

13. Modele VAR

13.1. Metodologia klasyczna

- Podział na zmienne z egzo i endogeniczne jest z góry znany
- Forma funkcyjna modelu jest z góry znana
- Zmienne w modelu są stacjonarne

13.2. Porażka modeli klasycznych

- Prognozy uzyskiwane z dużych modeli wielorównaniowych w praktyce często gorsze od prognoz uzyskiwanych za pomocą metodologii Boxa-Jenkinsa (*ARIMA*) mimo zawartej w nich *explicite* wiedzy *a priori*.

Źródła metodologii VAR

- Niezadowolenie z klasycznych modeli wielorównaniowych uzyskiwanych za pomocą metod Komisji Cowlesa
- Lovell - problemem w przypadku większości modeli jest uzyskiwanie wysokiego R^2 i statystyk t za pomocą nieustrukturyzowanego przekopywania danych
- Krytyka Simsa - niepowodzenie klasycznych modeli wielorównaniowych wzięło się z identyfikacji z równań za pomocą arbitralnych i nie wynikających z teorii ograniczeń narzucanych na poszczególne równania celem uzyskania ich identyfikacji

Odpowiedź Simsa

- Często nie da się z góry uznać jednych zmiennych za egzo a drugich za endogeniczne - wszystko zależy od wszystkiego
- Większą wagę trzeba przywiązać do analizy dynamicznych własności modelu - w szczególności jego reakcji dynamicznej na szoki a mniejszą do analizy form funkcyjnych

13.3. Model VAR

- **Strukturalny bez ograniczeń**

$$A\mathbf{x}_t = B_1\mathbf{x}_{t-1} + \dots + B_k\mathbf{x}_{t-k} + \Psi D_t + \mathbf{u}_t \quad (13.1)$$

$$\mathbf{u}_t \sim N(\mathbf{0}, \Sigma)$$

gdzie X_t jest wektorem losowym $p \times 1$, a D_t jest wektorem nielosowym o wymiarach $m \times 1$, Σ jest macierzą symetryczną $p \times p$ a wartości początkowe procesu są pewnymi zmiennymi losowymi. Zauważmy, że parametry w tym modelu są w ogólnym przypadku niezidentyfikowane.

- **Forma zredukowana**

Mnożąc (13.1) lewostronnie przez A^{-1} uzyskujemy:

$$\mathbf{x}_t = A^{-1}B_1\mathbf{x}_{t-1} + \dots + A^{-1}B_k\mathbf{x}_{t-k} + A^{-1}\Psi D_t + \mathbf{u}_t$$

definiując $\Pi_i = A^{-1}B_i$, $\Phi = A^{-1}\Psi$, $\varepsilon_t = A^{-1}\mathbf{u}_t$, $\Omega = A^{-1}\Sigma A'^{-1}$ otrzymujemy typową postać VAR, którą traktować można jako formę zredukowaną modelu (13.1):

$$\mathbf{x}_t = \Pi_1\mathbf{x}_{t-1} + \dots + \Pi_k\mathbf{x}_{t-k} + \Phi D_t + \varepsilon_t, \quad t = 1 \dots T \quad (13.2)$$

gdzie

$$\varepsilon_t \sim \text{NID}(0, \Omega)$$

- **Identyfikacja Simsa**

Załóżmy, że C jest macierzą Choleskiego dla macierzy Ω . Wtedy

$$C\Omega C' = \Sigma_D \quad (13.3)$$

i macierz Σ_D jest diagonalna. Mnożąc lewostronnie równanie (??) przez C otrzymujemy

$$\begin{aligned} C\mathbf{x}_t &= C\Pi_1\mathbf{x}_{t-1} + \dots + C\Pi_k\mathbf{x}_{t-k} + C\boldsymbol{\epsilon}_t \\ &= \Psi_1\mathbf{x}_{t-1} + \dots + \Psi_k\mathbf{x}_{t-k} + \boldsymbol{\eta}_t \end{aligned}$$

gdzie

$$\boldsymbol{\eta}_t = C\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_D)$$

Forma ta jest formą strukturalną dla $\Psi_i = C\Pi_i$ jeśli tylko ograniczenia implikowane przez (13.3) rzeczywiście identyfikują strukturę modelu. Macierz Choleskiego jest macierzą dolnotrójkątną, więc model ma następującą postać

$$\begin{aligned} x_{1t} &= \sum_{i=1}^k \psi_{1i}x_{t-i} + \eta_{1t} \\ x_{2t} &= -c_{21}x_{1t} + \sum_{i=1}^k \psi_{2i}x_{t-i} + \eta_{2t} \\ &\dots \\ x_{Gt} &= -\sum_{i=1}^{G-1} c_{Gi}x_{it} + \sum_{i=1}^k \psi_{Gi}x_{t-i} + \eta_{Gt} \end{aligned}$$

gdzie π_{ij} jest j -tym wierszem macierzy Π_i . W rezultacie x_{1t} zależy jedynie od zaburzenia losowego η_{1t} , x_{2t} zależy bezpośrednio od zaburzenia losowego η_{2t} a pośrednio, poprzez x_{1t} od zaburzenia losowego η_{1t} , etc. Oznacza to, że x_{it} jest zmienną z góry określoną w stosunku do x_{jt} jeśli $i < j$. Implikuje to pewną strukturę przyczynowości, w której x_{it} może natychmiastowo wpływać na x_{jt} ale nie na odwrót. Jest rzeczą sporną, czy tego rodzaju strukturę przyczynowości można wywnioskować z teorii ekonomii.

Jeśli jednak model jest prawidłowo zidentyfikowany, to elementy η_t są wzajemnie niezależne i η_{it} może być interpretowane jako strukturalny szok odnoszący się do zmiennej x_{it} . Oszacowane w tak zidentyfikowanym modelu reszty mogą być traktowane jako estymatory szoków strukturalnych i mogą być analizowane pod kątem ich wpływu na zachowanie zmiennych endogenicznych. Często porównuje się także przebieg tak oszacowanych szoków (np. podażowych i popytowych) z przebiegiem wypadków historycznych.

