) & \sum_{t=2}^n y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=2}^n y_t \\ \sum_{t=2}^n (x_t y_t) \\ \sum_{t=2}^n (y_{t-1} y_t) \end{bmatrix}. \quad (11)$$

Zatem estymator

$$\hat{\beta}_j = (\hat{\lambda}^j) \beta_0 \quad (j=1, \dots, k), \quad (12)$$

co po uwzględnieniu  $\alpha^* = (\alpha - \lambda\alpha)$  daje  $\alpha = [\alpha^*/(1-\lambda)]$ , Z czego wynika, że estymator  $\hat{\alpha}$  parametru  $\alpha$  jest

$$\hat{\alpha} = \frac{\hat{\alpha}^*}{1-\hat{\lambda}}. \quad (13)$$

Oszacowany model ma zatem postać

$$y_t = \hat{\alpha}^* + \hat{\beta}_0 x_t + \hat{\lambda} y_{t-1} + \hat{v}_t, \quad (14)$$

gdzie  $\hat{\alpha}^* = \hat{\alpha}(1-\hat{\lambda})$ , a  $\hat{v}_t = (u_t - \hat{\lambda}u_{t-1})$ .

Zapisany w postaci (10) model Koycka stanowi niewątpliwie postęp w stosunku do klasycznej procedury estymacyjnej opartej o model (2), niemniej jednak łatwo wykazać, iż przyjmowane założenia 2) - 4) nie są prawdziwe. Okazuje się, że opóźniona o jednostkę czasu zmienna endogeniczna  $Y_{t-1}$  jest skorelowana ze składnikiem losowym  $v_t = (u_t - \lambda u_{t-1})$ , a to — jak wiadomo — kryje w sobie bardzo istotne niebezpieczeństwo, gdyż można wykazać<sup>11</sup>, że wówczas estymatory  $\hat{\alpha}$ ,  $\hat{\beta}_0$  i  $\hat{\lambda}$  parametrów strukturalnych modelu (10) nie są ani nieobciążone, ani zgodne, a więc są praktycznie bezużyteczne. Istotnie, z zależności (10) otrzymujemy:

$$\begin{aligned} E(v_t Y_{t-1}) &= E[(u_t - \lambda u_{t-1})(\alpha^* + \beta_0 X_{t-1} + \lambda Y_{t-2} + v_{t-1})] = -\lambda \sigma_u^2 \neq 0, \\ E(v_t)^2 &= E(u_t - \lambda u_{t-1})^2 = (1 + \lambda^2) \sigma_u^2, \\ E(v_t v_{t-1}) &= E[(u_t - \lambda u_{t-1})(u_{t-1} - \lambda u_{t-2})] = \\ &= E(u_t u_{t-1} - \lambda u_t u_{t-2} - \lambda u_{t-1}^2 + \lambda^2 u_{t-1} u_{t-2}) = -\lambda \sigma_u^2 \neq 0, \end{aligned} \quad (15)$$

co oznacza, że składnik losowy równania (10) jest skorelowany w czasie. Wykorzystując (15), widzimy, że

<sup>11</sup> Dowód ten można znaleźć w wielu pracach. Por. np. pracę A. S. Goldbergera, *Teoria ekonometrii*, Warszawa 1972, ss. 354 - 356.

$$E(\mathbf{v}\mathbf{v}') = \sigma_u^2 \begin{bmatrix} 1+\lambda^2 & -\lambda & 0 & \dots & 0 \\ -\lambda & 1+\lambda^2 & -\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1+\lambda^2 \end{bmatrix}. \quad (16)$$

Elementy macierzy  $E(\mathbf{v}\mathbf{v}')$  (o wymiarach  $n-1 \times n-1$ ) wyznaczono na podstawie wyrażenia

$$E(v_t v_{t+s}) = \begin{cases} -\lambda \sigma_u^2 & \text{dla } s = \pm 1 \\ 0 & \text{dla } |s| \geq 2 \end{cases} \quad \text{dla wszystkich } t.$$

