

Paris

Session : JANVIER 2023 – 1^{er} semestre

Année d'étude : L3 – Economie – Gestion. Parcours Analyse Economique

Discipline : Econométrie

Titulaire(s) du cours : Professeur Georges Bresson

Durée de l'épreuve : 3 heures

Document(s) autorisé(s) :

Aucun document autorisé

Calculatrices et tables statistiques autorisées

Ce sujet comporte 6 pages. Avant de composer, veuillez vérifier que votre sujet est complet.

Econométrie

Licence Economie-Gestion. Parcours Analyse Economique. Label CMI EFIQuaS.

Professeur Georges Bresson

Session Janvier 2023

1. **Exercice 1** (4 pts.) - Soit le modèle de régression linéaire simple :

$$y_i = \alpha + \beta X_i + u_i, \quad i = 1, \dots, n (= 10)$$

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

- (a) Calculez la variance de l'estimateur des MCO $Var(\hat{\beta}_{MCO})$ pour $\delta = 0.5, 1, 1.5$ et 2 .
(b) Calculez la variance de l'estimateur BLUE $Var(\tilde{\beta}_{BLUE})$ pour $\delta = 0.5, 1, 1.5$ et 2 .
(c) Déterminez l'efficacité relative de l'estimateur des MCO sous l'hypothèse d'hétéroscédasticité :

$$\frac{Var(\tilde{\beta}_{BLUE})}{Var(\hat{\beta}_{MCO})}$$

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

2. **Exercice 2** (4 pts.) - On considère la régression linéaire simple :

$$y_i = \alpha + \beta X_i + u_i, \quad i = 1, \dots, n$$

où α et β sont des scalaires et $u_i \sim N(0, \sigma^2)$.

Dérivez la statistique du rapport de vraisemblance pour le test de l'hypothèse nulle $H_0 : \beta = 0$ et montrez qu'elle peut s'écrire : $n \log\left(\frac{1}{1-r^2}\right)$ où r^2 est le carré du coefficient de corrélation entre X et y .

3. **Exercice 3** (4 pts.) - Soit $Y = X\beta + u$ avec $u \sim N(0, \sigma^2 I_n)$ où $X_i = [X_{1i} \ X_{2i}]$, $i = 1, \dots, n$ et où I_n est une matrice identité de taille (n, n) . On suppose que la variable X_{2i} est corrélée avec les perturbations u_i et on définit un vecteur de variables instrumentales : $Z_i = [X_{1i} \ Z_{2i}]$. Les instruments Z sont supposés être non corrélés avec u : $E[Z'u] = 0$ et l'estimateur à variables instrumentales $\hat{\beta}_{IV}$ est défini par :

$$\hat{\beta}_{IV} = \left(\sum_{i=1}^n Z_i' X_i \right)^{-1} \sum_{i=1}^n Z_i' Y_i = (Z'X)^{-1} Z'Y$$

Posons

$$\frac{1}{n} Z'X = \Sigma_{ZX}, \quad \frac{1}{n} Z'Z = \Sigma_{ZZ} \text{ et } \frac{1}{n} X'Z = \Sigma_{XZ}$$

Montrez que l'estimateur centré dilaté converge en loi vers la loi normale :

$$\sqrt{n} (\hat{\beta}_{IV} - \beta) \sim N\left(0, \sigma^2 [\Sigma_{XZ} \cdot \Sigma_{ZZ}^{-1} \cdot \Sigma_{ZX}]^{-1}\right)$$

4. **Exercice 4** (8 pts.) - En France, les truffes sont souvent repérées par des collectionneurs qui utilisent des cochons pour flairer les truffes et les “désigner”. Considérons un modèle d’offre et de demande de truffes :

$$\begin{cases} \text{demande :} & q_t = \alpha_1 + \alpha_2 p_t + \alpha_3 ps_t + \alpha_4 di_t + \varepsilon_{d,t} & (\text{eq.1}) \\ \text{offre :} & q_t = \beta_1 + \beta_2 p_t + \beta_3 pf_t + \varepsilon_{o,t} & (\text{eq.2}) \end{cases}$$