13.4. Przyczynowość

- Nie istnieje natychmiastowa przyczynowość
- Nie istnieje natychmiastowa wzajemna przyczynowość
- Przyszłość nie może powodować teraźniejszości
- **Testowanie przyczynowości**
 - Test Grangera
 - Test Simsa

13.5. Stacjonarność, integracja i kointegracja

Definicja 13.1 *Proces stochastycznym jest silnie stacjonarny, jeśli łączna dystrybuanta tego procesu spełnia warunek*

$$F(\mathbf{x}_{t_1}, \mathbf{x}_{t_2}, \dots, \mathbf{x}_{t_m}) = F(\mathbf{x}_{t_1+h}, \mathbf{x}_{t_2+h}, \dots, \mathbf{x}_{t_m+h})$$

dla dowolnych t_1, t_2, \dots, t_m i h i $\mathbf{x}_{t_1} = \mathbf{x}_{t_1+h}, \mathbf{x}_{t_2} = \mathbf{x}_{t_2+h}, \dots, \mathbf{x}_{t_m} = \mathbf{x}_{t_m+h}$.

Definicja 13.2 *Słaba stacjonarność*

$$\text{Var}(\mathbf{x}_t) = \Sigma_X < \infty$$

$$\text{Cov}(\mathbf{x}_{t_1}, \mathbf{x}_{t_2}) = \text{Cov}(\mathbf{x}_{t_1+h}, \mathbf{x}_{t_2+h})$$

dla dowolnych t_1, t_2 i h .

Uwaga 13.3 Dla procesów gaussowskich mocna stacjonarność jest równoważna słabej stacjonarności.

Uwaga 13.4 O procesie \mathbf{x}_t , w którym występują trendy deterministyczne, przyjęło się mówić, że jest stacjonarny jeśli $\mathbf{x}_t - \mathbb{E}(\mathbf{x}_t)$ jest procesem stacjonarnym.

Definicja 13.5 Proces liniowy definiujemy jako $\mathbf{x}_t = \mathbf{C}(L)\boldsymbol{\epsilon}_t = \sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\epsilon}_{t-i}$, $t = 0, 1, \dots$ gdzie szereg $\mathbf{C}(z) = \sum_{i=0}^{\infty} \mathbf{C}_i z^i$ jest zbieżny dla $|z| \leq 1 + \delta$ dla pewnego $\delta > 0$ a $\boldsymbol{\epsilon}_i$ ciągiem niezależnych i o identycznych rozkładach p -wymiarowych zmiennych losowych o wartościach oczekiwanych równych zero i macierzach wariancji kowariancji równych $\boldsymbol{\Omega}$.

Definicja 13.6 Proces stochastyczny \mathbf{x}_t nazywamy procesem $I(1)$, jeśli $\mathbf{x}_t - \mathbb{E}(\mathbf{x}_t)$ jest procesem liniowym i $\|\mathbf{C}(1)\| < \infty$, gdzie $\mathbf{C}(1) = \sum_{i=0}^{\infty} \mathbf{C}_i$.

Lemat 13.7 Proces liniowy \mathbf{x}_t jest słabo stacjonarny rzędu 2.

Dowód. Wartość oczekiwana dla zmiennej $I(0)$ równa jest 0. Z kolei wariancja i kowariancje są równe

$$\begin{aligned} \mathbb{E}(\mathbf{x}_t, \mathbf{x}_{t+h}) &= \mathbb{E} \left[\left(\sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\epsilon}_{t-i} \right) \left(\sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\epsilon}_{t-i+h} \right)' \right] \\ &= \sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\Omega} \mathbf{C}'_{i+h} \end{aligned}$$

dla $h = 0, 1, \dots$ Zauważmy, że

$$\left\| \sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\Omega} \mathbf{C}'_{i+h} \right\| \leq \left\| \left(\sum_{i=0}^{\infty} \mathbf{C}_i \right) \boldsymbol{\Omega} \left(\sum_{i=0}^{\infty} \mathbf{C}_i \right)' \right\| = \|\mathbf{C}\boldsymbol{\Omega}\mathbf{C}\| < \infty$$

i $\sum_{i=0}^{\infty} \mathbf{C}_i \boldsymbol{\Omega} \mathbf{C}'_{i+h}$ zależy jedynie od h . ■

Definicja 13.8 *Proces stochastyczny \mathbf{x}_t uważamy za zintegrowany rzędu d , co oznaczamy jako $\mathbf{x}_t \sim I(d)$, $d = 0, 1, 2, \dots$ jeśli $\Delta^d \mathbf{x}_t$ jest $I(0)$.*

Definicja 13.9 *Niech \mathbf{x}_t będzie $I(d)$. \mathbf{x}_t nazywamy procesem skointegrowanym $CI(d, b)$ z wektorem kointegrującym β jeśli $\beta' \mathbf{x}_t$ jest $I(d - b)$.*

13.6. Warunki stabilności i stacjonarności dla modeli VAR

Z rozważań zawartych w podrozdziale 17.2 dotyczącym własności operatora opóźnień wynika, że jeśli wielomian charakterystyczny jest odwracalny, to

$$\mathbf{A}(L)^{-1} = \sum_{i=0}^{\infty} \Psi_i L^i$$

Oznacza to, że proces (13.2) można przekształcić do procesu

$$\mathbf{x}_t = \sum_{i=0}^{\infty} \Psi_i (\epsilon_{t-i} + \Phi \mathbf{D}_{t-1})$$

Wartość oczekiwana

$$\mathbf{x}_t = \mathbb{E} \left[\sum_{i=0}^{\infty} \Psi_i (\epsilon_{t-i} + \Phi \mathbf{D}_{t-1}) \right] = \sum_{i=0}^{\infty} \Psi_i \Phi \mathbf{D}_{t-1}$$

w rezultacie

$$\mathbf{x}_t - \mathbb{E}(\mathbf{x}_t) = \sum_{i=0}^{\infty} \Psi_i \epsilon_{t-i}$$

Szereg

$$\sum_{i=0}^{\infty} \Psi_i = \sum_{i=0}^{\infty} \mathbf{J}' \mathbb{A}^i \mathbf{J}$$

jest zbieżny jeśli $\mathbf{A}(L)$ jest odwracalny. W rezultacie mamy następujący wniosek:

Wniosek 13.10 *Jeśli wielomian operatora opóźnień procesu (13.2) jest odwracalny, to $\mathbf{x}_t \sim I(1)$*

13.7. Kointegracja a modele VAR

Proces (13.2) można za pomocą transformacji kointegrującej przekształcić (np. Lütkepohl 1991) do następującej postaci:

