Kinshasa, Mars 2018

Manuel d'Econométrie
(Inspiré de Regis bourbonnais 2009, 7è édition)

Économétrie Appliquée :
Recueil des cas pratiques sur EViews

Par

Jonas KIBALA KUMA

(DEA-PTC Economie/Unikin en cours)

--

Centre de Recherches Economiques et Quantitatives
(CREQ)

***

Mars 2018

Kinshasa, Mars 2018

Manuel d'Econométrie
(Inspiré de Regis bourbonnais 2009, 7è édition)

Économétrie Appliquée :
Recueil des cas pratiques sur EViews

Par

Jonas KIBALA KUMA

(DEA-PTC Economie/Unikin en cours)
--
Centre de Recherches Economiques et Quantitatives
(CREQ)

***

Mars 2018

– Note –

Ce manuel d'Econométrie consacré aux cas pratiques sur Logiciels (Eviews), qui s'inscrit dans le cadre de nos travaux/recherches sur « l'Econométrie appliquée » que nous nous efforçons de mettre à la disposition des chercheurs africains/congolais surtout et de tout bord, s'inspire essentiellement de l'ouvrage/manuel de Regis Bourbonnais (2009, 7è édition) intitulé « Econométrie : Manuel et Exercices corrigés ». En plus de notre propre expérience (leçons tirées des enseignements dispensés), un autre manuel du même auteur (2009) intitulé « Logiciel Eviews » – où Bourbonnais nous fait découvrir le logiciel Eviews hexa tétra – nous a servi de référence. La particularité de ce manuel, c'est qu'il initie l'utilisateur à la manipulation du logiciel Eviews avec les exercices à l'appui. L'objectif est donc qu'en fin de compte le lecteur comprenne l'économétrie appliquée en reproduisant les outputs obtenus par Bourbonnais, ce qui n'est pas tâche facile à mon avis. Toutefois, nous sollicitons beaucoup plus d'attention de nos lecteurs quant à certains petits détails pourtant important et souhaitons à tous une bonne pratique de l'Econométrie sur Eviews.

Merci au Professeur Regis Bourbonnais (Université de Paris-Dauphine) pour ses travaux de recherche qui nous inspirent.

Kinshasa, Mars 2018

I. Calcul d'un coefficient de corrélation (avec EViews 6) : Feuille « E1 »

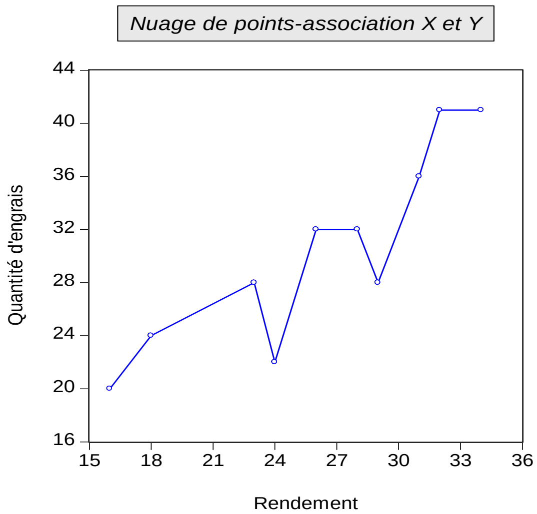
- **TD** : Sur base de données « **E1** » (*x=rendement de maïs en quintal d'une parcelle de terre et y=quantité d'engrais en kilo*), il est demandé de :
  - *Produire et commenter le nuage de points (association « x et y ») ;*
  - *Obtenir le coefficient de corrélation simple « r » et tester « r=0 » au seuil $\alpha = 0.05$.*

