Université d'Orléans - Master ESA 2
Macro-Econométrie

Examen Terminal Mars 2005. C. Hurlin
Tout Document Autorisé

Exercice 1 : Estimation d'une Courbe de Phillips Hybride par GMM

On considère le modèle suivant de détermination de l'inflation i_t :

$$i_t = \rho \, i_{t-1} + \beta \, z_t + \varepsilon_t \tag{1}$$

où z_t désigne une variable explicative non précisée (généralement l'output gap) et où le paramètre $|\rho| < 1$. On suppose que les résidus satisfont les deux conditions suivantes :

$$E\left(\varepsilon_t z_t\right) = 0 \qquad E\left(\varepsilon_t i_{t-1}\right) = 0 \tag{2}$$

Partie I : Un peu d'exercice (6 points)

*Question 1 (*1 point*)* : (i) Ecrivez précisément les deux conditions d'orthogonalité de ce modèle en fonction des deux paramètres ρ et β.
(ii) Construisez leur contre partie empirique, notée $g\left(Y_T, \theta\right)$ où $\theta = (\rho, \beta)'$ et Y_t désigne la matrice des observations de z_t et de i_{t-1} sur les T périodes d'observation.

Question 2 (1 point) : Précisez la nature de ce système (sur-identifié ou juste identifié) et écrivez le programme permettant de définir les estimateurs GMM $\widehat{\rho}$ et $\widehat{\beta}$ des paramètres ρ et β. A quoi correspondent ces estimateurs ?

Question 3 (1 point) : Ecrivez le même programme qu'à la question 2 en utilisant une méthode de *continous updating GMM*. Est ce la réalisation des estimateurs $\widehat{\theta} = \left(\widehat{\rho}, \widehat{\beta}\right)'$ sera différente de celle obtenue à la question précédente ?

Question 4 (2 point) : On suppose que les résidus ε_t sont *i.i.d.* $\left(0, \sigma_\varepsilon^2\right)$. Montrez que l'expression de la matrice de variance covariance **théorique** du **biais de l'estimateur GMM** est tel que :

$$\sqrt{T}\left(\widehat{\theta}_T - \theta_0\right) \underset{T\to\infty}{\overset{L}{\longrightarrow}} N\left(0, V\right) \tag{3}$$

avec

$$V = \left\{ \left(\text{plim} - \frac{1}{T} \begin{bmatrix} \sum_{t=1}^T i_{t-1}^2 & \sum_{t=1}^T i_{t-1} z_t \\ \sum_{t=1}^T i_{t-1} z_t & \sum_{t=1}^T z_t^2 \end{bmatrix} \right) \times \left(\sigma_\varepsilon^2 E \begin{bmatrix} i_{t-1}^2 & i_{t-1} z_t \\ i_{t-1} z_t & z_t^2 \end{bmatrix} \right)^{-1} \right.$$
$$\left. \times \left(\text{plim} - \frac{1}{T} \begin{bmatrix} \sum_{t=1}^T i_{t-1}^2 & \sum_{t=1}^T i_{t-1} z_t \\ \sum_{t=1}^T i_{t-1} z_t & \sum_{t=1}^T z_t^2 \end{bmatrix} \right) \right\}^{-1}$$

Question 5 (1 point) : En déduire que l'**estimateur** de la matrice de variance covariance **de l'estimateur GMM** correspond à l'estimateur de la matrice de variance covariance des estimateurs MCO :

$$\widehat{V} = \sigma_\varepsilon^2 \begin{bmatrix} \sum_{t=1}^T i_{t-1}^2 & \sum_{t=1}^T i_{t-1} z_t \\ \sum_{t=1}^T i_{t-1} z_t & \sum_{t=1}^T z_t^2 \end{bmatrix}^{-1} \tag{4}$$

Partie II : Estimation d'une Courbe de Phillips Hybride par GMM

On considère à présent un modèle de la forme :

$$i_t = \alpha E_t\left(i_{t+1}\right) + (1-\alpha)\, i_{t-1} + \beta\, z_t + \varepsilon_t \tag{5}$$

avec $0.5 < \alpha \leq 1$ et où $E_t\left(i_{t+1}\right) = E\left(i_{t+1} \middle| \Omega_t\right)$ où Ω_t désigne l'ensemble d'information disponible à la date t. Cette forme d'équation dans laquelle le niveau d'inflation dépend de sa valeur passée, d'un ensemble de variables explicatives mais aussi de son niveau anticipé correspond aux formulations usuelles d'une courbe de Phillips dite hybride introduite notamment par Fuhrer et Moore (1995), Gali et Gertler (1999) etc. Leur idée consiste à introduire de la persistance (via le terme retardé) dans des formulations purement forward de la courbe de Phillips proposées notamment par Taylor (1980). Nous ne commenterons pas ici les différentes implications de cette spécification de la courbe de Phillips, mais simplement les méthodes d'estimation des paramètres α, β et ρ. Deux grandes méthodes d'estimation sont utilisées dans ce cas :

- Maximum de vraisemblance à information complète (FIML
- Des méthodes de moment et en particulier les GMM

C'est cette dernière méthode que nous allons ici étuider.

