KM 19(1985)
pag 143-147

SINGULIERE COVARIANTIEMATRICES EN SUR-MODELLEN: EEN ANTWOORD AAN WANSBEEK

R. Harkema en P.M.C. de Boer

Econometrisch Instituut
Erasmus Universiteit Rotterdam
Postbus 1738
3000 DR ROTTERDAM

In KM 18 (p. 99-102) reageert Wansbeek op de door ons in KM 16 voorgestelde schattingsprocedure voor SUR-modellen met een optelrestrictie op de storingen. Vanzelfsprekend zijn wij hem erkentelijk voor zijn reactie en met name voor zijn elegante wiskundige analyse. Helaas heeft diezelfde elegante wiskundige analyse Wansbeek naar onze mening verleid tot een onjuiste interpretatie van de parameters van de door ons gehanteerde covariantiematrix en, daarmee samenhangend, tot een vereenvoudiging van onze schattingsprocedure, die onjuist is en bovendien de zin van de door ons voorgestelde inperking van de covariantiematrix ondergraaft.

De door ons voorgestelde specificatie van de covariantiematrix luidt

$$\Omega = D - (\iota_p'd)^{-1}dd' \tag{1}$$

met $D = \mathrm{diag}(d_1,\ldots,d_p)$, $d = D\iota_p$ en $d_1,\ldots,d_p$ een p-tal onbekende parameters. Deze matrix is positief semi-definiet, indien hetzij alle d_i's positief zijn, dan wel ten hoogste èén d_i, zeg d_1, negatief is en tevens geldt $|d_1| > \Sigma_{i=2}^{p} d_i$. Hieronder zullen wij nader ingaan op de volgende drie punten uit het commentaar van Wansbeek: (a) de interpretatie van het geval $d_1 < 0$; (b) de voorgestelde vereenvoudiging van de schattingsprocedure en (c) de interpretatie van de covarianties in de door ons voorgestelde specificatie (1).

(a) Zoals uit de door Wansbeek voorgestelde schattingsprocedure blijkt, tracht hij de d_i's te interpreteren als de varianties van de verschillende deelmodellen en dus acht hij het geval $d_1 < 0$ dan ook een dusdanig vreemde situatie, dat het zinnig is om $d_1 > 0$ van het begin af aan tot onderdeel van de specificatie te maken. Voor de correcte specificatie, i.c., $\omega_{ii} = d_i - d_i^2(\iota_p'd)^{-1}$ $(i = 1,\ldots,p)$, heeft het geval $d_1 < 0$ echter een duidelijke econometrische interpretatie. Immers, stel $d_1 < 0$, $d_i > 0$ $(i = 2,\ldots,p)$ en $|d_1| > \Sigma_{i=2}^p d_i$ (en dus $\iota_p'd < 0$). Uit $d_i > 0$ $(i = 2,\ldots,p)$ volgt

$$(\sum_{i=2}^{p} d_i)^2 > \sum_{i=2}^{p} d_i^2$$

en dus

$$\Sigma_{i=2}^p d_i - (\iota_p'd)^{-1}(\Sigma_{i=2}^p d_i)^2 > \Sigma_{i=2}^p d_i - (\iota_p'd)^{-1} \Sigma_{i=2}^p d_i^2 = \Sigma_{i=2}^p[d_i - (\iota_p'd)^{-1}d_i^2]$$

daar $-(\iota_p'd)^{-1} > 0$. Substitutie van $\Sigma_{i=2}^p d_i = (\iota_p'd) - d_1$ in het linkerlid van de ongelijkheid leidt, na enig herschrijven, tot

$$d_1 - (\iota_p'd)^{-1} d_1^2 > \sum_{i=2}^{p} [d_i - (\iota_p'd)^{-1} d_i^2].$$

Kortom, $d_1 < 0$ impliceert dat de variantie van het eerste deelmodel groter is dan de som van de varianties van alle andere deelmodellen (en vice versa). Alhoewel dit geval, economisch gezien, misschien duidt op een misspecificatie van het model, zouden wij er toch de voorkeur aan geven deze mogelijkheid open te laten, al was het maar als toets op de plausibiliteit van het model.

(b) In de door Wansbeek gebruikte representatie, wordt het SUR-model geschreven als

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad (2)$$

met

$$y_t' = [y_{t1} \ldots y_{tp}] \qquad u_t' = [u_{t1} \ldots u_{tp}]$$

$$X_t = \begin{bmatrix} x_{t1}^{(1)} \ldots x_{tk_1}^{(1)} & 0 \ldots 0 & \ldots 0 \ldots 0 \\ 0 \ldots 0 & x_{t1}^{(2)} \ldots x_{tk_2}^{(2)} & \ldots 0 \ldots 0 \\ \vdots \quad \vdots & \vdots \quad \vdots & \vdots \quad \vdots \\ 0 \ldots 0 & 0 \ldots 0 & \ldots x_{t1}^{(p)} \ldots x_{tk_p}^{(p)} \end{bmatrix}$$

en

$$\beta' = [\beta_{11} \ldots \beta_{1k_1}\ \beta_{21} \ldots \beta_{2k_2} \ldots \beta_{p1} \ldots \beta_{pk_p}].$$

In het bovenstaande model representeert y_{ti} de waarde van de te verklaren variabele in deelmodel i, $x_{tj}^{(i)}$ de waarde van verklarende variabele j in deelmodel i, u_{ti} de storing in deelmodel i, β_{ij} de coefficient van verklarende variabele j in deelmodel i en n het aantal waarnemingen met betrekking tot elk van de deelmodellen. De vectoren van storingen u_t voldoen aan de optelrestrictie, dat wil zeggen $\iota_p' u_t = 0$ $(t = 1,\ldots,n)$ en de covariantiematrix van de vector van storingen $u' = [u_1' \ldots u_n']$ is $I_n \otimes \Omega$, waarbij Ω is gespecificeerd volgens (1).

