Uitwerkingen Tentamen Econometrie 3, FEE-UvA
20 januari 2003, 9:30-11:30 uur, gebouw C, zaal 206

1. Gegeven is het model met slechts één verklarende variabele $y_i = x_i\beta + \varepsilon_i$, $i = 1, ..., N$, waarbij $x_i \equiv y_{i-1}$. Men beschikt louter over waarnemingen $\{y_0, ..., y_N\}$ en we nemen aan dat $N = 20$ en $\varepsilon_i \sim$i.i.d.N$(0, \sigma^2)$.

   (a) $\{x_1, ..., x_N\} = \{y_0, ..., y_{N-1}\}$, waarbij $y_1 = \beta y_0 + \varepsilon_1$, $y_2 = \beta^2 y_0 + \beta\varepsilon_1 + \varepsilon_2$, ... zodat $y_i = \sum_{j=0}^{i-1} \beta^j \varepsilon_{i-j} + \beta^i y_0$. Dus, y_i hangt samen met $\{\varepsilon_1, ..., \varepsilon_i\}$, zodat $E\{x_i\varepsilon_j\} = E\{y_{i-1}\varepsilon_j\} \neq 0$ voor $j = 1, ..., i-1$. We vinden dus dat x_i wel onafhankelijk is van $\{\varepsilon_i, ..., \varepsilon_N\}$, maar niet van $\{\varepsilon_1, ..., \varepsilon_{i-1}\}$, zodat (A2) niet geldt.

   (b) Het niet gelden van (A2) heeft voor de verdeling van de OLS schatter b van β tot gevolg dat hij niet zuiver is. Zijn verwachting en variantie zijn niet simpel analytisch te bepalen, en ondanks de normaliteit van de storingen is b niet normaal verdeeld, zodat de t-toets voor $\beta = \beta_0$ niet exact is.

   (c) Hier geldt wel (A7) $E\{x_i\varepsilon_i\} = 0$, zoals onder (a) al werd aangetoond.

   (d) Er geldt (A6) $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N x_i^2 = \Sigma_{xx} > 0$. Dan is b raak (consistent) en asymptotisch normaal verdeeld, want $\sqrt{N}(b-\beta) \to$N$(0, \sigma^2\Sigma_{xx}^{-1})$.

   (e) De raakheid is een pleister op de wonde dat de schatter niet zuiver is; als N maar heel groot is tendeert b toch naar β. Maar, dat sluit niet uit dat er bij $N = 20$ toch systematische fouten gemaakt worden door b, die bijvoorbeeld vaker negatief dan positief zijn. Als N heel erg groot is dan is b ontaard verdeeld (variantie gaat naar nul). Maar de asymptotische normaliteit van $\sqrt{N}(b-\beta)$ geeft hoop dat als N eindig is $b-\beta$ bij benadering normaal verdeeld is, met ongeveer verwachting nul en ongeveer variantie $N^{-1}\sigma^2\Sigma_{xx}^{-1} \approx N^{-1}\sigma^2(\frac{1}{N}\sum_{i=1}^N x_i^2)^{-1} = \sigma^2(\sum_{i=1}^N x_i^2)^{-1} = \sigma^2(\sum_{i=1}^N y_{i-1}^2)^{-1}$, en dat is de uitdrukking die van toepassing is als (A2) wel zou gelden. Of die benadering met de standaardformule al redelijk goed is als N slechts 20 is zou met Monte Carlo vastgesteld kunnen worden.