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{x}_{t-i} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\epsilon}_t, \quad t = 1 \dots T \quad (13.4)$$

gdzie $\mathbf{\Pi} = \sum_{i=1}^k \mathbf{\Pi}_i - \mathbf{I}$ i $\mathbf{\Gamma}_i = -\sum_{j=i+1}^k \mathbf{\Pi}_j$. Wielomian charakterystyczny tego procesu jest danym wzorem:

$$\mathbf{A}(z) = (1-z)\mathbf{I} - \mathbf{\Pi}z - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i (1-z)z^i \quad (13.5)$$

Zauważmy, że $|\mathbf{A}(1)| = 0$ może zajść tylko wtedy, gdy jeśli $\mathbf{\Pi}$ jest osobliwa. W takim przypadku $\mathbf{\Pi}$ można rozbić na iloczyn $\mathbf{\Pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$, gdzie $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ są macierzami $p \times r$ o pełnym rzędzie kolumnowym, gdzie $r < p$. Twierdzenie Grangera (1983) o reprezentacji podaje warunki, dla których proces (13.4) jest $I(1)$ oraz umożliwia przedstawienie tego procesu za pomocą wielowymiarowego procesu MA (moving average) i umożliwia interpretację macierzy $\boldsymbol{\beta}$ jako macierzy wektorów kointegrujących.

Twierdzenie 13.11 ((1995a)) *Jeśli $|\mathbf{A}(z)| = 0$ dla $|z| \geq 1$, istnieją $\boldsymbol{\alpha}$ i $\boldsymbol{\beta}$ o wymiarach $p \times r$ i $p \times r$ oraz rzędzie r , takie że*

$$\mathbf{\Pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$$

Koniecznym i dostatecznym warunkiem, by $\Delta \mathbf{x}_t - \mathbf{E}[\Delta \mathbf{x}_t]$ i $\boldsymbol{\beta}' \mathbf{x}_t - \mathbf{E}(\boldsymbol{\beta}' \mathbf{x}_t)$ były rzędu $I(0)$ jest, że

$$|\boldsymbol{\alpha}'_{\perp} \mathbf{\Gamma} \boldsymbol{\beta}_{\perp}| \neq 0, \quad (13.6)$$

gdzie $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$. W tym przypadku \mathbf{x}_t można zapisać w następującej formie:

$$\mathbf{x}_t = \mathbf{C} \sum_{i=1}^t (\boldsymbol{\epsilon}_t + \Phi \mathbf{D}_t) + \mathbf{C}_1(L) (\boldsymbol{\epsilon}_t + \Phi \mathbf{D}_t) + \mathbf{A}_0 \quad (13.7)$$

gdzie $\mathbf{C} = \beta_{\perp} (\boldsymbol{\alpha}'_{\perp} \Gamma \beta_{\perp})^{-1} \boldsymbol{\alpha}'_{\perp}$ a $\mathbf{C}_1(z)$ ma pierwiastki poza kołem jednostkowym i $\beta' \mathbf{A}_0 = 0$.

Z powyższego twierdzenia wynika, że

$$\beta' \mathbf{x}_t = \beta' \mathbf{C}_1(L) (\boldsymbol{\epsilon}_t + \Phi \mathbf{D}_t) = \beta' \sum_{i=1}^{\infty} \mathbf{C}_{1i} (\boldsymbol{\epsilon}_t + \Phi \mathbf{D}_t)$$

jest $I(0)$, ponieważ $\mathbf{C}_1(z)$ jest zbieżne. Wektory stanowiące kolumny macierzy β nazywamy wektorami kointegrującymi procesu \mathbf{x}_t . Za początkowe wartości procesu przyjmuje się takie zmienne losowe, dla których dystrybuanta $\beta' \mathbf{x}_t$ jest stała w czasie dla $t = 1, \dots, T$.

Twierdzenie Grangera umożliwia dekompozycję zaburzeń losowych na te, które mają wpływ permanentny na proces i te, które mają wpływ przejściowy. Zauważmy, że z reprezentacji (13.7) wynika, że trwały wpływ na \mathbf{x}_t ma $\boldsymbol{\alpha}'_{\perp} \sum_{i=1}^t \boldsymbol{\epsilon}_t$, zaś $\boldsymbol{\alpha}' \sum_{i=1}^t \boldsymbol{\epsilon}_t$ ma jedynie przejściowy wpływ. Element

$\boldsymbol{\alpha}'_{\perp} \sum_{i=1}^t \boldsymbol{\epsilon}_t$ związany jest z $p - r$ niezależnymi trendami stochastycznymi występującymi w procesie.

Wpływ elementów deterministycznych na zachowanie procesu zależy od tego, czy wektor parametrów z nimi związanych jest równoległy do $\boldsymbol{\alpha}$. Przykładowo, jeśli \mathbf{D}_t zawiera jedynie stałą, wtedy w \mathbf{x}_t występować będzie trend liniowy, ponieważ $\boldsymbol{\alpha}'_{\perp} \sum_{i=1}^t \phi = \boldsymbol{\alpha}_{\perp} \phi t$. Trend taki nie wystąpi jednak jeśli $\boldsymbol{\alpha}'_{\perp} \phi = 0$, czyli dla szczególnego przypadku, kiedy ϕ jest równoległy do $\boldsymbol{\alpha}$. Bezwarunkowa wartość oczekiwana \mathbf{x}_t jest równa

$$E(\mathbf{x}_t) = \mathbf{C} \sum_{i=1}^t \Phi \mathbf{D}_t + \mathbf{C}_1(L) \Phi \mathbf{D}_t + \mathbf{A}_0$$

Z kolei wariancja $\mathbf{x}_t - E(\mathbf{x}_t)$ jest liniową funkcją czasu:

$$\text{Var}(\mathbf{x}_t) = t\mathbf{C}\boldsymbol{\Omega}\mathbf{C}' + \mathbf{C}_1(1)\boldsymbol{\Omega}\mathbf{C}_1(1)$$

Wyrażenie $\beta' \mathbf{x}_t$ interpretujemy jako mechanizm korekty błędu (Error Correction Mechanism *ECM*). Z kolei macierz α nazywamy macierzą współczynników korygujących.