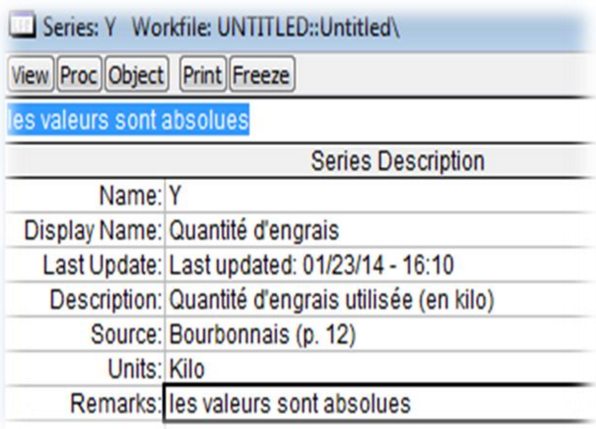
- ***Solution*** :

  a) Nuage des points

Les commandes EViews sont :

create u 1 10
data X Y
copier et coller les données (« Ctr+C » sur Excel et « Ctrl+V » sur EViews)
dans Workfile, doublecliquer sur « Y » ou « X » et suivre « View/label » pour décrire les variables « X et Y » (Cfr figure à gauche ci-dessous)
scat X Y
dans l'output du graphique, cliquer sur « Options→Line et Symbol » pour relier les points du nuage.





Commentaires : présomption d'une corrélation positive entre X et Y.

b) teste « r=0 » au seuil $\alpha = 0.05$ (Cfr « t de student »)

$$t_c = \frac{|\rho_{x,y}|}{\sqrt{\frac{(1-\rho_{x,y}^2)}{n-2}}} \sim t_t = t_{n-2}^{\alpha/2}$$

Avec :

- t_c : t de student calculé ;
- t_t : la valeur du t de student lue sur la table (à « n-2 » degrés de liberté et au seuil de « $\alpha/2$ ») ;

- « $\rho_{x,y}$ » : le coefficient de corrélation simple empirique (c'est une approximation du vrai « $r_{x,y}$ ») qui est obtenu sur *EViews* en tapant : *cor X Y*

	X	Y
X	1	0.8929014206598092
Y	0.8929014206598092	1

Les hypothèses à tester sont :

H0 : $\rho_{x,y} = 0$: absence de lien entre « X et Y » ($t_c < t_t$) ;

H1 : $\rho_{x,y} \neq 0$: Existence de lien entre « X et Y » ($t_c > t_t$) .

Vous obtiendrez : $t_c = \frac{0.89}{\sqrt{\frac{(1-0.79)}{10-2}}} = \frac{0.89}{0.1620} = 5.49$ et $t_t = t_8^{0.025} = 2.306.$

Conclusion : le lien positif entre X et Y est significatif (le coefficient de corrélation est significativement $\neq 0$, car : $t_c > t_t$).

II. Génération d'une consommation aléatoire (avec EViews 6) : Feuille « E2 »

- ***TD*** : Sur base de données « **E2** », il est demandé de :
  - *Calculer la consommation théorique sur 10 ans, sachant que : la propension marginale à consommer est « 0.8 » et la consommation autonome/incompressible est « 1.000 » ;*
  - *Soit l'erreur d'observation « $e_t \sim N(0,20000)$ » : générer la variable aléatoire et calculer une consommation observée en fonction de cette erreur.*

- ***Solution*** :

a) <u>Calcul de la consommation théorique</u>

Sur EViews, pour générer la consommation théorique, taper :

```
create a 1992 2001
data REVENU
genr CONS1=1000+0.8*REVENU
```

b) Variable aléatoire générée (E) et consommation observée calculée (CONS2) : Cfr tableau ci-dessous :

```
genr CONS1=1000+0.8*REVENU
genr E=nrnd*sqr(20000)
genr CONS2=CONS1+E
```

obs	*CONS1*	*REVENU*	*E*	*CONS2*
1992	7400.000	8000.000	43.72313	7443.723
1993	8200.000	9000.000	-149.3332	8050.667
1994	8600.000	9500.000	7.790322	8607.790
1995	8600.000	9500.000	-176.7368	8423.263
1996	8840.000	9800.000	-21.73846	8818.262
1997	9800.000	11000.00	273.4824	10073.48
1998	10600.00	12000.00	-203.8789	10396.12
1999	11400.00	13000.00	122.3745	11522.37
2000	13000.00	15000.00	170.7927	13170.79
2001	13800.00	16000.00	-212.9944	13587.01