   (f) Voor $r = 1, ..., R$ bepalen we $b^{(r)} = \sum_{i=1}^N y_i y_{i-1} / \sum_{i=1}^N y_{i-1}^2$. Er geldt $E\{b^{(r)}\} = E\{b\}$ en $V\{b^{(r)}\} = V\{b\}$ terwijl de $b^{(r)}$ onderling onafhankelijk zijn. Voor $\bar{b} = R^{-1}\sum_{r=1}^R b^{(r)}$ geldt daarom $E\{\bar{b}\} = E\{b\}$ met $V\{\bar{b}\} = V\{R^{-1}\sum_{r=1}^R b^{(r)}\} = R^{-2}V\{\sum_{r=1}^R b^{(r)}\} = R^{-1}V\{b\}$ en volgens de Centrale Limiet Stelling $\sqrt{R}(\bar{b} - E\{b\}) \to$N$(0, V\{b\})$. Bij benadering geldt dus $(\bar{b} - E\{b\})/\sqrt{R^{-1}V\{b\}} \sim$N$(0,1)$. Een bij benadering 95% betrouwbaarheidsinterval voor $E\{b\}$ is dus $\bar{b} \pm 2\sqrt{R^{-1}\hat{V}\{b\}}$. $V\{b\}$ schatten we met $\hat{V}\{b\} = (R-1)^{-1}\sum_{r=1}^R (b^{(r)} - \bar{b})^2$. Voor de lengte van het interval willen we dat $4\sqrt{R^{-1}\hat{V}\{b\}} < 0.8 \times 10^{-3}$. Dan moet gelden $\sqrt{R^{-1}\hat{V}\{b\}} < 0.2 \times 10^{-3}$ of $\hat{V}\{b\} < 2 \times 10^{-4}R$. In de simulatie kan men dus R op blijven voeren tot dat geldt $R > 0.5 \times 10^4 \times \hat{V}\{b\}$ (want de schatter $\hat{V}\{b\}$ varieert wel wat, maar gaat toch voor R groot naar de eindige constante $V\{b\}$).

2. We hebben panel data $\{y_{ij}, x_{ij}\}$ voor 2 bedrijfstakken, $i = 1, 2$. Voor de eerste bedrijfstak hebben we N_1 waarnemingen en voor de tweede N_2. We veronderstellen

$$y_{ij} = x_{ij}'\beta + \varepsilon_{ij} \qquad i = 1, 2;\ j = 1, ..., N_i$$

   waarbij x_{ij} een niet-stochastische vector van K elementen is. Voorts zijn alle ε_{ij} onderling onafhankelijk met $E\{\varepsilon_{ij}\} = 0$ en $E\{\varepsilon_{ij}^2\} = \sigma_i^2$.

(a) Neem $N = N_1 + N_2$ en maak de $N \times 1$ vectoren y en ε en de $N \times K$ matrix X

$$\begin{aligned} y' &= (y_{11}, ..., y_{1N_1}, y_2, ..., y_{2N_2}), \\ X' &= (x_{11}, ..., x_{1N_1}, x_{21}, ..., x_{2N_2}), \\ \varepsilon' &= (\varepsilon_{11}, ..., \varepsilon_{1N_1}, \varepsilon_2, ..., \varepsilon_{2N_2}). \end{aligned}$$

dan geldt $y = X\beta + \varepsilon$.

(b) $E\{\varepsilon\} = 0$ en

$$E\{\varepsilon\varepsilon'\} = \begin{pmatrix} \sigma_1^2 I_{N_1} & O \\ O & \sigma_2^2 I_{N_2} \end{pmatrix}.$$

Er is dus sprake van heteroskedasticiteit in dit regressiemodel?

(c) De dichtheid van y is

$$f(y; \beta, \sigma_1^2, \sigma_2^2) = \prod_{i=1}^{2} \prod_{j=1}^{N_i} \left(2\pi\sigma_i^2\right)^{-1/2} \exp\left\{-0.5(y_{ij} - x_{ij}'\beta)^2/\sigma_i^2\right\}.$$

Voor de log-likelihood geldt dus

$$\log L(\beta, \sigma_1^2, \sigma_2^2) = -\frac{1}{2} \sum_{i=1}^{2} \left[N_i \ln(2\pi) + N_i \ln \sigma_i^2 + \frac{1}{\sigma_i^2} \sum_{j=1}^{N_i} \left(y_{ij} - x_{ij}'\beta\right)^2 \right].$$

(d) Er geldt

$$\begin{aligned} \frac{\partial}{\partial\beta} \log L(\beta, \sigma_1^2, \sigma_2^2) &= -\frac{1}{2} \sum_{i=1}^{2} \left[\frac{1}{\sigma_i^2} \sum_{j=1}^{N_i} \frac{\partial}{\partial\beta} \left(y_{ij} - x_{ij}'\beta\right)^2 \right] \\ &= \sum_{i=1}^{2} \frac{1}{\sigma_i^2} \sum_{j=1}^{N_i} x_{ij} \left(y_{ij} - x_{ij}'\beta\right). \end{aligned}$$

Dus we zoeken de waarde $\hat{\beta}$ die voldoet aan

$$\sum_{i=1}^{2} \sum_{j=1}^{N_i} \left(\frac{1}{\hat{\sigma}_i^2} x_{ij} y_{ij} - \frac{1}{\hat{\sigma}_i^2} x_{ij} x_{ij}' \hat{\beta} \right) = 0.$$