Reasumując wyniki tej dyskusji można stwierdzić, że w przypadku, gdy mamy relację postaci (10), opóźniona o jednostkę czasu zmienna endogeniczna jest skorelowana ze składnikiem losowym. Zauważmy dalej, że składniki losowe rozpatrywanego modelu nie są od siebie stochastycznie niezależne<sup>12</sup>, a ich macierz  $E(\mathbf{v}\mathbf{v}')$  wariancji i kowariancji dana jest wzorem (16). W takich sytuacjach klasyczna metoda najmniejszych kwadratów zastosowana do modelu (10) daje estymatory parametrów, które nie tylko nie są nie obciążone, ale nawet nie są zgodne.

Aby otrzymać zgodne estymatory parametrów strukturalnych modelu (10), L. Koyck<sup>13</sup> proponuje dwustopniową metodę estymacji, która wymaga jednak wykonania wielu bardzo skomplikowanych obliczeń, gdyż oceny parametrów  $\alpha^*$ ,  $\beta_0$  oraz  $\lambda$  otrzymuje się przez rozwiązanie układu 3 równań nieliniowych i w związku z tym dla efektywnego korzystania z niej potrzebna jest elektroniczna maszyna cyfrowa.

Naszym zdaniem, istnieje jeszcze jedna możliwość wyjścia z tego dylematu, a mianowicie przez posłużenie się metodą zmiennych instrumentalnych<sup>14</sup>, która dostarcza zgodne estymatory w sytuacjach, w których zawodzi metoda najmniejszych kwadratów. Zastosujemy więc do estymacji modelu (10) metodę zmiennych instrumentalnych, w której  $X_t$  będzie własną zmienną instrumentalną, a opóźniona zmienna objaśniająca  $X_{t-1}$  będzie zmienną instrumentalną dla  $Y_{t-1}$ . Zmienne te mają pożądane własności dobrych zmiennych instrumentalnych, a mianowicie są nieskorelowane ze składnikiem losowym  $v_t$  w tym sensie, że

$$\text{plim} \left( \frac{1}{n} \mathbf{Z}'\mathbf{v} \right) = \begin{bmatrix} \text{plim} \frac{1}{n} \sum_t v_t \\ \text{plim} \frac{1}{n} \sum_t (x_t v_t) \\ \text{plim} \frac{1}{n} \sum_t (x_{t-1} v_t) \end{bmatrix} = \mathbf{0}, \quad (17)$$

<sup>12</sup> Zależność taką nazywa się zwykle w literaturze autokorelacją składnika losowego.

<sup>13</sup> L. M. Koyck, op. cit.

<sup>14</sup> Por. pracę J. Johnstona, *Econometric Methods*, New York 1972, ss. 278 - 281.

gdzie  $Z$  jest macierzą obserwacji na tych zmiennych.

Zakładamy ponadto, że zmienne instrumentalne są skorelowane ze zmiennymi objaśniającymi  $X_t$  i  $Y_{t-1}$ , czyli, że

$$\text{plim} \left( \frac{1}{n} \mathbf{Z}'\mathbf{X} \right) = \Sigma_{ZX} \quad (18)$$

istnieje i jest macierzą nieosobliwą.

We wzorze tym  $\Sigma_{ZX}$  oznacza macierz kowariancji między zmiennymi instrumentalnymi oraz zmiennymi objaśniającymi w populacji.

Niezależnie od tego przyjmujemy, że

$$\text{plim} \left( \frac{1}{n} \mathbf{Z}'\mathbf{Z} \right) = \Sigma_{ZZ} \quad (19)$$

istnieje, przy czym  $\Sigma_{ZZ}$  jest macierzą kowariancji zmiennych instrumentalnych.