III. Estimation des coefficients de régression (avec EViews 6) : Feuille « E3 »

- **TD** : Sur base de données « **E3** », considérant que « $Y_t = a_0 + a_1 X + e_t$ », il est demandé de calculer les estimateurs « $\hat{a}_0$ et $\hat{a}_1$ ».
- **Solution** : Rappelons ce qui suit (en bas : les résultats obtenus sur EViews 6 et les commandes utilisées) :

$$\hat{a}_1 = \frac{\sum_{t=1}^{T}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{t=1}^{T}(x_i - \bar{X})^2} = 0.78$$

$$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X} = 1176.09$$

Dependent Variable: Y
Method: Least Squares
Date: 01/26/14 Time: 13:29
Sample: 1992 2001
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1176.090	207.3921	5.670852	0.0005
X	0.780983	0.017939	43.53518	0.0000

R-squared	0.995797	Mean dependent var	9985.575
Adjusted R-squared	0.995271	S.D. dependent var	2089.553
S.E. of regression	143.6878	Akaike info criterion	12.95002
Sum squared resid	165169.4	Schwarz criterion	13.01054
Log likelihood	-62.75009	Hannan-Quinn criter.	12.88363
F-statistic	1895.312	Durbin-Watson stat	1.881665
Prob(F-statistic)	0.000000		

Commandes EViews :

create a 1992 2001
data Y X
copier et coller les données
ls Y c X

Présentation « standard » des résultats d'estimation :

$\hat{y}_t = 1176{,}09 + 0.78 * X$
$\quad\quad(5.67) \quad (43.54)$

$R^2 = 99.58$

$n = 10$

$(.) = t\ de\ student$

IV. Prévision dans un modèle de régression linéaire simple (avec EViews 6) : Feuille « E3 »

- **TD** : Sur base de données « **E3** », *Cfr output précédent*, il est demandé de :
  - *Calculer le coefficient de détermination et tester la significativité globale des paramètres (test de Fisher) ;*
  - *Si « X (revenu) » augmente de 8%, comment réagirait « Y (Consommation) » ?*
  - *Quels seraient les niveaux de consommation attendus pour des revenus projetés à 16 800 et 17 000 USD, pour les années 2002 et 2003 respectivement. Calculer aussi l'intervalle de prédiction au seuil de 95%.*

- **Solution** :

  a) *Calcul du coefficient de détermination et test de la significativité globale des paramètres (test de Fisher)*

Rappelons ce qui suit :

$$R^2 = \frac{SCE}{SCT} = \frac{\sum(\hat{y}_t - \bar{Y})^2}{\sum(y_t - \bar{Y})^2} = 1 - \frac{\sum e_t^2}{\sum(y_t - \bar{Y})^2} = 0.9958$$

$$F_c = \frac{R^2}{(1-R^2)/(T-2)} = \frac{r^2}{(1-r^2)/(T-2)} = (t_c)^2 = 1895.31$$

Sachant que $F_t = F_{1,8}^{0.05} = 5.32$, l'on constate que « $F_c > F_t$ » (*les paramètres sont globalement significatifs*).

b) *Réaction de Y (Consommation) à l'augmentation de X (revenu) pour 8%*

En effet : $\Delta\hat{y}_t = \hat{a}_1\Delta x_t \rightarrow \Delta\hat{y}_t = 0.78\Delta x_t \rightarrow\rightarrow \Delta\hat{y}_t = 0.78 \times 0.08 = 0.0624$;

Réponse : Si le revenu augmente de 8%, la consommation réagira dans le même sens pour 6.24%.

c) *Niveaux de consommation attendus pour des revenus projetés à 16 800 et 17 000 USD, pour les années 2002 et 2003 respectivement, et Calcul de l'intervalle de prédiction au seuil de 95%.*

Prévision :

$$\hat{y}_{2002} = 1176.09 + 0.78\, x_{2002} = 1176.09 + 0.78 * 16\,800 = 14\,280.09$$

$$\hat{y}_{2003} = 1176.09 + 0.78\, x_{2003} = 1176.09 + 0.78 * 17\,000 = 14\,436.09$$

Intervalle de prévision : en général :

$$y_{t+1} = \hat{y}_{t+1} \pm t_{t-2}^{\alpha/2}\hat{\sigma}_\varepsilon \sqrt{\frac{1}{t} + \frac{(x_{t+1}-\bar{X})^2}{\sum(x_{t+1}-\bar{X})} + 1}$$