Dans l’équation de demande, q_t est la quantité de truffes échangées sur un marché français particulier à l’année t , p est le prix de marché des truffes, ps est le prix de marché d’un substitut aux vraies truffes (un autre champignon beaucoup moins prisé), et di est le revenu mensuel disponible par habitant des résidents locaux. L’équation d’offre contient le prix du marché et la quantité fournie. Elle inclut également pf , le prix d’un facteur de production qui, dans ce cas, est le prix de location horaire des cochons truffiers utilisés dans le processus de recherche. On dispose de 30 observations annuelles et les unités de mesure sont les suivantes : en dollars par once pour p et ps , en onces pour q , en milliers de dollars pour di et en dollars pour pf .¹ Dans ce modèle, nous supposons que p et q sont des variables endogènes. Les variables exogènes sont ps , di et pf .

- On estime la demande de truffes (eq.1) à l’aide de la méthode de Prais-Winsten (MCQG). Commentez les résultats de la Table 1.
- Mais on suspecte l’endogénéité des prix. On ré-estime donc cette équation à l’aide de la méthode des variables instrumentales (doubles moindres carrés (DMC)). Puis on effectue un test d’Hausman pour comparer les modèles MCQG et DMC. Commentez les résultats associés de la Table 2.
- On ré-estime cette équation à l’aide de la méthode généralisée des moments (GMM). On suspecte cependant l’endogénéité de la variable pf . Commentez les résultats de la Table 3 et la différence de Sargan (*C statistics*). Que concluez-vous ?
- On estime l’équation d’offre (eq.2) à l’aide de la méthode généralisée des moments (GMM). On suspecte l’endogénéité de la variable ps_{t-1} . Commentez les résultats de la Table 4 et la différence de Sargan (*C statistics*). Que concluez-vous ?
- On s’intéresse maintenant au modèle à équations simultanées (eq.1 et eq.2)
 - Donnez les conditions d’ordre d’identification pour chacune des 2 équations du modèle.
 - On estime ce modèle à l’aide des triples moindres carrés (TMC). Commentez les résultats de la Table 5.
 - On ré-estime ensuite ce modèle à l’aide des GMM en utilisant une correction HAC pour gérer l’autocorrélation et l’hétéroscédasticité des résidus. Commentez les résultats de la Table 6.
 - A l’aide de la Table 7, calculez et interprétez les élasticités de la demande et de l’offre de truffes par rapport au prix de marché au point moyen de l’échantillon.

Aucun document autorisé.

Calculatrices et tables statistiques autorisées.

1. une once = 28.35 grammes.

```
. prais q p ps di
. estimates store prais_D
```

Prais-Winsten AR(1) regression with iterated estimates

Source	SS	df	MS	Number of obs	=	30
Model	338.194804	3	112.731601	F(3, 26)	=	10.03
Residual	292.191286	26	11.2381264	Prob > F	=	0.0001
				R-squared	=	0.5365
				Adj R-squared	=	0.4830
Total	630.38609	29	21.7374514	Root MSE	=	3.3523

q	Coefficient	Std. err.	t	P> t	[95% conf. interval]
p	.0368402	.0708375	0.52	0.607	-.1087684 .1824488
ps	.7002194	.2165862	3.23	0.003	.25502 1.145419
di	-.2846082	1.05934	-0.27	0.790	-2.462113 1.892897
_cons	1.729868	3.422743	0.51	0.618	-5.30568 8.765417
rho	-.2547271				

Durbin-Watson statistic (original) = 2.365947

Durbin-Watson statistic (transformed) = 1.910130

TABLE 1: Estimation de la demande de truffes (MCQG).

```
. ivregress 2sls q ps di (p = L(0).pf), first
. estimates store iv_D
. hausman iv_D prais_D
```