Weźmy następnie pod uwagę estymator wektora  $\boldsymbol{\beta}$  w omawianej metodzie. Wektor  $\mathbf{b}$  ocen parametrów  $\alpha^*$ ,  $\beta_0$  oraz  $\lambda$  modelu (10) znajdujemy rozwiązując następujący układ równań

$$\mathbf{Z}'\mathbf{X}\mathbf{b} = \mathbf{Z}'\mathbf{y} \quad (20)$$

względem  $\mathbf{b}$ , przy czym w naszym przypadku

$$\mathbf{Z} = \begin{bmatrix} 1 & x_2 & x_1 \\ 1 & x_3 & x_2 \\ \dots & \dots & \dots \\ 1 & x_n & x_{n-1} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_2 & y_1 \\ 1 & x_3 & y_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_{n-1} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \hat{\alpha}^* \\ \hat{\beta}_0 \\ \hat{\lambda} \end{bmatrix}.$$

Wektor  $\mathbf{b}$  jest więc równy

$$\mathbf{b} = (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{y}. \quad (21)$$

W prosty sposób można wykazać, że estymator  $\mathbf{b}$  jest w tym przypadku zgodny. Podstawiając  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v}$  do (21), otrzymujemy

$$\mathbf{b} = (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{v}) = \boldsymbol{\beta} + (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{v}.$$

Kontynuując nasze rozumowanie, otrzymujemy z kolei

$$\text{plim } \mathbf{b} = \boldsymbol{\beta} + \text{plim} \left( \frac{1}{n} \mathbf{Z}'\mathbf{X} \right)^{-1} \text{plim} \left( \frac{1}{n} \mathbf{Z}'\mathbf{v} \right),$$

skąd ze względu na (17) i (18) dochodzimy do wyrażenia

$$\text{plim } \mathbf{b} = \boldsymbol{\beta} + \Sigma_{ZX}^{-1} \cdot \mathbf{0} = \boldsymbol{\beta}.$$

Ocena asymptotycznej macierzy wariancji i kowariancji tych estymatorów dana jest wzorem

$$\text{asy var}(\mathbf{b}) = \sigma_v^2 (\mathbf{Z}'\mathbf{X})^{-1} (\mathbf{Z}'\mathbf{Z}) (\mathbf{X}'\mathbf{Z})^{-1}, \quad (22)$$

gdzie estymatorem zgodnym wariancji składnika losowego  $\sigma_v^2$  jest

$$\hat{\sigma}_v^2 = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) / (n' - p) = \hat{\mathbf{v}}' \hat{\mathbf{v}} / (n' - p), \quad (23)$$

gdzie  $\hat{\mathbf{v}} = \mathbf{y} - \mathbf{X}\mathbf{b}$ ,  $n' = n - 1$ , a  $p$  oznacza liczbę szacowanych parametrów występujących w modelu (10); w przedstawionym modelu  $p = 3$ .

Tak więc metoda zmiennych instrumentalnych dostarczyła estymatorów zgodnych parametrów  $\alpha^*$ ,  $\beta_0$  oraz  $\lambda$  równania (10). Aby metodę tę można było efektywnie wykorzystywać w badaniach empirycznych, muszą być oczywiście spełnione założenia (17), (18) i (19).

Na zakończenie tej części rozważań należy podkreślić, że z warunkiem (18) wiąże się wymóg występowania dostatecznie silnej korelacji między zmiennymi instrumentalnymi i zmiennymi objaśniającymi, w innym bowiem przypadku wprowadzone otrzymane estymatory będą zgodne, ale efektywność ich będzie bardzo niska.

W tym miejscu przedstawimy nieco inną metodę estymacji modelu (2), w którym zmienna endogeniczna zależy od opóźnionych zmiennych objaśniających. Metoda ta opracowana została przez S. Almon<sup>15</sup> i — jak sądzę — zdała egzamin w praktyce. Zaproponowana przez Almon metoda polega na tym, że zamiast założenia (6) rozpatruje się nowe, bardziej elastyczne założenie, a mianowicie

$$\beta_j \triangleq f(j) \quad (j = 0, 1, \dots, k), \quad (24)$$

gdzie  $f(j)$  jest wielomianem o postaci

$$f(j) \triangleq \alpha_0 + \alpha_1 j + \alpha_2 j^2 + \dots + \alpha_p j^p \quad (j = 0, 1, \dots, k), \quad (25)$$