- *Pour le 1er cas :*

$$y_{2002} = \hat{y}_{2002} \pm t_{t-2}^{\alpha/2}\hat{\sigma}_\varepsilon \sqrt{\frac{1}{t} + \frac{(x_{2002}-\bar{X})^2}{\sum(x_{t+1}-\bar{X})} + 1}$$

Sachant que (*Cfr output*) : t=10 ; $\hat{\sigma}_\varepsilon$=143.69 ; $\sum(x_i-\bar{X})^2$=64 156 000 ; $\bar{X}$=11280 ; $t_{t-2}^{\alpha/2}$=2.306 ; x_{2002}=16800, alors :

$$y_{2002}=14280{,}08 \pm 2{,}306 \times 180{,}32$$

Et l'intervalle de confiance (*la réalisation a 95% de chance d'apparaître dans cet intervalle, surtout au centre, la distribution étant normale*) :

$$IC = [13\,864.24\,;\,14\,695.91]$$

- *Pour le 2ème cas :*

$$y_{2003} = \hat{y}_{2003} \pm t_{t-2}^{\alpha/2}\hat{\sigma}_\varepsilon \sqrt{\frac{1}{t} + \frac{(x_{2003}-\bar{X})^2}{\sum(x_{t+1}-\bar{X})} + 1}$$

$$y_{2003}=14436{,}08 \pm 2{,}306 \times 182{,}32$$

$$IC = [14\,015.65\,;\,14\,856.51]$$

V. Tests à partir de l'analyse de la variance (avec EViews 6) : Feuille « E4 »

- ***TD*** : Sur base de données « **E4** », il est demandé de vérifier ce qui suit :
  - *La significativité des variables « x2 et x3 » ajoutées par rapport au modèle où seule « x1 » est explicative ;*
  - *Le modèle avec trois régresseurs (x1, x2 et x3) est-il stable sur tout l'échantillon ? Peut-on effectuer deux estimations : de 1 à 7 et de 8 à 14 ?*
  - $b_1 = 1$ *et* $b_2 = b_3$ *?*
- ***Solution :***

a) Vérification de la significativité des variables « x2 et x3 » ajoutées par rapport au modèle où seul « x1 » est explicative (bonté du modèle à trois variables)

Dependent Variable: Y
Method: Least Squares
Date: 01/26/14 Time: 15:53
Sample: 1 14
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	32.89132	11.66331	2.820068	0.0182
X1	0.801901	0.298436	2.687012	0.0228
X2	-0.381362	0.156581	-2.435564	0.0351
X3	-0.037132	0.052023	-0.713768	0.4917

R-squared	0.702687	Mean dependent var	17.71429
Adjusted R-squared	0.613493	S.D. dependent var	4.177385
S.E. of regression	2.597069	Akaike info criterion	4.981600
Sum squared resid	67.44767	Schwarz criterion	5.164188
Log likelihood	-30.87120	Hannan-Quinn criter.	4.964698
F-statistic	7.878181	Durbin-Watson stat	3.186886
Prob(F-statistic)	0.005452		

Dependent Variable: Y
Method: Least Squares
Date: 01/26/14 Time: 16:11
Sample: 1 14
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.57116	1.889095	6.125240	0.0001
X1	1.011809	0.281386	3.595797	0.0037

R-squared	0.518647	Mean dependent var	17.71429
Adjusted R-squared	0.478535	S.D. dependent var	4.177385
S.E. of regression	3.016597	Akaike info criterion	5.177699
Sum squared resid	109.1983	Schwarz criterion	5.268993
Log likelihood	-34.24389	Hannan-Quinn criter.	5.169248
F-statistic	12.92975	Durbin-Watson stat	1.958215
Prob(F-statistic)	0.003674		

Rappels : théoriquement, le tableau d'analyse de la variance est construit comme suit :

Source de variation	Somme des carrés	Degré de liberté	Carrés moyens
$x_1, x_2, \dots, x_k$	$SCE = \sum_t (\hat{y}_n - \bar{y})^2$	k	SCE/k
Résidu	$SCR = \sum_n e_t^2$	$n-k-1$	$SCR/n-k-1$
Total	$SCT = (y_n - \bar{y})^2$	$n-1$	