13.8. Własności estymatorów dla modeli *VAR*

• MNK dla zmiennych stacjonarnych

Jeśli \mathbf{x}_t jest słabo stacjonarne, to z definicji wiemy, że macierz wariancji kowariancji jest skończona i stała w czasie a więc $\sup_t \text{Var}(\mathbf{x}_t) < \infty$. Z kolei

$$\begin{aligned} \text{Var}(\mathbf{X}'\mathbf{X}) &= \text{Var}\left(\sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t'\right) = \text{Var}\left(\sum_{t=1}^T \mathbf{C}(L) \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mathbf{C}'(L)\right) \\ &= \sum_{t=1}^T [\mathbf{C}(L) \text{Var}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \mathbf{C}'(L)] = T\mathbf{C}(1) \text{Var}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \mathbf{C}'(1) \end{aligned}$$

Jeśli $\|\text{Var}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')\| < \infty$, to

$$\begin{aligned} \text{Var}\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right) &= T^{-1}\mathbf{C}(1) \text{Var}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \mathbf{C}'(1) \longrightarrow 0 \\ \frac{\mathbf{X}'\mathbf{X}}{T} &\xrightarrow{m.s.} E\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right) = E[\mathbf{C}(L) \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mathbf{C}'(L)] = \mathbf{C}(1)\boldsymbol{\Omega}\mathbf{C}'(1) \end{aligned}$$

Z drugiej strony jeśli $E(\mathbf{x}_t \boldsymbol{\varepsilon}_t) = 0$, to

$$E(\mathbf{x}_t' \boldsymbol{\varepsilon}_t)^2 = \sigma^2 \text{Var}(\mathbf{x}_t)$$

a więc $\mathbf{x}_t \varepsilon_t = O_p \left(T^{-\frac{1}{2}} \right)$

$$\mathbf{X}'\varepsilon = \sum_{t=1}^T \mathbf{x}_t' \varepsilon_t = \sum_{t=1}^T O_p \left(T^{-\frac{1}{2}} \right) = O_p \left(T^{\frac{1}{2}} \right)$$

13.9. Warunkowe skointegrowane procesy VAR

Parametry procesu VAR są trudne do estymacji ze względu na ilość parametrów, które musimy wyestymować. Jeśli w procesie (13.4) występuje stała i trend liniowy, to jego ilość parametrów jest równa

$$pr + (k-1)p^2 + p + p + \frac{p(p+1)}{2} = \left(k + \frac{1}{2}\right)p^2 + \left(r + \frac{5}{2}\right)p \quad (13.8)$$

Całkowita liczba obserwacji dla wektorowego szeregu czasowego o długości T i wymiarze p wynosi Tp . Zwiększając liczbę analizowanych zmiennych powiększamy wprost proporcjonalnie liczbę obserwacji, ale tracimy stopnie swobody proporcjonalnie do kwadratu liczby zmiennych. Ograniczenie ilości estymowanych parametrów można uzyskać estymując zamiast procesu bezwarunkowego proces warunkowy względem części elementów wektora \mathbf{X}_t . Estymatory uzyskane na podstawie dystrybuanty procesu warunkowego mogą być efektywne jedynie wtedy, gdy zmienne generowane przez proces brzegowy są słabo egzogeniczne względem parametrów procesu warunkowego. Hendry (1995) podaje następującą definicję słabej egzogeniczności:

Definicja 13.12 *W procesie losowym generowanym przez PGD (Proces Generujący Dane) $D_{\mathbf{x}}(\mathbf{x}_t | \mathbf{x}_{t-1}, \boldsymbol{\theta})$, $\Pr(\mathbf{x}_t | \mathbf{x}_{t-1}, \boldsymbol{\theta})$, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, $\mathbf{X}_{t-1} = (\mathbf{x}_0, \dots, \mathbf{x}_{t-1})$ o p_1 wymiarowym wektorze zmiennych \mathbf{x}_2 mówimy, że jest egzogeniczny względem parametru $\boldsymbol{\theta}_1$ jeśli*

1. *PGD tego procesu można zdekomponować w taki sposób, że*

$$D_{\mathbf{x}}(\mathbf{x}_t | \mathbf{X}_{t-1}, \boldsymbol{\theta}) = D_{\mathbf{x}_1 | \mathbf{x}_2}(\mathbf{x}_{1t} | \mathbf{X}_{t-1}, \mathbf{x}_{2t}, \boldsymbol{\theta}_1) D_{\mathbf{x}_2}(\mathbf{x}_{2t} | \mathbf{X}_{t-1}, \boldsymbol{\theta}_2)$$

gdzie $\mathbf{X}_t = [\mathbf{x}_{t-1}, \dots, \mathbf{x}_0]$

2. Zbiory dopuszczalnych wartości parametrów $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$ są swobodnie zmienne, to jest $\Theta_1 \cup \Theta_2 = \Theta_1 \times \Theta_2$.

Ograniczenia jakie muszą spełniać parametry procesu, by p_1 zmiennych było słabo ezogenicznych względem β przeanalizowali, ?. Jeśli macierz wariancji zapiszemy w zgodzie z podziałem \mathbf{x}_t na $(\mathbf{x}_{1t}, \mathbf{x}_{2t})$ jako macierz blokową $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$ i pomnożymy równanie (13.4) przez macierz $\begin{bmatrix} I & -\omega \\ 0 & I \end{bmatrix}$, gdzie $\omega = \Omega_{12}\Omega_{22}^{-1}$, to otrzymamy proces warunkowy

$$\Delta \mathbf{x}_{1t} = \omega \Delta \mathbf{x}_{2t} + (\alpha_1 - \omega \alpha_2) \beta' \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} (\Gamma_{1i} - \omega \Gamma_{2i}) \Delta \mathbf{x}_{t-i} + (\Phi_1 - \omega \Phi_2) \mathbf{D}_t + \varepsilon_{1t}, \quad (13.9)$$

i proces brzegowy

$$\Delta \mathbf{x}_{2t} = \alpha_2 \beta' \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} \Gamma_{2i} \Delta \mathbf{x}_{t-i} + \Phi_2 \mathbf{D}_t + \varepsilon_{2t}, \quad (13.10)$$