*Question 1 (*1 point*)* : (i) Ecrivez précisément les conditions d'orthogonalité de ce modèle en fonction des deux paramètres α et β en utilisant comme instruments les variables i_{t-1} et z_t et z_{t-1} sous la forme

$$E\left[h\left(\theta_0, w_t\right)\right] = 0 \tag{6}$$

(ii) Construisez leur contre partie empirique, notée $g\left(Y_T, \theta\right)$ où $\theta = (\alpha, \beta)'$ et Y_t désigne la matrice des observations de z_t et de i_{t-1} sur les T périodes d'observation.

Question 2 (1 point) : Précisez la nature de ce système (sur-identifié ou juste identifié) et écrivez le programme permettant de définir les estimateurs GMM $\widehat{\alpha}$ et $\widehat{\beta}$ des paramètres α et β.

Question 3 (2 points) : Donnez l'expression de la matrice de poids optimale en fonction des paramètres α et β. On suppose pour cela que le processus $h\left(\theta_0, w_t\right)$ est stationnaire et que par conséquent la matrice de variance covariance de long terme des résidus des conditions d'orthogonalité vérifie :

$$\underset{(r,r)}{S} = \sum_{v=-\infty}^{\infty} E\left[h\left(\theta_0, w_t\right) h\left(\theta_0, w_{t-v}\right)'\right] \tag{7}$$

Que devient cette matrice lorsque les résidus ε_t ne sont pas auto-corrélés ?

Question 4 (2 points) : On suppose que les résidus ε_t sont auto-corrélés. On définit $\widehat{\Gamma}_v$ comme l'estimateur de la matrice d'auto-covariance d'ordre v du processus $h\left(\theta_0, w_t\right)$.

$$\begin{aligned}
\widehat{\Gamma}_v &= \frac{1}{T}\sum_{t=v+1}^{T}\left\{h\left(\widehat{\theta}, w_t\right) h\left(\widehat{\theta}, w_{t-v}\right)'\right\} \\
&= \frac{1}{T}\sum_{t=v+1}^{T}\left[\varepsilon_t\varepsilon_{t-v}\begin{pmatrix} i_{t-1}i_{t-v-1} & i_{t-1}z_{t-v} & i_{t-1}z_{t-v-1} \\ z_t i_{t-1-v} & z_t z_{t-v} & z_t z_{t-v-1} \\ z_{t-1}i_{t-1-v} & z_{t-1}z_{t-v} & z_{t-1}z_{t-v-1}\end{pmatrix}\right]
\end{aligned} \tag{8}$$

avec $\varepsilon_{t-v} = i_{t-v} - \alpha\, i_{t-v+1} - (1-\alpha)\, i_{t-v-1} - \beta\, z_{t-v}$. Donnez l'expression **des estimateurs de Newey West (1987) et de Gallant (1987)** de la matrice de poids optimale de l'estimateur GMM pour un paramètre de troncature $q = 3$.

Question 5 (3 points) : Ecrivez le programme SAS permettant d'estimer une courbe de Phillips Hybride :

Par GMM itératifs (Hansen et Singleton, 1992)

En utilisant un estimateur kernel *à* la Andrews (1991) de la matrice de poids optimale.

En utilisant une correction de petit échantillon (Gallant, 1987) pour l'estimateur de la matrice de poids optimale

En spécifiant que la variable z_t est exogène

En imposant les contraintes $0.5 < \alpha \leq 1$

Exercice 2 : Hansen et Singleton (1982)

On considère les résultats d'estimation sous SAS du modèle de Hansen- Singleton (1982). Dans le cas d'une fonction d'utilité de type CRAA, la relation estimée est de la forme :

$$1 - \beta E_t \left[(1 + r_{t+1}) \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \right] = 0 \tag{9}$$

où γ désigne le coefficient d'aversion relative pour le risque et β le facteur d'escompte psychologique. On considère dans cette application les données de Ferson et Harvey (1992). Les auteurs considèrent des données trimestrielles pour les Etats Unis allant du deuxième trimestre 1947 (codé 1947.6) au quatrième trimestre de 1987 (1987.12). Comme mesure de la consommation réelle, les auteurs utilisent la consommation de bien non durables corrigée des variations saisonnières rapportée à un déflateur de la consommation en données CVS. La croissance de la consommation représentée par le ratio $(C_{t+1} - C_t) / C_t$ est désigné par la variable CONRAT. On considère ici un seul rendement d'actif exprimé sous la forme de rendements, à savoir le rendement réel sur les obligations émises par le gouvernement. Variable (variable GB).