Dans notre cas, ce tableau se présente comme suit :

Source de variation	Somme des carrés	Degré de liberté	Carrés moyens
x_1	$SCE1 = 117.65$	1	117.65
$x_2, \dots, x_k$	$SCE = 159.41$	3	53.14
Résidu	$SCR = 67.45$	10	6.14
Total	$SCT = 226.85$	13	

Les détails :

Modèle complet	**Modèle à une variable**
(1) $SCR = 67.45$	*(4)* $SCR1 = 109.20$
(2) $1 = 0.702 + \frac{67.45}{SCT} \rightarrow$ $(1 - 0.702)SCT = 67.45 \rightarrow$ $SCT = \frac{67.45}{0.298} = 226.85$	*(5)* $1 = 0.519 + \frac{109.20}{SCT1} \rightarrow$ $(1 - 0.519)SCT1 = 109.20 \rightarrow$ $SCT1 = \frac{109.20}{0.482} = 226.85$
(3) $SCE = SCT - SCR = 226.85 - 67.45$ $SCT = 159.41$	*(6)* $SCE = SCT1 - SCR1 = 226.85 - 109.20$ $SCT1 = 117.65$

La statistique du test est (test de Fisher contraint) :

$$F_c = \frac{(SCE - SCE1)/(k - k1)}{SCR/(n - k - 1)} = \frac{41.67/(3 - 1)}{67.45/10} = 3.09 < F_{2,10}^{0.05} = 4.10$$

Ou

$$F_c = \frac{(SCR1 - SCR)/(k - k1)}{SCR/(n - k - 1)} = \frac{(109.2 - 67.45)/2}{67.45/10} = 3.09$$

Avec : *k=nombre de régresseurs du modèle complet (à trois variables) ; k1=nombre de régresseurs du modèle à une variable ; SCE=Somme des Carrés Expliqués du modèle complet et SCR1=Somme des Carrés des Résidus du modèle à un régresseur.*

Les Hypothèses du test sont :

H0 : $a_2 = a_3 = 0$ *: les variables ajoutées n'améliorent pas le modèle* ($F_c < F_t$) ;
H1 : au moins 1 de ces deux coefficients est non nul – amélioration ($F_c > F_t$).

Décision : nous acceptons H0 et disons qu'il n'y a pas de différences significatives entre les deux variances estimées. D'où, il n'y a pas amélioration avec l'ajout de nos deux variables « x2 et x3 ».

b) *Stabilité du modèle avec trois régresseurs (X1, X2 et X3) sur tout l'échantillon ? Peut-on effectuer deux estimations : de 1 à 7 et de 8 à 14 ?*

Soient les modèles :

$$y^a = \beta^a x + \varepsilon \dots\dots [E4.1], \qquad avec\ n = 1, .., 14$$
$$y^b = \beta^b x + \varepsilon \dots\dots [E4.2], \qquad avec\ n1 = 1, .., 7$$
$$y^c = \beta^c x + \varepsilon \dots\dots [E4.3], \qquad avec\ n2 = 8, \dots, 14$$

Tester la stabilité d'un modèle revient à (i) effectuer des estimations sur de sous-périodes différentes (généralement 2), (ii) calculer la statistique de Fisher comme suit (*test de Chow*) :

$$F_c = \frac{[SCR - (SCR1 + SCR2)]/dn}{(SCR1 + SCR2)/dd}$$

Avec ($NB: n = n1 + n2$) :

- $dn = (t - k - 1) - [(n1 - k - 1) + (n2 - k - 1)] = k + 1 = 4$;
- $dd = (n1 - k - 1) + (n2 - k - 1) = n - 2(k + 1) = 6$.
- SCR2 : Somme des Carrés des Résidus du modèle [2].

Et (iii) tester les hypothèses jointes suivantes :

H0 : $\beta^a = \beta^b = \beta^c$: pas de différence entre « SCR » et la somme « SCR1+SCR2 », et donc le modèle est stable ($F_c < F_t$) ;

H1 : $\beta^a \neq \beta^b \neq \beta^c$: il y a une différence entre « SCR » et la somme « SCR1+SCR2 », et donc le modèle n'est pas stable ($F_c > F_t$).