$$k \geq p$$

co oznacza, iż przyjmujemy, że ciąg parametrów strukturalnych  $\beta_j$  ( $j = 0, 1, \dots, k$ ) równania (2) można aproksymować za pomocą wielomianu stopnia skończonego<sup>16</sup>. Podstawowym problemem do rozwiązania jest określenie stopnia wielomianu, to znaczy określenie wartości  $p$ . Ustalenie wartości  $p$  nie jest jednak proste. Decyzje co do wyboru stopnia wielomianu najczęściej oparte są na metodzie kolejnych przybliżeń. W zastosowaniach zwykle funkcja (25) jest wielomianem stopnia drugiego, trzeciego lub wyjątkowo wielomianem stopnia czwartego.

<sup>15</sup> S. Almon, *The Distributed Lag between Capital Appropriations and Expenditures*, *Econometrica* 1965, vol. 30.

<sup>16</sup> Wynika to z twierdzenia Weierstrassa głoszącego, że każdą funkcję ciągłą i ograniczoną można aproksymować z dowolną dokładnością w przedziale skończonym za pomocą wielomianu stopnia skończonego. Dowód tego twierdzenia znaleźć można np. w pracy B. R. Mortona, *Numerical Approximation*, London 1964, s. 4.



Wprowadzamy następujące oznaczenia:

$$Z_{1t} = \sum_{j=0}^k X_{t-j}, \quad Z_{2t} = \sum_{j=1}^k jX_{t-j} \quad \text{oraz} \quad Z_{3t} = \sum_{j=1}^k j^2 X_{t-j}. \quad (30)$$

Wówczas równanie (29) możemy napisać w równoważnej postaci

$$Y_t = \alpha + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t. \quad (31)$$

Zakładamy, że dla równania (31) spełnione są wymienione na początku pracy założenia 1) - 5). Dla wyznaczenia wartości ocen parametrów posłużymy się więc metodą najmniejszych kwadratów. Poszukiwany wektor ocen jest równy

$$\mathbf{a} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}, \quad (32)$$

gdzie  $\mathbf{a} = \{\hat{\alpha} \hat{\alpha}_0 \hat{\alpha}_1 \hat{\alpha}_2\}$  oznacza wektor  $(p+2 \times 1)$  ocen parametrów strukturalnych modelu (31),  $\mathbf{Z}$  jest macierzą (o wymiarach  $n-k \times p+2$ ) wyznaczonych wartości zmiennych objaśniających<sup>18</sup>, a  $\mathbf{y}$  oznacza wektor  $(n-k \times 1)$  zaobserwowanych wartości zmiennej endogenicznej.

Macierz wariancji i kowariancji estymatorów  $\hat{\alpha}$ ,  $\hat{\alpha}_0$ ,  $\hat{\alpha}_1$  oraz  $\hat{\alpha}_2$  przedstawić można za pomocą równania

$$\text{var}(\mathbf{a}) = \sigma_u^2 (\mathbf{Z}'\mathbf{Z})^{-1}. \quad (33)$$

Nie obciążonym estymatorem parametru  $\sigma_u^2$  jest estymator  $\hat{\sigma}_u^2$ :

$$\hat{\sigma}_u^2 = (\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}}) / (n' - p - 2), \quad (34)$$

gdzie  $\mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{Z}\mathbf{a}$ , przy czym  $\mathbf{a}$  jest zdefiniowane wzorem (32),  $n' = n - k$ , a  $p$  oznacza stopień wielomianu.

