

Uitwerkingen en Verslag Tentamen Econometrie 3, FEE-UvA
27 oktober 2003, 9:30-11:30 uur, gebouw B, tentamenzaal

Er schreven zich 19 studenten in, waarvan er 18 op kwamen dagen, en 17 hun werk inleverden. Van de maximaal 75 punten werden er de volgende scores gerealiseerd: 74, 73, 72, 71, 69, 64, 64, 63, 57, 51, 51, 47, 46, 45, 39, 24, 23. Met de practicum cijfers samen leidde dat tot de volgende eindresultaten: 2x9.5, 3x9, 1x8.5, 1x8, 1x7, 4x6.5, 3x6, 2x4.

Hieronder compacte uitwerkingen, met vermelding van enkele typische fouten.

1. Beschouw het regressie model met slechts één coëfficiënt $y_i = \beta \frac{1}{x_i} + \varepsilon_i$, waarbij $x_i \neq 0$, $i = 1, \dots, N$. Neem aan dat de vector $x = (x_1, \dots, x_N)'$ niet stochastisch is en dat $\varepsilon_i \sim \text{i.i.d.} N(0, \sigma^2)$. Het model heeft dus 2 parameters $\theta = (\beta, \sigma^2)'$ en die gaan we schatten met maximum likelihood. De cumulatieve dichtheid voor de waarnemingen $y = (y_1, \dots, y_N)'$ luidt

$$f(y | x; \beta, \sigma^2) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \frac{\left(y_i - \frac{\beta}{x_i}\right)^2}{\sigma^2} \right\}.$$

- (a) {5} De log likelihood functie is

$$\begin{aligned} \ln L(\beta, \sigma^2) &= \ln \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \frac{\left(y_i - \frac{\beta}{x_i}\right)^2}{\sigma^2} \right\} \\ &= -\frac{1}{2} \sum_{i=1}^N \left(\ln(2\pi) + \ln \sigma^2 + \frac{\left(y_i - \frac{\beta}{x_i}\right)^2}{\sigma^2} \right) \\ &= -\frac{1}{2} \left(N \ln(2\pi) + N \ln \sigma^2 + \frac{1}{\sigma^2} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right)^2 \right). \end{aligned}$$

- (b) {10} De twee elementen van de score vector (voor de hele steekproef) zijn

$$\begin{aligned} \frac{\partial \ln L(\beta, \sigma^2)}{\partial \beta} &= -\frac{1}{2} \frac{2}{\sigma^2} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right) \left(-\frac{1}{x_i}\right) = \frac{1}{\sigma^2} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right) \frac{1}{x_i} \\ \frac{\partial \ln L(\beta, \sigma^2)}{\partial \sigma^2} &= -\frac{N}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{1}{\sigma^4} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right)^2. \end{aligned}$$

Fout: $\frac{\partial \ln L(\beta, \sigma^2)}{\partial \beta} = \frac{1}{\sigma^2} \frac{1}{x_i} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right)$.

- (c) {5} De score vector heeft verwachting nul want $(\beta, \sigma^2$ en $1/x_i$ zijn niet stochastisch):

$$\begin{aligned} E \frac{\partial \ln L(\beta, \sigma^2)}{\partial \beta} &= \frac{1}{\sigma^2} E \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i}\right) \frac{1}{x_i} = \frac{1}{\sigma^2} \sum_{i=1}^N E \left(\varepsilon_i \frac{1}{x_i} \right) = \frac{1}{\sigma^2} \sum_{i=1}^N \frac{1}{x_i} E(\varepsilon_i) = 0 \\ E \frac{\partial \ln L(\beta, \sigma^2)}{\partial \sigma^2} &= -\frac{N}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{1}{\sigma^4} \sum_{i=1}^N E \left(y_i - \frac{\beta}{x_i}\right)^2 = -\frac{N}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{1}{\sigma^4} \sum_{i=1}^N \sigma^2 \\ &= -\frac{N}{2} \frac{1}{\sigma^2} + \frac{N}{2} \frac{1}{\sigma^2} = 0. \end{aligned}$$

Fout: Zondermeer opschrijven dat $E\{\varepsilon_i \frac{1}{x_i}\} = E\{\varepsilon_i\}/E\{x_i\}$ (wat hier overigens wel geldt omdat x_i niet stochastisch is, dus $E\{x_i\} = x_i$)

