Econométrie

Licence Economie-Gestion. Parcours Analyse Economique

Professeur Georges Bresson

Session Janvier 2017

1. **Exercice 1** (3 points) - Soient deux variables aléatoires X et Y de moyennes μ_X et μ_Y et de variances σ_X^2 et σ_Y^2. Le coefficient de corrélation est donné par: $\rho = Cov\,(X,Y)\,/\sqrt{\sigma_X^2.\sigma_Y^2} = \sigma_{XY}/\,(\sigma_X\sigma_Y)$. On considère la relation linéaire $Y = \alpha + \beta X$ où α et β sont des scalaires.

   (a) Montrez que le meilleur estimateur linéaire[1] de Y, basé sur X, est donné par:

   $$\widehat{Y} = \widehat{\alpha} + \widehat{\beta}X \text{ avec } \widehat{\alpha} = \mu_Y - \widehat{\beta}\mu_X \text{ et } \widehat{\beta} = \rho\sigma_Y/\sigma_X$$

   (b) Montrez que $Var\left(\widehat{Y}\right) = \rho^2\sigma_Y^2$ et que l'erreur d'estimation $\widehat{u}\left(= Y - \widehat{Y}\right)$ est de moyenne nulle et de variance $Var\,(\widehat{u}) = \left(1-\rho^2\right)\sigma_Y^2$.

   (c) Montrez que $Cov\left(\widehat{Y}, \widehat{u}\right) = 0$.

2. **Exercice 2** (4 points) - Soit le modèle de régression simple:

   $$Y_i = \alpha + \beta X_i + u_i \ , \ i = 1, ..., n(=10)$$

   On suppose la présence d'hétéroscédasticité $E\left(u_i^2\right) = \sigma_i^2 = \sigma^2 X_i^\delta$ où $X_i = 1, 2, ..., 10$.

   (a) Calculez la variance de l'estimateur des MCO $Var\left(\widehat{\beta}_{MCO}\right)$ pour $\delta = 0.5, 1, 1.5$ et 2.

   (b) Calculez la variance de l'estimateur BLUE $Var\left(\widetilde{\beta}_{BLUE}\right)$ pour $\delta = 0.5, 1, 1.5$ et 2.

   (c) Déterminez l'efficacité relative de l'estimateur des MCO sous l'hypothèse d'hétéroscédasticité:

   $$\frac{Var\left(\widetilde{\beta}_{BLUE}\right)}{Var\left(\widehat{\beta}_{MCO}\right)}$$

   pour $\delta = 0.5, 1, 1.5$ et 2. Que se passe-t-il quand δ augmente?

3. **Exercice 3** (5 points) - On considère le modèle suivant:

   $$y_i = x_i\beta + u_i$$

   pour seulement deux observations: $i = 1, 2$ et où $|x_1| \neq |x_2|$ sont scalaires et non aléatoires. On suppose que:

   $$\begin{aligned} u_i &\sim N\left(0, \sigma^2\right) \\ u_1 &= \rho u_2 + \varepsilon, |\rho| < 1 \\ \varepsilon &\sim N\left(0, \left(1-\rho^2\right)\sigma^2\right) \end{aligned}$$

   où ε et u_2 sont indépendants.

   (a) Calculez l'estimateur MCO de β.

[1] au sens de la minimisation de l'erreur quadratique moyenne $E\left[(Y-\alpha-\beta X)^2\right]$.

(b) Calculez l'estimateur du maximum de vraisemblance de β.

(c) Montrez que l'estimateur du maximum de vraisemblance de ρ est non aléatoire.

(d) Comment se comportent les estimateurs du maximum de vraisemblance de β et de ρ quand $x_1 \to x_2$ (avec $x_2 \neq 0$)? pour $y_1 = y_2$ puis pour $y_1 \neq y_2$?