Z kolei wykorzystujemy zależność (27): podstawiając w niej w miejsce  $\alpha_0$ ,  $\alpha_1$  oraz  $\alpha_2$  wyznaczone oszacowania  $\hat{\alpha}_0$ ,  $\hat{\alpha}_1$  i  $\hat{\alpha}_2$ , otrzymujemy oceny parametrów strukturalnych modelu (2). Mamy zatem

$$\begin{aligned} \hat{\beta}_0 &= \hat{\alpha}_0, \\ \hat{\beta}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2, \\ \hat{\beta}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2, \\ &\dots \dots \dots \\ \hat{\beta}_k &= \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2. \end{aligned} \quad (35)$$

W rezultacie, po dokonaniu tych obliczeń, oszacowane równanie (2) przybiera postać

$$y_t = \hat{\alpha} + \hat{\beta}_0 x_t + \hat{\beta}_1 x_{t-1} + \dots + \hat{\beta}_k x_{t-k} + \hat{u}_t. \quad (36)$$

Na koniec, wyznaczamy wariancje estymatorów  $\hat{\beta}_j (j=0, 1, \dots, k)$  danych wzorem (35). Otrzymujemy wówczas

<sup>18</sup> Ponieważ w modelu (31) występuje wyraz wolny  $\alpha$ , należy przyjąć, że mamy w rzeczywistości cztery zmienne objaśniające, a mianowicie  $Z_{0t}$ ,  $Z_{1t}$ ,  $Z_{2t}$ ,  $Z_{3t}$ , przy czym ta pierwsza zmienna może przyjmować tylko wartość 1.

$$\begin{aligned} \text{var}(\hat{\beta}_j) &= \sigma_u^2 \mathbf{h}_j (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{h}_j' = \\ &= \mathbf{h}_j \begin{bmatrix} \text{var}(\hat{\alpha}_0) & \text{cov}(\hat{\alpha}_0, \hat{\alpha}_1) & \text{cov}(\hat{\alpha}_0, \hat{\alpha}_2) \\ \text{cov}(\hat{\alpha}_1, \hat{\alpha}_0) & \text{var}(\hat{\alpha}_1) & \text{cov}(\hat{\alpha}_1, \hat{\alpha}_2) \\ \text{cov}(\hat{\alpha}_2, \hat{\alpha}_0) & \text{cov}(\hat{\alpha}_2, \hat{\alpha}_1) & \text{var}(\hat{\alpha}_2) \end{bmatrix} \mathbf{h}_j' \quad (j=0, 1, \dots, k), \end{aligned} \quad (37)$$

gdzie estymatorem parametru  $\sigma_u^2$  jest estymator  $\hat{\sigma}_u^2$  wyznaczony na podstawie wzoru (34), a  $\mathbf{h}_j (j=0, 1, \dots, k)$  oznacza wektor wierszowy współczynników występujących przy  $\hat{\alpha}_0, \hat{\alpha}_1$  oraz  $\hat{\alpha}_2$  w (35). Np. dla  $\hat{\beta}_2$  mamy

$$\mathbf{h}_2 = [1 \ 2 \ 4].$$

Drogą bezpośredniego mnożenia znajdujemy, że wariancja np. estymatora  $\hat{\beta}_2$  jest

$$\begin{aligned} \text{var}(\hat{\beta}_2) &= \text{var}(\hat{\alpha}_0) + 4 \text{var}(\hat{\alpha}_1) + 16 \text{var}(\hat{\alpha}_2) + 4 \text{cov}(\hat{\alpha}_0, \hat{\alpha}_1) + 8 \text{cov}(\hat{\alpha}_0, \hat{\alpha}_2) + \\ &+ 16 \text{cov}(\hat{\alpha}_1, \hat{\alpha}_2) = \text{var}(\hat{\alpha}_0) + 4 \text{var}(\hat{\alpha}_1) + 16 \text{var}(\hat{\alpha}_2) + \\ &+ 4\hat{\rho}_{01} \sqrt{\text{var} \hat{\alpha}_0 \text{var} \hat{\alpha}_1} + 8\hat{\rho}_{02} \sqrt{\text{var} \hat{\alpha}_0 \text{var} \hat{\alpha}_2} + \\ &+ 16\hat{\rho}_{12} \sqrt{\text{var} \hat{\alpha}_1 \text{var} \hat{\alpha}_2}, \end{aligned} \quad (38)$$

gdzie przez  $\hat{\rho}_{01}, \hat{\rho}_{02}$  oraz  $\hat{\rho}_{12}$  oznaczono współczynniki korelacji liniowej między estymatorami parametrów  $\alpha_0, \alpha_1$  oraz  $\alpha_2$ . Współczynniki te przy przyjętych oznaczeniach, są zdefiniowane następująco:

$$\hat{\rho}_{ij} = \frac{\text{cov}(\hat{\alpha}_i, \hat{\alpha}_j)}{\sqrt{\text{var} \hat{\alpha}_i \text{var} \hat{\alpha}_j}} \quad \begin{matrix} (i, j=0, 1, 2) \\ (i \neq j) \end{matrix} \quad (39)$$