(d) {5} De formule voor de maximum likelihood schatter $\hat{\beta}$ van β wordt gevonden uit

$$\frac{1}{\hat{\sigma}^2} \sum_{i=1}^N \left(y_i - \frac{\hat{\beta}}{x_i} \right) \frac{1}{x_i} = 0.$$

Dit geeft $\sum_{i=1}^N \left(y_i - \frac{\hat{\beta}}{x_i} \right) \frac{1}{x_i} = 0$, zodat $\sum_{i=1}^N \frac{y_i}{x_i} = \hat{\beta} \sum_{i=1}^N \frac{1}{x_i^2}$, en dat geeft

$$\hat{\beta} = \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{y_i}{x_i}.$$

Veel gemaakte fout: Uit $\sum_{i=1}^N \left(y_i - \frac{\hat{\beta}}{x_i} \right) \frac{1}{x_i} = 0$ concluderen dat $\sum_{i=1}^N \left(y_i - \frac{\hat{\beta}}{x_i} \right) = 0$ (wat alleen mag als x_i constant is) en uit $\sum_{i=1}^N \left(y_i - \frac{\hat{\beta}}{x_i} \right) = 0$ (wat dus niet waar is) ook nog concluderen dat $\hat{\beta} = \sum_{i=1}^N y_i x_i$ inplaats van $\hat{\beta} = \sum_{i=1}^N y_i \sum_{i=1}^N x_i$.

(e) {5} De schatter $\hat{\beta}$ is zuiver voor β omdat

$$\begin{aligned} \hat{\beta} &= \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{y_i}{x_i} = \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{1}{x_i} \left(\beta \frac{1}{x_i} + \varepsilon_i \right) \\ &= \beta + \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{1}{x_i} \varepsilon_i \end{aligned}$$

zodat, omdat x_i *niet* stochastisch is

$$E(\hat{\beta}) = \beta + \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{1}{x_i} E(\varepsilon_i) = \beta + \left(\sum_{i=1}^N \frac{1}{x_i^2} \right)^{-1} \sum_{i=1}^N \frac{1}{x_i} 0 = \beta.$$

Fout: zij die met de foute schatter $\hat{\beta} = \sum_{i=1}^N y_i x_i$ op de proppen kwamen, meenden dat die toch zuiver was, omdat $E(\hat{\beta}) = \sum_{i=1}^N E\{y_i x_i\} = \sum_{i=1}^N \beta \frac{x_i}{x_i} + \sum_{i=1}^N x_i E(\varepsilon_i) = \sum_{i=1}^N \beta$ simpel gelijk gesteld wordt aan β .

(f) {5} Eigenlijk is dit een gewone regressie van y_i op de variabele $z_i = 1/x_i$ zonder constante. De schatter van β wordt gewoon gevonden met kleinste kwadraten op basis van maar 1 normaalvergelijking, namelijk

$$\sum_{i=1}^N \left(y_i - \hat{\beta} z_i \right) z_i = 0$$

ofwel $\sum_{i=1}^N \hat{\varepsilon}_i z_i = 0$. Als er een constante in de regressie zat, dan sommeerden de residuen tot nul. Nu sommeert alleen hun gewogen som (met gewichten z_i) tot nul. Alleen als alle x_i aan elkaar gelijk zijn, dus $x_i = c \neq 0$ dan geldt $\sum_{i=1}^N \hat{\varepsilon}_i \frac{1}{x_i} = c \sum_{i=1}^N \hat{\varepsilon}_i = 0$ en sommeren de residuen tot nul.

(g) {5} De informatie matrix is een 2×2 matrix. Het niet-diagonale element is gelijk aan

$$\begin{aligned} -E \frac{\partial^2 \ln L(\beta, \sigma^2)}{\partial \beta \partial \sigma^2} &= -E \left[\frac{\partial}{\partial \sigma^2} \frac{1}{\sigma^2} \sum_{i=1}^N \left(y_i - \frac{\beta}{x_i} \right) \frac{1}{x_i} \right] \\ &= E \left[\frac{1}{\sigma^4} \sum_{i=1}^N \varepsilon_i \frac{1}{x_i} \right] = \frac{1}{\sigma^4} \sum_{i=1}^N \frac{1}{x_i} E(\varepsilon_i) = 0, \end{aligned}$$

dus de informatiematrix is blok-diagonaal.