Dat zal in een iteratie moeten gebeuren. Bijvoorbeeld, eerst wordt aangenomen $\sigma_1^2 = \sigma_2^2$ zodat β gewoon met OLS geschat wordt, dus $\hat{\beta}^{(1)} = b$. Dan, zie (e), kan de eerste schatting van σ_1^2 en σ_2^2 bepaald worden, $\hat{\sigma}_i^{2(1)} = \frac{1}{N_i} \sum_{j=1}^{N_i} \left(y_{ij} - x_{ij}'\hat{\beta}^{(1)}\right)^2$, $i = 1, 2$. Die worden weer gebruikt om $\hat{\beta}^{(2)}$ uit $\sum_{i=1}^{2} \sum_{j=1}^{N_i} \left(\hat{\sigma}_i^{-2(1)} x_{ij} y_{ij} - \hat{\sigma}_i^{-2(1)} x_{ij} x_{ij}' \hat{\beta}^{(2)}\right) = 0$ te berekenen, enz., totdat er convergentie optreedt.

(e) Er geldt

$$\begin{aligned} \frac{\partial}{\partial\sigma_i^2} \log L(\beta, \sigma_1^2, \sigma_2^2) &= -\frac{1}{2} \frac{\partial}{\partial\sigma_i^2} \left[N_i \ln(2\pi) + N_i \ln \sigma_i^2 + \frac{1}{\sigma_i^2} \sum_{j=1}^{N_i} \left(y_{ij} - x_{ij}'\beta\right)^2 \right] \\ &= -\frac{1}{2} \left[\frac{N_i}{\sigma_i^2} - \frac{1}{\sigma_i^4} \sum_{j=1}^{N_i} \left(y_{ij} - x_{ij}'\beta\right)^2 \right], \quad i = 1, 2 \end{aligned}$$

zodat

$$-\frac{1}{2}\left[\frac{N_i}{\hat{\sigma}_i^2}-\frac{1}{\hat{\sigma}_i^4}\sum_{j=1}^{N_i}\left(y_{ij}-x_{ij}'\hat{\beta}\right)^2\right]=0,\ i=1,2$$

en voor de maximum likelihood schatters van σ_1^2 en σ_2^2 vinden we dus

$$\hat{\sigma}_i^2=\frac{1}{N_i}\sum_{j=1}^{N_i}\left(y_{ij}-x_{ij}'\hat{\beta}\right)^2,\ i=1,2.$$

De $\hat{\beta}$ is een gewogen kleinste kwadraten schatter die iteratief, met de $\hat{\sigma}_1^2$ en $\hat{\sigma}_2^2$ bepaald moet worden.

(f) We vinden

$$\frac{\partial}{\partial\sigma_1^2}\frac{\partial}{\partial\sigma_2^2}\log L(\beta,\sigma_1^2,\sigma_2^2)=0$$

en

$$\begin{aligned}\frac{\partial^2}{(\partial\sigma_i^2)^2}\log L(\beta,\sigma_1^2,\sigma_2^2) &= -\frac{1}{2}\frac{\partial}{\partial\sigma_i^2}\left[\frac{N_i}{\sigma_i^2}-\frac{1}{\sigma_i^4}\sum_{j=1}^{N_i}\left(y_{ij}-x_{ij}'\beta\right)^2\right]\\ &= \frac{1}{2}\left[\frac{N_i}{\sigma_i^4}-\frac{2}{\sigma_i^6}\sum_{j=1}^{N_i}\left(y_{ij}-x_{ij}'\beta\right)^2\right],\ i=1,2\end{aligned}$$

zodat

$$\begin{aligned}E\left\{\frac{\partial^2}{(\partial\sigma_i^2)^2}\log L(\beta,\sigma_1^2,\sigma_2^2)\right\} &= \frac{1}{2}E\left[\frac{N_i}{\sigma_i^4}-\frac{2}{\sigma_i^6}\sum_{j=1}^{N_i}\left(y_{ij}-x_{ij}'\beta\right)^2\right]\\ &= \frac{1}{2}\left[\frac{N_i}{\sigma_i^4}-\frac{2N_i\sigma_i^2}{\sigma_i^6}\right]\\ &= -\frac{N_i}{2\sigma_i^4},\ i=1,2.\end{aligned}$$