gdzie $\varepsilon_{1t} = \epsilon_{1t} - \Omega_{12}\Omega_{22}^{-1}\epsilon_{2t}$ i $\varepsilon_{2t} = \epsilon_{2t}$ są rzeczywiście niezależne. Warunkiem koniecznym do tego, by brzegowe PGD procesu \mathbf{x}_{2t} nie zależało od β jest warunek, by $\alpha_2 = 0$. W takim przypadku PGD (13.4) można sformułować jako

$$\begin{aligned} D_{\mathbf{x}}(\mathbf{x}_t | \mathbf{x}_{t-1}, \Gamma_1, \dots, \Gamma_{k-1}, \alpha, \beta, \Phi, \Omega) \\ = D_{\mathbf{x}_1 | \mathbf{x}_2}(\mathbf{x}_1 | \mathbf{x}_{t-1}, \mathbf{x}_{2t}, \Gamma_{1,1}^*, \dots, \Gamma_{1,k-1}^*, \alpha_1, \beta, \Phi_1^*, \Omega_{11}^*) \times \\ D_{\mathbf{x}_2}(\mathbf{x}_{2t} | \mathbf{x}_{t-1}, \Gamma_{2,1}, \dots, \Gamma_{2,k-1}, \Phi_2, \Omega_{22}), \end{aligned}$$

gdzie $\Gamma_{1i}^* = \Gamma_{1i} - \omega \Gamma_{2i}$, $\Phi_1^* = \Phi_1 - \omega \Phi_2$, $\Omega_{11}^* = \Omega_{11} - \omega \Omega_{21}$.

Johansen (1992a) pokazał, że jeśli ograniczenie narzucone na α jest prawdziwe i rząd β jest znany, to można efektywnie estymować β przy użyciu procesu warunkowego (13.10). Ograniczenie $\alpha_2 = 0$ oznacza, że wyrażenie związane z korektą błędu nie wchodzi do procesu generującego x_{2t} ale wektor x_{2t} może być w dalszym ciągu skointegrowany. Do efektywnej estymacji β nie jest nam potrzebna silna egzogeniczność x_{2t} , która wymagałaby braku przyczynowości w sensie Grangera. Ważne jest także to, że słaba egzogeniczność x_{2t} względem parametru β implikuje słabą egzogeniczności tej zmiennej w stosunku do α_1 ale nie względem pozostałych parametrów. Z tego powodu uzyskane z procesu warunkowego (13.10) estymatory Γ_i , Φ_1 oraz Ω_{11} są nieefektywne i obciążone.

Założenie o niewystępowaniu w procesie brzegowym elementów związanych z korektą błędów może być oczywiście fałszywe. W związku z tym nasuwa się pytanie, czy możliwe jest testowanie słabej egzogeniczności w procesie *VECM*. ? zaproponowali następującą procedurę estymacji w przypadku wektorowych *ECM*. Odpowiednio zmodyfikowaną wersję tej procedury będziemy stosować przy okazji estymacji przykładu empirycznego. Dwa ostatnie¹ jej punkty można traktować jako test diagnostyczny słabej egzogeniczności x_{2t} względem parametru β :

1. Stwórz proces warunkowy (13.9) taki, by reszty z tego procesu miały rozkład *NID*
2. Stwórz proces brzegowy (13.10) taki, by reszty z tego procesu miały rozkład *NID*
3. Ustal na podstawie testów jaki jest rząd macierzy β w procesie warunkowym (13.9)
4. Przetestuj założenia dotyczące formy trendów deterministycznych
5. Przetestuj słabą egzogeniczność x_{2t} sprawdzając, czy wyestymowane dla procesu warunkowego (13.9) zależności długookresowe są rzeczywiście nieistotne w procesie brzegowym
6. Przetestuj rząd macierzy kointegrującej dla procesu brzegowego x_{2t}

¹Drugi z tych punktów zaproponowali ?

13.10. Problemy związane z identyfikacją $VECM$

W rozdziale tym rozpatrujemy problem identyfikacji. W $VECM$ można rozdzielić ograniczenia służące identyfikacji parametrów długiego i krótkookresowych. W mechanizmie korekty błędów (13.4) sposób generowania obserwacji nie zostanie zmieniony, jeśli macierze α i β zastąpimy macierzami $\alpha = \alpha A'$ i $\beta = \beta A^{-1}$ gdzie A jest dowolną nieosobliwą macierzą $r \times r$. Do identyfikacji parametrów długookresowych zawartych w macierzy β konieczne jest więc nałożenie na nią r^2 ograniczeń. Do uzyskania identyfikacji parametrów krótkookresowych zwykle wykorzystuje się ograniczenia implikowane przez brak występowania po prawej stronie równania (13.4) zmiennych równoczesnych ze zmiennymi po lewej stronie. Ograniczenia są całkowicie arbitralne, ale zauważmy, że do identyfikacji zależności krótkookresowych używamy innych ograniczeń, niż do identyfikacji relacji długookresowych. Jeśli zainteresowani jesteśmy jedynie zależnościami długookresowymi, wtedy jedynie będzie miało dla nas znaczenie jakie ograniczenia zastosujemy do identyfikacji parametrów krótkookresowych.

Rząd α i β nie zależy od sposobu identyfikacji tych macierzy. Jak wiemy z Twierdzenia Grangera rzędy tych macierzy związane są z ilością relacji długookresowych (kointegrujących). Tą cechę macierzy β można przetestować stosując w procesie estymacji dowolne, arbitralne ograniczenia. Często używamy układu ograniczeń służących do identyfikacji zaproponowanych przez Johansena (1988):

$$\beta' S_{11} \beta = I, \quad (13.11)$$

gdzie S_{11} jest macierzą momentów empirycznych wektora x_t .