4. **Exercice 4** (8 points) - On dispose de données mensuelles relatives au marché de la viande bovine aux USA entre 1977 et 1986. Les variables sont: la quantité de viande de boeuf (q_t) en milliers de livres (*pounds* = 0.4533 kg), le prix réel (déflaté par l'indice des prix à la consommation) de la livre de boeuf (p_t), le revenu disponible réel des ménages (inc_t) (en milliers de dollars 2005), le prix réel du poulet entier ($chickp_t$) et le prix réel des aliments pour bétail (au kg) pour nourrir les boeufs ($feedp_t$).
On souhaite estimer un système d'offre et de demande de la viande bovine à l'équilibre:

$$\begin{aligned}
\log q_{d,t} &= \alpha_1 \log p_t + \alpha_2 \log inc_t + \alpha_3 \log chickp_t + \alpha_4 trend_t \quad \text{demande de viande de boeuf} \\
\log q_{o,t} &= \beta_1 \log p_t + \beta_2 \log feedp_t + \beta_3 trend_t \quad \text{offre de viande de boeuf} \\
q_{d,t} &= q_{o,t} = q_t \quad \text{équilibre du marché}
\end{aligned}$$

C'est un modèle volontairement sans constante et où la variable $trend_t$ est un trend ($trend_t = 1, 2, ..., T$). Les variables endogènes sont q_t et p_t, les autres sont considérées comme exogènes.

(a) Etudiez les conditions d'identification (conditions d'ordre seulement) de chacune des équations du modèle. Qu'en déduisez-vous?

(b) Rappelez succinctement le principe de la méthode des triples moindres carrés (3SLS) et des GMM.

(c) On estime le système complet par les 3SLS en utilisant les instruments suivants: $\log inc_t$, $\log chickp_t$, $\log feedp_t$, $\log p_{t-1}$, $trend_t$, $\log chickp_{t-1}$. Interprétez. Les résultats vous semblent-t-il cohérents?

(d) On estime le système complet par les GMM en supposant des jeux d'instruments différents pour chacune des équations. Les instruments pour la fonction de demande sont: $\log inc_t$, $\log chickp_t$, $\log feedp_t$, $\log p_{t-1}$, $trend_t$, $\log chickp_{t-1}$. Les instruments pour la fonction d'offre sont: $\log inc_t$, $\log chickp_t$, $\log feedp_t$, $trend_t$, $\log chickp_{t-1}$. On utilise d'abord une correction de White afin d'obtenir des estimateurs robustes, puis une correction HAC pour gérer l'hétéroscédasticité et l'autorrélation des résidus.

   i. Interprétez successivement les deux estimations GMM.

   ii. Quelles conclusions tirez-vous des tests d'Hansen-Sargan de restriction de sur-identification?

Aucun document autorisé.

Calculatrices et tables statistiques autorisées.

1.

/* 3SLS */

```
reg3 (log_q log_p log_inc log_chickp trend, noconstant) (log_q log_p log_feedp
trend, noconstant), 3sls inst(log_inc log_chickp log_feedp log_p_1 log_chickp_1
trend)
```

```
Three-stage least-squares regression
----------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
log_q                119       4    .0306045    1.0000   1.90e+07   0.0000
2log_q               119       3    .3546229    0.9992  141292.22   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
log_q        |
       log_p |  -.2179803   .0210379   -10.36   0.000    -.2592138   -.1767467
     log_inc |   1.600236   .0214653    74.55   0.000     1.558165    1.642308
  log_chickp |   .0052755   .0382088     0.14   0.890    -.0696124    .0801634
       trend |  -.0027882    .000084   -33.18   0.000    -.0029529   -.0026235
-------------+----------------------------------------------------------------
2log_q       |
       log_p |   1.975372   .1631691    12.11   0.000     1.655566    2.295178
   log_feedp |   .5700344   .2105796     2.71   0.007     .1573059    .9827629
       trend |   .0001223   .0011601     0.11   0.916    -.0021515     .002396
------------------------------------------------------------------------------
Endogenous variables:  log_q log_p
Exogenous variables:   log_inc log_chickp log_feedp log_p_1 log_chickp_1
     trend
------------------------------------------------------------------------------
```