Pour ce faire :

- *Obtenir les résultats d'estimation sur chacune de deux sous-périodes (récurer les sommes des carrés des résidus) ;*
- *Calculer la statistique de Fisher et tester les hypothèses du test.*

- *Pour les estimations, sur **EViews 6**, faire (à gauche : les commandes, à droite : les objets générés dans le fichier de travail/Work file, et en bas : les résultats) :*

```
create u 1 14
data y x1 x2 x3
equation modelcomplet.ls y c x1 x2 x3
scalar SCRcomplet=modelcomplet.@ssr
smpl 1 7
equation modelcontraint1.ls y c x1 x2 x3
scalar SCRcont1=modelcontraint1.@ssr
smpl 8 14
equation modelcontraint2.ls y c x1 x2 x3
scalar SCRcont2=modelcontraint2.@ssr
show SCRcont2
```

```
Range:  1 14  --  14 obs
Sample: 8 14  --  7 obs
c
modelcomplet
modelcontraint1
modelcontraint2
resid
scrcomplet
scrcont1
scrcont2
x1
x2
x3
y
```

Scalar SCRCONT2 = 20.7218883326

Dependent Variable: Y
Method: Least Squares
Date: 02/01/14 Time: 13:43
Sample: 1 7
Included observations: 7

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	25.27560	16.66755	1.516456	0.2267
X1	0.773916	0.529033	1.462887	0.2397
X2	-0.293176	0.313677	-0.934644	0.4189
X3	-0.012503	0.100809	-0.124029	0.9091

R-squared	0.692567	Mean dependent var	15.14286
Adjusted R-squared	0.385135	S.D. dependent var	3.848314
S.E. of regression	3.017591	Akaike info criterion	5.342354
Sum squared resid	27.31757	Schwarz criterion	5.311446
Log likelihood	-14.69824	Hannan-Quinn criter.	4.960331
F-statistic	2.252746	Durbin-Watson stat	3.458992
Prob(F-statistic)	0.261019		

Dependent Variable: Y
Method: Least Squares
Date: 02/01/14 Time: 13:47
Sample: 8 14
Included observations: 7

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	62.33574	37.23454	1.674137	0.1927
X1	1.228196	0.685233	1.792376	0.1710
X2	-0.620833	0.522363	-1.188508	0.3201
X3	-0.184339	0.152831	-1.206160	0.3142

R-squared	0.543858	Mean dependent var	20.28571
Adjusted R-squared	0.087716	S.D. dependent var	2.751623
S.E. of regression	2.628174	Akaike info criterion	5.066015
Sum squared resid	20.72189	Schwarz criterion	5.035106
Log likelihood	-13.73105	Hannan-Quinn criter.	4.683992
F-statistic	1.192299	Durbin-Watson stat	2.948266
Prob(F-statistic)	0.444230		

- *La statistique du test (Fisher) se calcul comme suit :*

$$F_c = \frac{[67.45 - (27.31 + 20.73)]/4}{(27.31 + 20.73)/6} = \frac{4.852}{8} = 0.606 < F_{4;6}^{0.05} = 4.53$$

Décision : nous acceptons « Ho » et disons que notre modèle estimé est stable sur l'échantillon considéré (14 observations). NB : l'hétéroscédasticité implique bien souvent le rejet de « Ho » (biais du test de Chow dans ce cas).

c) Tests : $b_1 = 1$ et $b_2 = b_3$?

Cette hypothèse amène à écrire l'équation :

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{3i} + e_i \ldots\ldots [E4.4]$$

Comme suit :

$$y_i = b_0 + 1x_{1i} + b_2 x_{2i} + b_2 x_{3i} + e_i \ldots\ldots [E4.5]$$

Autrement :

$$(y_i - x_{1i}) = b_0 + b_2(x_{2i} + x_{3i}) + e_i$$

Voir même :

$$p_i = b_0 + b_2 q_i + e_i \ldots\ldots [E4.6]$$

Avec : $p_i = (y_i - x_{1i})$ et $q_i = (x_{2i} + x_{3i})$

Il est question (i) d'estimer le modèle « E4.6 » pour obtenir les « SCT1 », « SCE1 » et « SCR1 » relatives à « E4.6 », et (ii) de calculer la statistique F de Fisher comme suit :

$$F_c = \frac{[SCR1 - SCR]/dn}{SCR/(n-k-1)}$$

Avec : $dn = (n - k' - 1) - (n - k - 1) = k - k'$

Pour ce faire, sur **EViews 6**, faire (*à gauche : les commandes et statistiques calculées et, à droite : les résultats d'estimation*) :

```
smpl 1 14
genr P=Y-X1
genr Q=X2+X3
equation modelcontraint3.ls P c Q
scalar SCT1=@ssr/(1-@r2)
scalar SCE1=SCT-@ssr
```