Identyfikacja ta, choć technicznie wygodna, nie da się z reguły przenieść na ograniczenia interpretowalne w ramach konkretnego modelu. Jednak estymatory uzyskane dla jednego zbioru ograniczeń dokładnie identyfikujących parametry mogą być łatwo przekształcone w taki sposób, by spełniały inny zbiór ograniczeń dokładnie identyfikujących parametry. Mając estymatory wyprowadzone za pomocą powyższej *technicznej* identyfikacji możemy łatwo uzyskać estymatory spełniające szerszą klasę ograniczeń. Johansen (1991) zaproponował, by liniowe ograniczenia dokładnie identyfikujące

β miały formę równania $c'\widehat{\beta}_c = I$, gdzie c jest dowolną $p \times r$ macierzą o pełnym rzędzie. Po to, by narzucić te ograniczenia na dowolny estymator β policzony dla dokładnie ale inaczej zidentyfikowanego układu relacji długookresowych, wystarczy zastosować wzór

$$\beta_c = \beta (c'\beta)^{-1} \quad (13.12)$$

i $\alpha_c = \alpha\beta'c$. Szczególnym przypadkiem identyfikacji za pomocą warunku (13.12) jest przyjęcie takiej macierzy c , że

$$\beta'_c = [I, \beta^{*'}] \quad (13.13)$$

13.11. Metody estymacji i wnioskowania dla procesów VAR

Parametry procesu VAR można estymować za pomocą MNK² zastosowaną do indywidualnych równań (Lütkepohl 1991). Metoda ta daje zgodne estymatory zarówno dla przypadku, kiedy zmienne są $I(0)$, jak i przypadku gdy zmienne są $I(1)$. Problemy związane z estymacją VAR ze zmiennymi $I(1)$ wiążą się więc z wnioskowaniem statystycznym a nie z trudnością z uzyskaniem zgodnych estymatorów parametrów procesu (13.2). Problemy z wnioskowaniem wynikają z różnych stóp zbieżności estymatorów parametrów związanych ze zmiennymi o rzędzie integracji $I(0)$ i $I(1)$. Jeśli w procesie nie występuje kointegracja, estymatory parametrów związanych ze zmiennymi $I(0)$ zbiegają z szybkością $T^{-\frac{1}{2}}$ a parametry przy zmiennych $I(1)$ z szybkością T^{-1} . Sytuacja ulega dalszej komplikacji, jeśli pewne kombinacje zmiennych są $I(1)$ a inne $I(0)$ i wektor estymowanych parametrów zbiega w pewnych kierunkach z szybkością $T^{-\frac{1}{2}}$ a w innych z szybkością T^{-1} . Ustalanie szybkości zbieżności estymatorów ma kluczowe znaczenie dla znalezienia takich formuł dla statystyk testowych, które mają niezdegenerowane rozkłady asymptotyczne.

Pierwszy zwrócił na ten problem uwagę Fuller (1976), który udowodnił, że zbiór koniecznych założeń, na których opiera się dowód, że statystyka t ma rozkład t -Studenta, nie jest spełniony

²MNK jest dla tego przypadku równoważny estymatorowi MNW.

dla zmiennych niestacjonarnych i że dla takich zmiennych test ten ma zupełnie inny rozkład prawdopodobieństwa. W kolejnych pracach pokazano, że rząd integracji procesu można testować za pomocą testu DF a rozkład statystyki t w tym teście jest dany funkcjami ruchów Browna (Phillips 1987, Phillips 1988) i zależy od tego, jak kształtuje się w zmiennej $I(0)$ trend deterministyczny (Dickey i Fuller 1979, Dickey i Fuller 1981). W szerszym kontekście pokazano, że twierdzenia graniczne dla zmiennych tworzących proces niestacjonarny wyglądają zupełnie inaczej niż dla zmiennych stacjonarnych. (np. Billingsley 1979, Davidson 1994).

Dalszy rozwój badań nad testami na istnienie pierwiastka jednostkowego poszedł w kierunku opracowywania testów asymptotycznie zgodnych przy słabszych założeniach niż w klasycznym teście DF . Said i Dickey (1984) pokazali, że dodając w teście DF opóźnione różnice uzyskujemy test zgodny dla procesów $ARIMA$ (test ADF). Problemem przy zastosowaniach testu ADF pozostaje dobór ilości opóźnień (np. ?). Jednym z proponowanych rozwiązań (Phillips i Perron 1988) jest zastosowanie nieparametrycznej poprawki zamiast dodawania opóźnionych różnic.

Drugim nurtem badań były prace nad modyfikacją samego kształtu hipotezy zerowej i alternatywnej. Sargan i Bhargava (1983) pokazali, że możliwe jest takie przeprowadzanie testu na istnienie pierwiastka jednostkowego, że zarówno przy hipotezie zerowej jak i przy alternatywnej zakładamy jedynie istnienie trendu liniowego. Z drugiej strony ? udowodnili, że możliwe jest nie tylko testowanie hipotezy zerowej o niestacjonarności ale także hipotezy zerowej o stacjonarności (test $KPSS$).

Na ile literatura na temat testowania rzędu integracji jest przydatna przy estymacji parametrów procesu VAR ? Proces (13.4) jest tak sformułowany, że elementy x_t mogą być zarówno stacjonarne jak i niestacjonarne. Stacjonarność elementów x_t implikuje jedynie pewne ograniczenia dotyczące macierzy β . Integracje zmiennych można testować poprzez testowanie tych ograniczeń. Z tego punktu widzenia wydaje się, że testowanie integracji zmiennych przed wprowadzeniem ich do procesów VAR jest nieuzasadnione, ponieważ testy wykorzystujące pełen proces powinny mieć wyższą moc. Testowanie takie może mieć jednak sens jeśli estymujemy proces warunkowy, ponieważ wt-

edy testowanie rzędu integracji zmiennych objaśniających można interpretować jako test na słabą egzogeniczność tych zmiennych względem parametrów długookresowych.

Historycznie pierwszymi testami kointegracji były testy opracowane dla procesów jednorównaniowych. Można je podzielić na te, które opierają się na uogólnionych testach na istnienie pierwiastka jednostkowego i te, które opierają się na testowaniu *ECM*. Testy oparte na uogólnionych testach pierwiastka jednostkowego wychodzą z obserwacji, że testowanie kointegracji dla procesu skalarnego można sprowadzić do testowania stacjonarności reszt z regresji x_t na y_t , gdzie x_t jest wektorem zmiennych słabo egzogenicznych względem β . Tablice wartości krytycznych opracowano dla testu Dickey-Fullera (Engle i Yoo 1987b, Charemza i Deadman 1997), testu Perrona-Phillipsa (Phillips i Ouliaris 1990b) oraz testu *KPSS* (Shin 1994).