/* GMM robustes */

```
gmm (eq1:log_q  - {xb: log_p log_inc log_chickp trend}) ///
    (eq2:log_q  - {xc: log_p log_feedp trend}),  ///
    instruments(eq1 : log_inc log_chickp log_feedp log_p_1 log_chickp_1 trend) ///
    instruments(eq2:  log_inc log_chickp log_feedp log_feedp_1 trend) ///
    winitial(unadjusted, independent) wmatrix(robust) twostep
estat overid
```

```
Step 1
Iteration 0:   GMM criterion Q(b) =  298.77242
I
teration 1:   GMM criterion Q(b) =  .01381732
Iteration 2:   GMM criterion Q(b) =  .01381732

Step 2
Iteration 0:   GMM criterion Q(b) =  .59888847
Iteration 1:   GMM criterion Q(b) =  .45581462
Iteration 2:   GMM criterion Q(b) =  .45581462

GMM estimation

Number of parameters =   7
Number of moments    =  13
Initial weight matrix: Unadjusted                 Number of obs   =        119
```

```
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
xb           |
       log_p |  -.2474424    .016845   -14.69   0.000    -.2804581   -.2144268
     log_inc |   1.595988   .0183267    87.09   0.000     1.560068    1.631907
  log_chickp |   .0485736   .0357554     1.36   0.174    -.0215057     .118653
       trend |  -.0028631   .0000588   -48.69   0.000    -.0029783   -.0027478
-------------+----------------------------------------------------------------
xc           |
       log_p |   3.716254    .284267    13.07   0.000     3.159101    4.273407
   log_feedp |  -1.679958   .3773771    -4.45   0.000    -2.419603   -.9403121
       trend |  -.0068176   .0017814    -3.83   0.000    -.0103091   -.0033262
------------------------------------------------------------------------------
Instruments for equation eq1: log_inc log_chickp log_feedp log_p_1 log_chickp_1 trend _cons
Instruments for equation eq2: log_inc log_chickp log_feedp log_feedp_1 trend _cons

Test of overidentifying restriction:

  Hansen's J chi2(6) = 54.2419 (p = 0.0000)
```

```
/* GMM HAC */

gmm (eq1:log_q  - {xb: log_p log_inc log_chickp trend}) ///
    (eq2:log_q  - {xc: log_p log_feedp trend}),  ///
    instruments(eq1 : log_inc log_chickp log_feedp log_p_1 log_chickp_1 trend) ///
    instruments(eq2:  log_inc log_chickp log_feedp log_feedp_1 trend) ///
    winitial(unadjusted, independent) ///
    wmatrix(hac nwest opt) twostep
estat overid
```

```
Step 1
Iteration 0:   GMM criterion Q(b) =  298.77242
Iteration 1:   GMM criterion Q(b) =  .01381732
Iteration 2:   GMM criterion Q(b) =  .01381732

Step 2
Iteration 0:   GMM criterion Q(b) =   .0592286
Iteration 1:   GMM criterion Q(b) =  .04933705
Iteration 2:   GMM criterion Q(b) =  .04933705

GMM estimation

Number of parameters =   7
Number of moments    =  13
Initial weight matrix: Unadjusted                 Number of obs   =        119
GMM weight matrix:     HAC Bartlett 17
                       (lags chosen by Newey-West)
```

	Coef.	HAC Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
log_p	-.2483922	.0328089	-7.57	0.000	-.3126964	-.1840879
log_inc	1.596268	.0149736	106.61	0.000	1.56692	1.625616
log_chickp	.0481432	.0476918	1.01	0.313	-.045331	.1416175
trend	-.0027916	.0000889	-31.39	0.000	-.0029659	-.0026173
xc						
log_p	3.18263	.2976799	10.69	0.000	2.599188	3.766072
log_feedp	-.9917417	.3894863	-2.55	0.011	-1.755121	-.2283626
trend	-.0043996	.0034946	-1.26	0.208	-.0112488	.0024497

HAC standard errors based on Bartlett kernel with 16 lags.
(Lags chosen by Newey-West method.)
Instruments for equation eq1: log_inc log_chickp log_feedp log_p_1 log_chickp_1 trend _cons
Instruments for equation eq2: log_inc log_chickp log_feedp log_feedp_1 trend _cons

Test of overidentifying restriction:

Hansen's J chi2(6) = 5.87111 (p = 0.4378)