First-stage regressions

p	Coefficient	Std. err.	t	P> t	[95% conf. interval]
ps	1.708147	.3508806	4.87	0.000	.9869017 2.429393
di	7.602492	1.724336	4.41	0.000	4.058069 11.14691
pf	1.353906	.2985062	4.54	0.000	.7403173 1.967494
_cons	-32.51242	7.984235	-4.07	0.000	-48.92425 -16.10059

Number of obs = 30
F(3, 26) = 69.19
Prob > F = 0.0000
R-squared = 0.8887
Adj R-squared = 0.8758
Root MSE = 6.5975

TABLE 2: Estimation de la demande de truffes (DMC).

```
Instrumental variables 2SLS regression          Number of obs =      30
                                                Wald chi2(3)  =     20.43
                                                Prob > chi2   =     0.0001
                                                R-squared    =      .
                                                Root MSE    =     4.5895
```

q	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
p	-.3744592	.1533755	-2.44	0.015	-.6750696	-.0738487
ps	1.296033	.3306669	3.92	0.000	.6479382	1.944129
di	5.013979	2.125875	2.36	0.018	.8473401	9.180618
_cons	-4.279473	5.161076	-0.83	0.407	-14.395	5.83605

```
Instrumented: p
Instruments: ps di pf
```

---- Coefficients ----				
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	iv_D	prais_D	Difference	Std. err.
p	-.3744592	.0368402	-.4112994	.1360371
ps	1.296033	.7002194	.595814	.249862
di	5.013979	-.2846082	5.298587	1.843134

```
b = Consistent under H0 and Ha; obtained from ivregress.
B = Inconsistent under Ha, efficient under H0; obtained from prais.
Test of H0: Difference in coefficients not systematic
```

```
chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
          = 9.19
Prob > chi2 = 0.0269
```

TABLE 2 (suite) - Estimation de la demande de truffes (DMC).

```
. ivreg2 q ps di (p = L(0/1).pf L.di L.ps), gmm2s robust orthog(pf)
```

```
2-Step GMM estimation
```

```
Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity
```

```
Number of obs =      29
F( 3, 25) =      7.63
Prob > F =      0.0009
Centered R2 =      0.2072
Uncentered R2 =      0.9533
Root MSE =      4.101
```

q	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
p	-.2703387	.1361595	-1.99	0.047	-.5372064	-.0034709
ps	1.2048	.3317696	3.63	0.000	.5545434	1.855056
di	3.566866	1.58554	2.25	0.024	.4592647	6.674467
_cons	-3.616193	4.426976	-0.82	0.414	-12.29291	5.060521

```
Hansen J statistic (overidentification test of all instruments):      3.248
Chi-sq(3) P-val =      0.3549
```

```
-orthog- option:
```

```
Hansen J statistic (eqn. excluding suspect orthog. conditions):      2.784
Chi-sq(2) P-val =      0.2486
```

```
C statistic (exogeneity/orthogonality of suspect instruments):      0.464
Chi-sq(1) P-val =      0.4957
```

```
Instruments tested: pf
```

```
Instrumented: p
Included instruments: ps di
Excluded instruments: pf L.pf L.di L.ps
```

TABLE 3: Estimation de la demande de truffes (GMM).

```
. ivreg2 q pf (p = L(1/2).pf L.di L.ps), gmm2s robust orthog(L.ps)
```

2-Step GMM estimation

Estimates efficient for arbitrary heteroskedasticity
 Statistics robust to heteroskedasticity

```

                                Number of obs =      28
                                F( 2, 25) =      52.33
                                Prob > F      =      0.0000
                                Centered R2    =      0.8884
                                Uncentered R2  =      0.9936
                                Root MSE    =      1.527

Total (centered) SS      = 585.1611793
Total (uncentered) SS  = 10272.78523
Residual SS             = 65.29044221
```

		Robust				
q	Coefficient	std. err.	z	P> z	[95% conf. interval]	
p	.3014396	.0298038	10.11	0.000	.2430252	.3598539
pf	-.9015108	.0853524	-10.56	0.000	-1.068798	-.7342231
_cons	20.2422	1.18002	17.15	0.000	17.9294	22.555

```
Hansen J statistic (overidentification test of all instruments):      3.511
Chi-sq(3) P-val =      0.3194
```