Jednorównaniowe testy oparte na estymacji *ECM* weryfikują istnienie kointegracji na podstawie wyników estymacji jednego z równań (13.4). (?) pokazali, że zmienna x_{it} jest słabo egzogeniczna przy estymacji parametru β na podstawie i -tego równania procesu (13.4) wtedy i tylko wtedy gdy i -te równanie jest jedynym, w którym występuje kointegracja. W tym przypadku test *LR* na kointegrację sprowadza się do testowania istotności współczynnika α przy wektorze kointegrującym β w jednorównaniowym procesie *ECM*. Testy oparte na podobnej zasadzie zaproponowali ?.

Dla procesów wielowymiarowych, odpowiednikiem testowania kointegracji jest testowanie liczby wektorów kointegrujących stanowiących kolumny macierzy β . Wszystkie metody estymacji wektorów kointegrujących opierają się na obserwacji, że własności βx_t są różne od własności x_t i $\beta'_{\perp} x_t$ ponieważ $\beta x_t \sim I(0)$ a x_t i $\beta'_{\perp} x_t$ są $I(1)$. Omówimy tu trzy najpopularniejsze metody, to jest metodę czynników głównych Stocka i Watsona (Stock i Watson 1988) oraz dwie metody oparte na korelacjach kanonicznych: metodę Johansena (Johansen 1988, Johansen i Juselius 1990, Johansen 1991, Johansen 1992b, Johansen 1994b) i metodę Boxa i Tiao (Box i Tiao 1977, ?).

Metoda czynników głównych Stocka i Watsona opiera się na fakcie, że $\beta' x_t$ jest stacjonarne i jego wariancja jest stała, podczas gdy $\beta'_{\perp} x_t$ jest niestacjonarne i jego wariancja rośnie proporcjonalnie do T . Można więc znaleźć estymator macierzy β_{\perp} szukając takich znormalizowanych kombi-

nacji liniowych elementów x_t , które mają największą wariancję. Do testowania rzędu macierzy β Harris (1997) proponuje użyć uogólnionego testu *KPSS*.

Metoda Boxa i Tiao *LCCA* (*Level Canonical Correlation Analysis*) opiera się na tym, że kowariancja empiryczna między x_t a βx_{t-1} jest asymptotycznie równa zeru jeśli tylko wszystkie elementy x_t są $I(1)$. Za wektory kointegrujące przyjmujemy więc takie znormalizowane kombinacje liniowe zmiennych niestacjonarnych, które są najmniej skorelowane z oryginalnymi zmiennymi. Rozkład testu na ilość wektorów kointegrujących wykorzystujący tę metodę estymacji można znaleźć w artykule Bewleya i Yanga (1995, 1996).

Inną metodą testowania rzędu macierzy kointegrującej jest metoda Johansena. Można ją zaliczyć do metod opartych na analizie korelacji kanonicznych, ponieważ polega na znajdowaniu takich kombinacji liniowych $\beta' x_{t-1}$, które są najsilniej skorelowane z Δx_t . Johansen pokazał (np. Johansen 1995a), że jego metoda estymacji jest *MNW*. Bliższe szczegóły dowodu zaprezentuję w dalszej części pracy. Wartości krytyczne dla procesu bez elementów deterministycznych podał Johansen (1988), tablice zawierające także wartości krytyczne dla procesów z elementami deterministycznymi opublikował Osterwald-Lenum (1992), najnowsze tablice podał Johansen (1995a). Wadą testu Johansena jest, że jego rozkład zależy od formy trendów deterministycznych. W efekcie, gdy testujemy hipotezę o rzędzie kointegracji, robimy to warunkowo względem założonej formy trendów deterministycznych.

Modyfikacja testu Johansena zaproponowana przez Saikkonen i Lütkepohl (2000) jest w pewnym stopniu pozbawiona tej wady. Rozpatrują oni klasę zmodyfikowanych testów *LM* i *LR* zastosowanych do zmiennych z usuniętymi trendami. Uzyskane testy mają nieco inny rozkład niż standardowy test Johansena i rozkład ten jest niezależny od założonej struktury trendów deterministycznych. Dla procesu jedynie ze stałą Saikkonen i Lütkepohl (2000) pokazują metodą Monte Carlo, że test ten ma, dla lokalnej alternatywy, wyższą moc od klasycznego testu *LR*.

Wariantem metody Johansena jest metoda Kleibergena (1996). Kleibergen zauważył, że jeśli zidentyfikujemy β za pomocą sztucznej identyfikacji (13.13), to względnie łatwo jest uzyskać esty-

mator α . Estymator α można z kolei wykorzystać do uzyskania estymatorów β i $\tilde{\Omega}$. Metodę tę można interpretować jako odmianę Dwustopniowej Metody Najmniejszych Kwadratów (2MNK)³. Estymatory uzyskane w ten sposób są równie efektywne i mają te same rozkłady co estymatory MNW, dla przypadku kiedy znane jest Ω . Kiedy Ω jest nieznane, uzyskany estymator jest zgodny ale nieefektywny. Estymator efektywny można uzyskać iterując 2MNK z wykorzystaniem macierzy $\tilde{\Omega}$ zamiast macierzy Ω . Test na rząd macierzy kointegrującej można uzyskać analizując rozkład minimalizowanej funkcji celu metody. Kleibergen (1996) pokazuje, że w pewnych przypadkach, takich jak heteroskedastyczność, czy załamania w macierzach α i β uzyskane w dzięki tej metodzie estymatory mają znacznie prostszą postać, niż estymatory MNW.

Który ze sposobów estymacji relacji kointegrujących i testowania ich liczby jest najlepszy? Wszystkie omawiane estymatory i testy są asymptotycznie zgodne. Wszystkie estymatory poza dwuetapowym Engla-Grangera i estymatorem metody czynników głównych są także asymptotycznie efektywne. Wady tej nie ma opracowana przez Phillipsa i Hansena (1990a) zmodyfikowana wersja procedury Engla-Grangera (*Fully Modified Ordinary Least Squares*), i zmodyfikowany estymator metody czynników głównych Harrisa (1997), które za pomocą nieparametrycznych poprawek usuwają wpływ parametrów zakłócających (*nuisance parameters*) na asymptotyczny rozkład testów. Kamieniem probierczym jakości estymatorów powinny być więc ich własności w małych próbach. Przegląd artykułów porównujących te własności za pomocą metody Monte Carlo znajduje się w książce Maddali i Kima (1998). W przypadku estymacji parametrów warunkowych procesów skalarnych wydaje się, że metody estymacji oparte na jednorównaniowym ECM wykazują lepsze własności niż dwuetapowa metoda Engla-Grangera.