2. Een onderzoeker heeft OLS schatters bepaald voor de regressie van een tijdreeks variabele y_t ($t = 1, \dots, T$) op een $K \times 1$ vector van regressoren x_t , waaronder geen vertraagd-afhankelijke variabelen voorkomen.

- (a) {5} Opdat de OLS schatters voor de coëfficiënten raak ('consistent') zijn voor β moet aangenomen worden dat $E\{(y_t - x_t'\beta)x_t\} = 0$ (dus $\varepsilon_t = y_t - x_t'\beta$ en alle K elementen van x_t zijn ongecorrleerd) en dat $\text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_t x_t' = \Sigma_{xx}$, waarbij Σ_{xx} een eindige niet singuliere matrix is. Fout: zonder blikken of blozen $\sum_{t=1}^T x_t^2$ opschrijven.
- (b) {5} Als de Durbin-Watson statistic kleiner is dan 2 en in de 'inconclusive region' ligt dan is er bij het betreffende significantie niveau α geen uitsluitel of de waarde nu wel of niet in het exacte kritieke gebied ligt van de toets $H_0 : \rho = 0$ tegen $H_1 : \rho > 0$ waarbij $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$ met $v_t \sim i.i.d.N(0, \sigma_v^2)$. Deze toets mag alleen worden toegepast als x_t onafhankelijk is van $\varepsilon_s = y_s - x_s'\beta$ voor alle $t, s = 1, \dots, T$. Er mogen dus niet alleen geen vertraagd afhankelijke variabelen y_{t-l} ($l > 0$) in x_t voorkomen, maar ook niet voor $l \leq 0$, bijv. verklarende variabelen x_t (zoals bij simultaneïteit) die met ε_t gecorrleerd zijn.
- (c) {10} Als de onderzoeker vervolgens besluit de met deze Durbin-Watson statistic getoetste nul-hypothese te verwerpen, dan rekent hij het onbesliste gebied dus ook tot het kritieke gebied, en daarmee creëer je dan een toets met een actueel significantie niveau dat groter of hooguit gelijk is aan het nominale significantie niveau α waarvoor de onder- en bovengrenzen van de exacte kritieke waarde in de gehanteerde tabel zijn gebruikt. Want, voor de onbekende exacte kritieke waarde d^α geldt $d_L^\alpha \leq d^\alpha \leq d_U^\alpha$. Voor de waargenomen waarde d geldt ook $d_L^\alpha \leq d \leq d_U^\alpha$. En, als H_0 waar is, geldt dus $\Pr\{d \leq d_U^\alpha\} \geq \Pr\{d \leq d^\alpha\} = \alpha$.
- (d) {5} Vervolgens schat hij het model opnieuw met OLS, maar nadat hij de regressoren y_{t-1} en (de vector) x_{t-1} aan de lijst van verklarende variabelen heeft toegevoegd. Argumenten daarvoor zijn de volgende: Hij vat de significante DW op als een signaal dat de specificatie verkeerd was en er ontbrekende regressoren zijn, bijv. de vertragingen y_{t-1} en x_{t-1} . Bovendien is het zo dat het model $y_t = x_t'\beta + \rho y_{t-1} + x_{t-1}'\delta + v_t$ ook ruimte biedt voor het oorspronkelijke model $y_t = x_t'\beta + \varepsilon_t$ met 1-orde autocorrelatie $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, want dat impliceert $(y_t - x_t'\beta) = \rho(y_{t-1} - x_{t-1}'\beta) + v_t$ ofwel $y_t = x_t'\beta + \rho y_{t-1} - x_{t-1}'\beta\rho + v_t$ dus dit komt overeen met het door de onderzoeker geschatte model wanneer de restricties $\delta = \beta\rho$ gelden.
- (e) {5} Alternatieve wegen (anders dan (d)) die de onderzoeker, op grond van de significant bevonden Durbin-Watson statistic, had kunnen gaan bewandelen zijn: (i) de functionele vorm aanpassen (misschien logs nemen van de variabelen in y_t en x_t); (ii) het model schatten met EGLS, dus aannemen dat er inderdaad sprake is van AR(1) storingen en de ρ samen met de β schatten, bijvoorbeeld zoals in de Cochran-Orcutt methode; (iii) op zoek gaan naar andere ontbrekende variabelen; (iv) de oorspronkelijke OLS schatters handhaven en hun variantie consistent schatten met Newey-West (HAC).
- (f) {5} De schatters van het in (d) aangeduide uitgebreide model zijn onder geen beding BLUE (best linear unbiased estimators), want door de vertraagd afhankelijke variabele zijn ze niet meer zuiver (Unbiased), maar hooguit wel raak, als $E\{(y_t - x_t'\beta - \rho y_{t-1} - x_{t-1}'\delta)x_t\} = 0$.
- (g) {5} Als je het onder (d) geschatte model op zijn beurt zou willen toetsen op eerste-orde autocorrelatie in de storingen dan schat je de volgende hulpregressie. Laten we de OLS residuen bij (d) \hat{v}_t noemen. Toets dan de hypothese $H_0 : \gamma = 0$ (gewoon met de t -toets) in de regressie $y_t = x_t'\beta + \rho y_{t-1} + x_{t-1}'\delta + \gamma \hat{v}_{t-1} + \text{storing}$. Fout: regressie van \hat{v}_t^2 op \hat{v}_{t-1}^2 ; \hat{y}_t^2 aan regressie toevoegen; White toets uitvoeren; minder fout: regressie doen van \hat{v}_t op \hat{v}_{t-1} .