Het gedeelte van de blok-diagonale informatie matrix dat met de parameters σ_1^2 en σ_2^2 correspondeert is dus

$$\begin{pmatrix}\frac{N_1}{2\sigma_1^4} & 0\\ 0 & \frac{N_2}{2\sigma_2^4}\end{pmatrix}.$$

(g) De Wald toetsgrootheid voor $H_0:\sigma_1^2=\sigma_2^2$ is dus, met $R=(0',1,-1)$ en $\theta'=(\beta',\sigma_1^2,\sigma_2^2)$,

$$\begin{aligned}W &= (R\hat{\theta}-0)'\left[R\hat{I}_{\theta\theta}^{-1}R'\right]^{-1}(R\hat{\theta}-0)\\ &= (\hat{\sigma}_1^2-\hat{\sigma}_2^2)\left[\frac{2\hat{\sigma}_1^2}{N_1}+\frac{2\hat{\sigma}_2^2}{N_2}\right]^{-1}(\hat{\sigma}_1^2-\hat{\sigma}_2^2)\\ &= \frac{1}{2}\frac{(\hat{\sigma}_1^2-\hat{\sigma}_2^2)^2}{\frac{\hat{\sigma}_1^2}{N_1}+\frac{\hat{\sigma}_2^2}{N_2}}.\end{aligned}$$

Voor de likelihood ratio toetsgrootheid moeten we ook het gerestricteerde model schatten. Dat impliceert gewoon OLS in het $y = X\beta + \varepsilon$ model, waarvoor geldt

$$\log L(\beta, \sigma^2) = -\frac{1}{2}\left[N \ln(2\pi) + N \ln \sigma^2 + \frac{1}{\sigma^2} \sum_{l=1} (y_l - x_l'\beta)^2 \right]$$

met maximum

$$\begin{aligned}
\log L(b, \hat{\sigma}^2) &= -\frac{1}{2}\left[N \ln(2\pi) + N \ln \hat{\sigma}^2 + \frac{1}{\hat{\sigma}^2} \sum_{l=1} (y_l - x_l'b)^2 \right] \\
&= -\frac{1}{2}\left\{ N \ln(2\pi) + N \ln \hat{\sigma}^2 + N \right\} \\
&= -\frac{N}{2}\left\{ \ln(2\pi) + 1 + \ln \hat{\sigma}^2 \right\},
\end{aligned}$$

zodat

$$\begin{aligned}
LR &= -\sum_{i=1} \left\{ N_i \ln(2\pi) + N_i \ln \hat{\sigma}_i^2 + N_i \right\} + N\left\{ \ln(2\pi) + 1 + \ln \hat{\sigma}^2 \right\} \\
&= N \ln \hat{\sigma}^2 - N_1 \ln \hat{\sigma}_1^2 - N_2 \ln \hat{\sigma}_2^2 \\
&= N_1 \ln \frac{\hat{\sigma}^2}{\hat{\sigma}_1^2} + N_2 \ln \frac{\hat{\sigma}^2}{\hat{\sigma}_2^2}.
\end{aligned}$$

3. Voor een zeker model wordt de log-likelihood gegeven door $\log L(\theta)$, waarbij $\theta = (\theta_1', \theta_2')'$ een parameter vector is die uit $K_1 + K_2$ elementen bestaat. We willen toetsen $H_0 : \theta_2 = \theta_2^0$. De typische kenmerken (zowel overeenkomsten als verschillen) van respectievelijk de Wald, likelihood ratio en de Lagrange multiplier toetsgrootheid zijn de volgende. Bij LR moet het model zowel onder als zonder de restricties geschat worden, bij W alleen zonder, en bij LM alleen onder de restricties. Onder H_0 zijn de toetsgrootheden ieder asymptotisch chi-kwadraat met K_2 vrijheidsgraden verdeeld. De toetsgrootheden zijn asymptotisch equivalent, ook onder de alternatieve hypothese. Alle drie de toetsen zijn raak (een foute H_0 wordt als de steekproef maar groot genoeg is altijd verworpen). Maar, in eindige steekproef zijn ze, behoudens in heel simplistische modellen, onderling verschillend, en niet exact. De een kan dus, als H_0 waar is, systematisch vaker verwerpen dan de andere(n). Overwegingen bij het maken van een keuze zijn dus de overeenkomst tussen het nominale en actuele significantie niveau, en het gemak waarmee de grootheid kan worden uitgerekend. Als het gerestricteerde model lineair is is LM gemakkelijker uit te rekenen dan W als de alternative hypothese niet-lineair is, en vice versa.