W przypadku estymacji procesów wielowymiarowych ustalono, że metoda Johansena daje estymatory stosunkowo mało obciążone ale o wysokiej wariancji. Metoda Boxa-Tiao daje estymatory o niższej wariancji ale większym obciążeniu. Metoda Johansena jest szybciej zbieżna niż metoda Boxa-Tiao, pod warunkiem, że szybkość dostosowania systemu do równowagi długookresowej jest

³Inną formę estymatora 2MNK można znaleźć w artykule Hsiao (1997).

szybka, korelacja błędów między równaniami wysoka a próba stosunkowo duża.

Z porównania testów wielorównaniowych, dokonanego przez Lütkepohla i Saikkonena (1998) wynika, że dla małych prób, testy kointegracji typu LM , LR i W mają lepsze własności niż pozostałe testy, takie jak test $LCCA$ i Boxa-Tiao. Test W ma jednak w małych próbach gorsze własności niż testy LM i LR . W przypadku testu LM i LR Lütkepohl i Saikkonen porównali test Johansena z różnymi wariantami testów LM i LR różniącymi się między sobą sposobem usuwania trendów ze zmiennych. Na podstawie eksperymentów Monte Carlo ustalili oni, że istotne dla mocy testu jest właściwe zdefiniowanie elementów deterministycznych. Dla przypadku, kiedy nie pewności, jak w rzeczywistości kształtują się trendy w procesie, testy oparte na statystyce LR wykazują lepsze własności niż testy LM . Jeśli w danych nie występuje trend liniowy, najlepsze własności wykazuje test LR zaproponowany przez Saikkonen i Lütkepohl (2000). W przypadku, gdy jeden z pierwiastków procesu jest bliski jedności, testy oparte na zmiennych z usuniętymi trendami wydają się mieć lepsze własności niż klasyczny test LR . W przypadku testowania rzędu kointegracji w procesach o większej ilości wektorów kointegrujących (Lütkepohl i Saikkonen 1998) autorzy nie otrzymali jednoznacznych wniosków. Test LR Johansena miał najwyższą moc w przypadku testowania hipotezy $H_0 : r = 0$ przeciw hipotezie $H_1 : r > 0$, miał jednak niższą moc niż testy oparte na zmiennych z usuniętymi trendami w przypadku testowania istnienia pozostałych wektorów kointegrujących. Ponieważ procedura ustalania rzędu macierzy kointegrującej jest sekwencyjna, więc otrzymane wnioski na temat efektywności testów są niejednoznaczne.

Poza testowaniem ograniczeń związanych z badaniem kointegracji zmiennych często testujemy inne ograniczenia nałożone na współczynniki procesu. Testy takie są ułatwione przez fakt, że asymptotycznie $\hat{\beta}$ ma rozkład jest mieszany normalny⁴ a dla pozostałych współczynników normalny. W rezultacie asymptotyczne rozkłady testów dla ograniczeń nakładanych na β , jak i na pozostałe współczynniki, są standardowe (np. w przypadku testu *Wald* rozkład χ^2). Trudniejsze jest testowanie hipotez zawierających zarówno parametry krótko jak i długookresowe ze względu

⁴Wyjątkami są tu estymatory β dwustopniowej metody Engla-Grangera i estymator metody czynników głównych.

na różne stopy zbieżności estymatorów dla tych dwóch rodzajów parametrów. W tym przypadku można jednak skorzystać z asymptotycznej niezależności tych estymatorów.

Przy realistycznych założeniach wszystkie omówione metody estymacji oraz testy na integracje i kointegrację mają wyprowadzone rozkłady estymatorów i testów jedynie dla przypadków asymptotycznych. Publikowane rozkłady testów w małych próbach (np. Charemza i Deadman 1997) są prawdziwe tylko wtedy, gdy w procesie nie występuje autokorelacja reszt. Trudno je uznać za szczególnie użyteczne, skoro autokorelacja błędów losowych jest regułą w przypadku rzeczywistych szeregów czasowych. Atrakcyjną alternatywą są tutaj metody polegające na wtórnym próbkowaniu (*resampling*) takie jak *bootstrap* (np. Li i Maddala 1997).

Inną rozwiązaniem jest zastosowanie estymatorów Bayesowskich. Nie będziemy ich tu szczegółowo omawiać a jedynie zwrócimy uwagę na główne problemy dyskutowane w literaturze. Najbardziej kontrowersyjnym problemem przy estymacji parametrów procesów niestacjonarnych tą metodą jest wybór rozkładu *a priori*. W przypadku tego rodzaju estymacji przyjęcie rozkładu jednostajnego jako rozkładu *a priori* dla parametru związanego pierwiastkami wielomianu opóźnień nie jest satysfakcjonujące, ponieważ rozkład ten nie jest neutralny względem różnych wartości tego pierwiastka. Phillips (1991) zaproponował rozkład *a priori* Jeffreya jako właściwą reprezentację braku informacji na temat cech szeregu czasowego. Niestety pokazano, że wyniki estymacji przy użyciu tego rozkładu są wrażliwe na sposób zdefiniowania elementów deterministycznych w procesie (Schotman i Dijk 1991) i sposób potraktowania wartości początkowych (Lubrano 1995). Teoretycznie największą zaletą podejścia Bayesowskiego jest możliwość policzenia rozkładów estymatorów i testów w małych próbach. Jak na razie, uzyskanie takich estymatorów dla parametrów procesów niestacjonarnych możliwe jest jedynie przy użyciu skomplikowanych metod numerycznych i przyjęciu pewnych ryzykownych założeń na temat rozkładu *a priori*. Inna sytuacja ma miejsce, jeśli skłonni jesteśmy przyjąć, że mamy pewną wiedzę *a priori* na temat zachowania ekonomicznych szeregów czasowych a w tym także informacje na rzędu ich integracji. W takim przypadku problem doboru neutralnego rozkładu *a priori* automatycznie znika. Interesujący przykład tego rodzaju

Bayesowskiego procesu VAR ($BVAR$) wraz z opisem zastosowanego rozkładu *a priori*⁵ można znaleźć w pracy Todda (1990).

⁵Jest to wariant tak zwanych *Minnesota priors* zaproponowanych między innymi przez Doana, Littermana i Simsa (?)