Commentez ces résultats SAS d'estimation des paramètres β et γ. Est ce que l'utilisateur aurait pu dans ce cas utiliser une procédure d'estimation plus simple ?

```
                    The MODEL Procedure

                       Model Summary

                Model Variables          2
                Endogenous               1
                Exogenous                1
                Parameters               2
                Equations                1
                Number of Statements     7
                Program Lag Length       3


           Model Variables  conrat gb
         Parameters(Value)  beta(0.1) alpha(0.1)
                 Equations  h1

                The Equation to Estimate is

                       h1 =  F(beta, alpha)
               Instruments   1 lc1

              The estimation lag length    3

NOTE: At ITGMM Iteration 0 convergence assumed because
OBJECTIVE=7.971944E-17 is almost zero (<1E-12).
```

The MODEL Procedure
ITGMM Estimation Summary

Data Set Options

DATA= HARVEY

Minimization Summary

Parameters Estimated	2
Kernel Used	PARZEN
l(n)	2.759459
Method	Gauss
Iterations	0

Final Convergence Criteria

R	1
PPC(alpha)	1.638E-8
RPC(alpha)	1.638E-8
Object	.
Trace(S)	0.007321
Objective Value	7.97E-17
S	0

Observations Processed

Read	163
Solved	160
First	4
Last	163
Lagged	3

Exercice 3 : Hansen et Singleton (1982), suite

Commentez les résultats empiriques de l'article original d'Hansen et Singleton. Le début de la section débute en bas de la page 1280.

The MODEL Procedure

Nonlinear ITGMM Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square
h1	2	158	1.1567	0.00732	0.0856	

Nonlinear ITGMM Parameter Estimates

Parameter	Estimation	Approx Std Err	t Value	Approx Pr > |t|
beta	0.971194	0.0345	28.16	<.0001
alpha	-8.41505	10.4909	-0.80	0.4237

Number of Observations	
Used	160
Missing	0

Statistics for System	
Objective	7.972E-17
Objective*N	1.276E-14

(4.4), gives the restrictions

$$\log\beta + \frac{\alpha^2\sigma_{kk} + \sigma_{jj} + 2\alpha\sigma_{kj}}{2} + \alpha\mu_k + \mu_j = 0, \tag{4.5}$$
$$\left[\frac{\Psi_j(L)}{L}\right]_+ = -\alpha\left[\frac{\Psi_k(L)}{L}\right]_+,$$

for $j = 1, 2, \ldots, m$. We can estimate the parameters of Σ, Ψ, μ, α, and β by imposing the restrictions (4.5), using maximum likelihood with a Gaussian density function, and employing observations on $X_1, X_2, \ldots, X_{T+1}$ (see, e.g., Hansen and Singleton **[14]**).

The restrictions given in (4.5) together with relation (4.2) imply a logarithmic form of relation (4.4), namely,

$$E\left[\log\beta + \alpha\log x_{kt+1} + \log x_{jt+1} \mid X_t, X_{t-1}, \ldots\right] = \frac{-\alpha^2\sigma_{kk} - \sigma_{jj} - 2\alpha\sigma_{kj}}{2}. \tag{4.6}$$

By the law of iterated expectations, it follows that the random variable

$$U_{jt+1} = \log\beta + \alpha X_{kt+1} + X_{jt+1} + \frac{\left(\alpha^2\sigma_{kk} + \sigma_{jj} + 2\alpha\sigma_{kj}\right)}{2}$$

satisfies the orthogonality conditions

$$E\left[U_{jt+1}\right] = 0; \qquad E\left[U_{jt+1}X_{t-s}\right] = 0; \tag{4.7}$$

for $j = 1, 2, \ldots, m$ and $s \geq 0$.