-orthog- option:

```
Hansen J statistic (eqn. excluding suspect orthog. conditions):      3.248
Chi-sq(2) P-val =      0.1971
```

```
C statistic (exogeneity/orthogonality of suspect instruments):      0.263
Chi-sq(1) P-val =      0.6081
```

Instruments tested: L.ps

```
Instrumented:      p
Included instruments: pf
Excluded instruments: L.pf L2.pf L.di L.ps
```

TABLE 4: Estimation de l'offre de truffes (GMM).

```
. reg3 (eq_D: q p ps di ) (eq_0: q p pf ) , inst(L(0/2).pf ps L(0/1).di) 3sls ireg3
```

Three-stage least-squares regression, iterated

Equation	Obs	Params	RMSE	"R-squared"	chi2	P>chi2
eq_D	28	3	4.331559	0.1022	18.86	0.0003
eq_0	28	2	1.459807	0.8980	195.01	0.0000

	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
eq_D						
p	-.3007782	.1443648	-2.08	0.037	-.5837281	-.0178283
ps	1.101252	.3135473	3.51	0.000	.4867102	1.715793
di	4.372164	1.892391	2.31	0.021	.6631463	8.081182
_cons	-2.299883	4.937007	-0.47	0.641	-11.97624	7.376474
eq_0						
p	.3336489	.0244524	13.64	0.000	.285723	.3815748
pf	-.9756035	.0818926	-11.91	0.000	-1.13611	-.8150969
_cons	19.74256	1.362891	14.49	0.000	17.07134	22.41378

Endogenous variables: q p

Exogenous variables: pf L.pf L2.pf ps di L.di

TABLE 5: Estimation de l'offre et de la demande de truffes (TMC).

```

. gmm (eq_D: q - {a0} - {xa: p ps di}) ///
    (eq_0: q - {b0} - {xb: p pf}), ///
    instruments(eq_D: L(1/3).p L(0/1).pf L.di L.ps) ///
    instruments(eq_0: L(1/3).p L(1/2).pf L.di L.ps) ///
    winitial(unadjusted, independent) wmatrix(hac nwest opt) twostep

GMM estimation

Number of parameters = 7
Number of moments = 16
Initial weight matrix: Unadjusted          Number of obs = 27
GMM weight matrix:    HAC Bartlett 11
                    (lags chosen by Newey-West)

```

		HAC		z	P> z	[95% conf. interval]	
		Coefficient	std. err.				
a0	_cons	7.090662	1.862173	3.81	0.000	3.44087	10.74045
xa	p	-.2109091	.0602037	-3.50	0.000	-.3289061	-.092912
	ps	.630979	.1544496	4.09	0.000	.3282634	.9336946
	di	3.089707	1.009984	3.06	0.002	1.110175	5.069238
b0	_cons	20.2918	.4615654	43.96	0.000	19.38715	21.19645
xb	p	.3292354	.0102784	32.03	0.000	.3090902	.3493806
	pf	-.9896121	.0405087	-24.43	0.000	-1.069008	-.9102165

HAC standard errors based on Bartlett kernel with 11 lags.
Lags chosen by Newey-West method.
Instruments for equation eq_D: L.p L2.p L3.p pf L.pf L.di L.ps _cons
Instruments for equation eq_0: L.p L2.p L3.p L.pf L2.pf L.di L.ps _cons

```
. estat overid
```

Test of overidentifying restriction:

Hansen's J chi2(9) = 2.44014 (p = 0.9825)

TABLE 6: Estimation de l'offre et de la demande de truffes (GMM - HAC).

```
. sum q p ps di pf
```

Variable	Obs	Mean	Std. dev.	Min	Max
q	30	18.45833	4.613088	6.37	26.27
p	30	62.724	18.72346	29.64	105.45
ps	30	22.022	4.077237	15.21	28.98
di	30	3.526967	1.040803	1.525	5.125
pf	30	22.75333	5.329654	10.52	34.01

TABLE 7: Statistiques descriptives.