3. Gegeven zijn de volgende relaties: (1) $A_t = \beta_1 + \beta_2 S_t + \varepsilon_t$ en (2) $S_t = \gamma_1 + \gamma_2 A_t + \gamma_3 P_t + \gamma_4 Y_t + \xi_t$. Hierbij is A_t de bestedingen aan advertenties (van een grote landelijke supermarktketen) per week, S_t is de behaalde omzet (β_2 zal waarschijnlijk wel negatief zijn, maar dat moet juist onderzocht gaan worden), P_t is een index voor de relatieve prijzen ten opzichte van andere supermarkten en Y_t is een index voor de gezinsinkomens. De storingen ε_t en ξ_t zijn in de tijd (en ook onderling) ongecorrleerd.

- (a) {10} Het is geen goed idee om vergelijking (1) met OLS te schatten om de volgende reden. Substitutie van (1) in (2) geeft

$$S_t = \gamma_1 + \gamma_2(\beta_1 + \beta_2 S_t + \varepsilon_t) + \gamma_3 P_t + \gamma_4 Y_t + \xi_t$$

ofwel

$$S_t = \frac{1}{1 - \gamma_2 \beta_2} [\gamma_1 + \gamma_2 \beta_1 + \gamma_3 P_t + \gamma_4 Y_t + (\xi_t + \gamma_2 \varepsilon_t)].$$

Dat is de herleide vorm vergelijking voor S_t en die laat zien dat S_t gecorreleerd is met ξ_t en met ε_t . Dat laatste is funest voor OLS schatters van (1) omdat niet zal gelden $E\{\varepsilon_t S_t\} = 0$, zodat de OLS schatters niet alleen niet zuiver, maar ook niet raak zullen zijn.

- (b) {10} Stel we schatten vergelijking (1) met instrumentele variabelen en gebruiken de constante en Y_i als instrument. Dat gaat als volgt. We definiëren $y = (A_1, \dots, A_T)'$, $x'_t = (1, S_t)$, $X = (x_1, \dots, x_T)'$, $z_t = (1, Y_t)$, $Z = (z_1, \dots, z_T)'$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)'$. Vergelijking (1) luidt nu $y = X\beta + \varepsilon$ en, als de z_t geschikte instrumenten zijn, dan geldt $E\{Z'\varepsilon\} = E\{Z'(y - X\beta)\} = 0$, zodat $E\{Z'y\} = E\{Z'X\}\beta$. Dit suggereert de volgende schatter: $Z'y = Z'X\hat{\beta}$, zodat, als $Z'X$ inverteerbaar is, $\hat{\beta} = (Z'X)^{-1}Z'y$. Die schatter is niet zuiver, maar bij voldoende regulariteit wel raak en asymptotisch normaal. Een geschikte schatter voor zijn variantie is $\hat{\sigma}^2(Z'X)^{-1}Z'Z(X'Z)^{-1} = \hat{\sigma}^2[X'Z(Z'Z)^{-1}Z'X]^{-1}$, waarbij $\hat{\sigma}^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/T$.