From the first-order conditions of the likelihood function, it can be seen that the method of maximum likelihood implicitly uses the logarithmic orthogonality conditions (4.7). The validity of these orthogonality conditions is crucially dependent on x being lognormally distributed. For other distributional assumptions, the logarithmic form of the orthogonality conditions (4.7) will generally not hold. Consequently, the maximum likelihood estimators of α and β obtained from (4.7) under the assumption of lognormality will generally not be consistent if this distributional assumption is incorrect. In contrast, the procedures which we proposed in Section 3 do not require that the distribution of x be specified *a priori* and, in particular, they do not require that the logarithmic form of the orthogonality conditions hold. Instead, we work directly with the orthogonality conditions implied by (4.1).

5. EMPIRICAL RESULTS

To illustrate the use of the generalized instrumental variables estimator, we estimated the parameters of preferences, α and β, for the model of stock prices discussed in Section 4. Two different measures of consumption were considered:

nondurables plus services (NDS) and nondurables (ND).[11] The monthly, seasonally adjusted observations on aggregate real consumption of nondurables and services were obtained from the Federal Reserve Board. Real per capita consumption series were constructed by dividing each observation of these series by the corresponding observation on population, published by the Bureau of Census. Each measure of consumption was paired with three sets of stock returns: the equally-weighted average return on all stocks listed on the New York Stock Exchange (EWR), the value-weighted average of returns on the New York Stock Exchange (VWR), and equally-weighted average returns on the stocks of three two-digit SEC industries. The industries chosen were chemicals (SEC code 28), transportation and equipment (SEC code 37), and other retail trade (SEC codes 50–52 and 54–59). The aggregate return data were obtained from the CRSP tapes and the industry return data were obtained from Stambaugh [**31**]. Nominal returns were converted to real returns, which appear in (4.1), by dividing by the implicit deflator associated with the measure of consumption.

Following the notation adopted in Section 4, we let

$$x'_{t+1} = \left[\frac{P_{1t+1} + D_{1t+1}}{P_{1t}} \quad \frac{C_{t+1}}{C_t} \right]$$

and

$$h(x_{t+1}, b_0) = \beta (x_{2t+1})^{\alpha} x_{1t+1} - 1,$$

where $b'_0 = (\alpha, \beta)$. The vector of instruments z_t was formed using lagged values of x_{t+1}. For this specification of the process $\{(x_{t+1}, z_t);\ t = 1, 2, \dots \}$, the stationarity assumption accommodates certain types of real growth in consumption. The number of lagged values of x_{t+1} included in z_t, NLAG, was chosen to be 1, 2, 4, or 6. As NLAG is increased, more orthogonality conditions are employed in the estimation. Furthermore, the asymptotic covariance matrix becomes smaller, and the number of overidentifying restrictions being tested increases.[12]

Table I displays the parameter estimates obtained using the aggregate return series for the period February, 1959 through December, 1978.

[11] Using a separation argument like that for labor supply discussed in note 4, it is possible to argue that the restrictions hold for a measure of a subset of aggregate consumption. For instance, suppose that C_{1t} and C_{2t} are two different components of consumption at time period t and that the function U is given by

$$U(C_{1t}, C_{2t}) = \frac{(C_{1t})^{\gamma}}{\gamma} + U_2(C_{2t}).$$

In this case, the restrictions we test are appropriate when C_1 is used as the measure of consumption. The two choices for C_1 considered correspond to two potentially different assumptions about the separability of U.

[12] To be more precise, the asymptotic covariance matrix will not increase as more orthogonality conditions are used. Using more orthogonality conditions may, at some point, lead to estimators with less desirable small sample properties.

TABLE I

INSTRUMENTAL VARIABLE ESTIMATES FOR THE PERIOD 1959:2–1978:12

Cons	Return	NLAG	$\hat{\alpha}$	$\widehat{SE}(\hat{\alpha})$	$\hat{\beta}$	$\widehat{SE}(\hat{\beta})$	χ^2	DF	Prob
NDS	EWR	1	−.9457	.3355	.9931	.0031	4.9994	1	.9746
NDS	EWR	2	−.9281	.2729	.9929	.0031	7.5530	3	.9438
NDS	EWR	4	−.7895	.2527	.9925	.0031	9.1429	7	.7574
NDS	EWR	6	−.8927	.2138	.9934	.0030	15.726	11	.8484
NDS	VWR	1	−.9001	.3130	.9979	.0025	1.1547	1	.7174
NDS	VWR	2	−.8133	.2298	.9981	.0025	3.2654	3	.6475
NDS	VWR	4	−.6795	.1855	.9973	.0024	6.3527	7	.5008
NDS	VWR	6	−.7958	.1763	.9980	.0023	14.179	11	.7767
ND	EWR	1	−.9737	.1245	.9922	.0031	5.9697	1	.9854
ND	EWR	2	−.9664	.1074	.9919	.0031	8.9016	3	.9694
ND	EWR	4	−.9046	.0926	.9918	.0031	11.084	7	.8650
ND	EWR	6	−.9466	.0793	.9422	.0030	15.663	11	.8459
ND	VWR	1	−.8985	.1057	.9971	.0025	1.5415	1	.8756
ND	VWR	2	−.8757	.0856	.9974	.0025	3.2654	3	.6475
ND	VWR	4	−.8174	.0742	.9967	.0024	7.8776	7	.5008
ND	VWR	6	−.8514	.0629	.9973	.0024	14.938	11	.8147