Vermenigvuldigen we nu het hele stelsel (2) met $I_n \otimes E_p$, waarbij $E_p = I_p - p^{-1}\iota_p\iota_p'$, dan blijkt dat de matrix van verklarende variabelen voor het i-de deelmodel na de transformatie geschreven kan worden als

$$Z_i = [-p^{-1}X^{(1)} \ldots -p^{-1}X^{(i-1)}\ (1-p^{-1})\ X^{(i)}\ -p^{-1}X^{(i+1)} \ldots -p^{-1}X^{(p)}]$$

waarbij

$$X^{(j)} = \begin{bmatrix} x_{11}^{(j)} & \ldots & x_{1k_j}^{(j)} \\ \vdots & & \vdots \\ x_{n1}^{(j)} & \ldots & x_{nk_j}^{(j)} \end{bmatrix} \qquad j = 1,\ldots,p \qquad (3)$$

Uit (3) volgt, dat de door Wansbeek voorgestelde transformatie van model (2) leidt tot een toename van het aantal verklarende variabelen in het i-de deelmodel van k_i tot $\Sigma_{i=1}^{p} k_i \equiv k$. Om te voorkomen dat de s^2 van het i-de deelmodel nul wordt, dienen er dus minstens k + 1 waarnemingen beschikbaar te zijn. Ruwweg gesproken, neemt in de schattingsprocedure van Wansbeek het minimaal benodigde aantal waarnemingen dus lineair toe met het aantal deelmodellen, terwijl in onze schattingsprocedure het minimaal benodigde aantal waarnemingen, i.c. $\max_i\{k_i + 1\}$, onafhankelijk is van het aantal deelmodellen. Daar de door ons voorgestelde inperking van de covariantiematrix juist ontworpen is om grote stelsels te kunnen schatten met kleine aantallen waarnemingen, lijkt Wansbeek's transformatie de zin ervan te ondergraven.

Voorts blijkt dat de vector van storingen u invariant is tegen de transformatie met de matrix $I_n \otimes E_p$. Immers

$$[I_n \otimes E_p] \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} E_p u_1 \\ \vdots \\ E_p u_n \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

daar $\iota_p' u_t = 0$ $(t = 1,\ldots,n)$ vanwege de optelrestrictie. Omdat de storingen niet gevoelig zijn voor deze transformatie, wordt via de s^2 van het i-de deelmodel niet d_i, maar de variantie van u_{ti}, i.c. $d_i - (\iota_p' d)^{-1} d_i^2$ geschat. Om schatters te bepalen voor de d_i's, zal ook in Wansbeek's schattingsprocedure hetzelfde stelsel van vergelijkingen moeten worden opgelost als in onze schattingsprocedure, te weten

$$s_i^2 = d_i - (\iota_p' d)^{-1} d_i^2 \qquad i = 1,\ldots,p \tag{4}$$

Daar, zoals door ons aangetoond, het stelsel (4) - op één geval na, dat zich met kans nul voordoet - een unieke oplossing heeft, zal ook Wansbeek vooralsnog de mogelijkheid dat één van de d_i's negatief wordt niet kunnen uitsluiten.

(c) Met betrekking tot de interpretatie van de covarianties dient opgemerkt te worden, dat, hoewel de door Wansbeek gehanteerde Moore-Penrose inverse uniek is, dit niet geldt voor de door hem gehanteerde decompositie $\Omega^+ = E_p D^{-1} E_p$. Immers, daar $\iota_p' E_p = 0$, geldt bij voorbeeld ook voor

$$\Sigma^{-1} \equiv D^{-1} + \xi \iota_p \iota_p'$$

dat

$$\Omega^+ = E_p \Sigma^{-1} E_p = E_p D^{-1} E_p .$$

Gemakkelijk valt na te gaan dat de inverse van Σ^{-1} wordt gegeven door

$$\Sigma = D - \frac{\xi}{1 + \xi(\iota_p'd)} dd' \qquad \xi \neq -(\iota_p'd)^{-1} \tag{5}$$

en dat, voor $\xi > 0$, Σ positief-definiet is indien alle d_i's positief zijn, maar ook als $d_1 < 0$, $d_i > 0$ $(i = 2,\ldots,p)$ en $(\iota_p'd) < -\xi^{-1}$, hetgeen weer iets meer licht werpt op de mogelijkheid van een negatieve d_1. De suggestie van Wansbeek volgend, kun je dus enerzijds zeggen, dat de covarianties uitsluitend bepaald worden door de optelrestrictie ($\Sigma = D$), maar anderzijds evengoed dat er sprake is van "echte" covarianties (b.v. $\Sigma = D - (1+\iota_p'd)^{-1}dd'$, i.c. (5) met $\xi = 1$). Kortom, het is niet mogelijk via het construeren van een denkbeeldig niet-singulier SUR-model te komen tot een éénduidige interpretatie van de covarianties.

Hoewel uit het voorgaande duidelijk zal zijn, dat de door Wansbeek voorgestelde vereenvoudiging naar onze mening onjuiste schatters voor de d_i's oplevert, zijn wij hem, het zij nogmaals herhaald, erkentelijk voor zijn commentaar, dat ons heeft gestimuleerd dieper na te denken over de interpretatie van de door ons voorgestelde specificatie.