Można wykazać, iż opisane dwuetapowe postępowanie daje estymatory parametrów strukturalnych modelu (2) efektywniejsze (o mniejszych wariancjach) w porównaniu z oszacowaniami uzyskanymi według metody najmniejszych kwadratów zastosowanej bezpośrednio do rozpatrywanego w niniejszej pracy modelu.

Na zakończenie tych rozważań zauważmy jeszcze, iż w praktyce często rozpatrujemy model postaci

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_k X_{t-k} + \gamma_1 W_{1t} + \dots + \gamma_r W_{rt} + u_t, \quad (40)$$

gdzie  $W_{1t}, \dots, W_{rt}$  są dodatkowymi zmiennymi objaśniającymi. Przykładowo przyjmiemy, że  $\beta_j = \alpha_0 + \alpha_1 j + \alpha_2 j^2$  ( $j=0, 1, \dots, k$ ), to postępując analogicznie jak poprzednio, można — podobnie jak tam — wykazać, że zależność (40) redukuje się do równania (31), rozszerzonego o nowo wprowadzone zmienne  $W_{it}$  ( $i=1, \dots, r$ ), czyli<sup>19</sup>

<sup>19</sup> To rozwiązanie nie może być jednak przyjęte tam, gdzie liczba obserwacji w próbie jest mała, gdyż wtedy wariancje poszczególnych estymatorów mogłyby okazać się bardzo wysokie.

$$Y_t = \alpha + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + \sum_{i=1}^r \gamma_i W_{it} + u_t. \quad (41)$$

Stosując metodę najmniejszych kwadratów do tego równania otrzymuje się efektywniejsze estymatory nieznanych parametrów  $\beta_j$  ( $j=0, 1, \dots, k$ ).

## ESTIMATION PROBLEMS OF PARAMETERS OF ECONOMETRIC MODELS WITH DISTRIBUTED LAGS

### Summary

Let us assume that a linear relationship exists between a variable  $Y_t$  and  $k-1$  explanatory variables  $X_t, X_{t-1}, \dots, X_{t-k}$  and a disturbance term  $u$ . If we have a sample of  $n$  observations on  $Y$  and  $X$ 's we can write

$$Y_t = \alpha + \sum_{j=0}^k \beta_j X_{t-j} + u_t \quad (t=1, \dots, n). \quad (1)$$

The  $\alpha, \beta_0, \beta_1, \dots, \beta_k$  coefficients and the parameters of the  $u$  distribution are unknown and our problem is to obtain estimates of these unknowns. Under the usual assumption about the distribution of  $u$  and the independence of  $X$  and  $u$  there are in principle no new estimation problems in this model. Least-squares will give best linear unbiased estimates, if the model has been specified correctly. Several difficulties, however, are likely to arise in practice. First of all one cannot really expect any precise and firm indication from theory of the length of lag to be incorporated; rather one hopes to determine the lag from the data by fitting a fairly lag and then examining the significance of the coefficients of various lagged values of  $X$ . But this in turn raises two main statistical difficulties; one is that observations are lost due to the lags and the other is that typically the various lagged values of  $X$  will be highly intercorrelated leading to very imprecise estimates of the lagged coefficients and great difficulty in making useful inferences about them.

These difficulties have led to the a priori imposition of various assumptions about the form of the weights  $\beta_0, \dots, \beta_k$  in an attempt to produce a more amenable estimation problem involving fewer than  $k+1$  parameters. The above paper discusses three models for distributed lag analysis that either reduce the number of observations lost due to lagging and/or reduce the number of parameters to be estimated.