The estimates of α range from −.95 to −.68 when ND is used as the measure of consumption, and from −.97 to −.82 when ND is used as the measure of consumption. The estimated standard errors for α, $\widehat{SE}(\hat{\alpha})$, are smaller when consumption is measured as ND than when consumption is measured as NDS. As expected, all of the estimates of β exceed .99 but are less than unity. The chi-square tests are also displayed in Table I, where the number of overidentifying restrictions is indicated by DF and Prob is the probability that a χ^2(DF) random variate is less than the computed value of the test statistic under the hypothesis that the restrictions (3.1) are satisfied. These tests provide greater evidence against the model when EWR is included as the return, and when the instrument vector is formed from a small number of lagged values of x.

For comparison, we present some results in Table II from estimating α and β using the method of maximum likelihood under the assumption that x is lognormally distributed. They were obtained using the procedure described in Section 4 assuming that log x has a sixth-order vector autoregressive representation. The corresponding estimates of α and β from the two methods of estimation are similar. However, the estimated standard errors of α and β from the instrumental variables procedure are smaller than the corresponding standard errors from the maximum likelihood procedure.[13] Three possible explanations for this result are that the asymptotic standard errors are being estimated imprecisely, the economic model of stock returns is misspecified, or the auxiliary assumptions underlying the maximum likelihood procedure are incorrect. The maximum likelihood procedure assumes that x is lognormally distributed and that the lag length specification of the vector autoregression is correct. Since the

[13] Analytical differentiation, as opposed to numerical differentiation, was used to calculate the standard errors.

TABLE II
MAXIMUM LIKELIHOOD ESTIMATES FOR THE PERIOD 1959:2–1978:12

	Nondurables Plus Services		Nondurables	
	Equally Weighted Returns	Value Weighted Returns	Equally Weighted Returns	Value Weighted Returns
$\hat{\alpha}$	−.5194 (.4607)	−.9349 (.3341)	−.7643 (.2203)	−.9327 (.1579)
$\hat{\beta}$	.9957 (.0037)	.9995 (.0028)	.9957 (.0036)	.9988 (.0028)
χ^2	12.5854	17.8716	14.4456	18.8846
DF	11	11	11	11
Prob.	.6787	.9154	.7907	.9368

sum of lognormally distributed random variables is not lognormally distributed, it seems implausible that both (and perhaps either) the value weighted and (or) the equally weighted stock return indices are lognormally distributed. Also, the lag length specification does not emerge from any theoretical consideration. The generalized instrumental variables procedure does not require that either one of these auxiliary assumptions be satisfied.[14] On the other hand, the maximum likelihood estimates provide a more complete characterization of the stochastic process x when the lognormality and the lag length specifications are correct.

Table II also presents the likelihood ratio tests of the restrictions implied by the lognormal version of the model, where the unrestricted model is an unrestricted sixth-order vector autoregression. In contrast to the instrumental variables results, the maximum likelihood results provide more evidence against the restrictions when VWR is used than when EWR is used. For a more complete description of the maximum likelihood estimation and a more comprehensive set of results, see Hansen and Singleton [**14**].

If the Euler equation (4.1) holds for a given measure of consumption and all of the stocks listed on the NYSE, then versions of (4.1) must also hold simultaneously for the equally- and value-weighted aggregate returns. In Table III we present the results from estimating α and β using the orthogonality conditions implied by the respective versions of (4.1) for EWR and VWR, with the z vector formed using lagged values of EWR, VWR, and the consumption ratio. Note that these estimates of α are smaller than the corresponding estimates in Table I. Also, the estimated standard errors for $\hat{\alpha}$ are smaller than the corresponding

[14]Another attractive feature of the generalized instrumental variables procedure is that it requires a numerical search over a smaller parameter space than is required from the maximum likelihood procedure.