$SCR1 = 108.78$

$SCE1 = 0.425$

$SCT1 = 109.21$

Dependent Variable: P
Method: Least Squares
Date: 02/01/14 Time: 14:55
Sample: 1 14
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	13.73517	9.691559	1.417230	0.1819
Q	-0.011155	0.051490	-0.216638	0.8321

R-squared	0.003896	Mean dependent var	11.64286
Adjusted R-squared	-0.079113	S.D. dependent var	2.898465
S.E. of regression	3.010936	Akaike info criterion	5.173943
Sum squared resid	108.7888	Schwarz criterion	5.265236
Log likelihood	-34.21760	Hannan-Quinn criter.	5.165492
F-statistic	0.046932	Durbin-Watson stat	1.879728
Prob(F-statistic)	0.832130		

Notre statistique se calcul comme suit :

$$F_c = \frac{[SCR1 - SCR]/dn}{SCR/(n-k-1)} = \frac{(108.78 - 67.45)/2}{67.45/10} = 3.06 < F^{0,05}_{2;10} = 4.10$$

Avec : $dn = (n - k' - 1) - (n - k - 1) = k - k' = 2$

Et les hypothèses du test sont :

H0 : $(SCR1 = SCR)$: *toutes les restrictions sont valides/vérifiées* $(F_c < F_t)$;
H1 : $SCR1 \neq SCR$: *les restrictions ne sont pas valides/non vérifiées* $(F_c > F_t)$.

Décision : nous acceptons « H0 » et disons que les deux restrictions imposées aux paramètres sont adaptées aux données faisant l'objet de l'étude.

VI. Etude de saisonnalité par variables indicatrices (avec EViews 6) : Feuille « E5 »

- **TD** : Sur base de données « **E5** », il est demandé de :
  - *Estimer la relation « dépenses de publicité et ventes » (sur de données trimestrielles pour 5ans) et commenter les résultats :*

$$VENTE_t = b_0 + b_1 PUB_t + e_t \dots\dots [E5.1]$$

  - *Représenter graphiquement la série des ventes et commenter son évolution ;*
  - *Spécifier et estimer le modèle adéquat.*

- ***Solution*** :

Estimation de la relation « E5.1 ». Sur **EViews**, *faire :*

```
create q 2000:01 2004:04
data VENTE PUB
ls VENTE c PUB
```

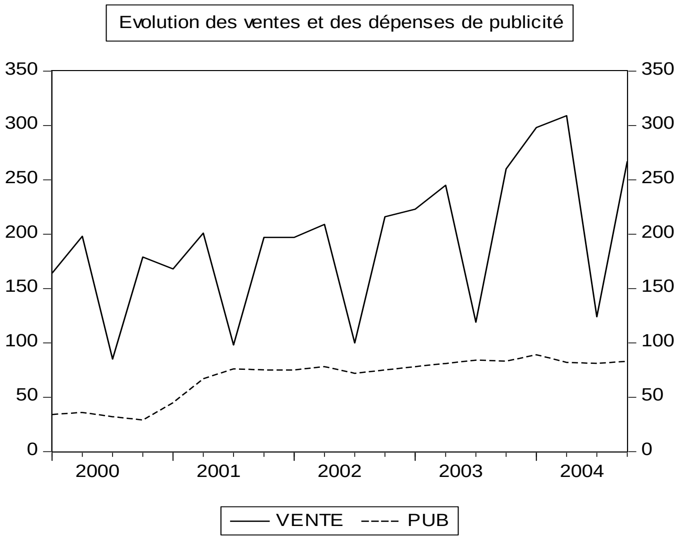
Dependent Variable: VENTE
Method: Least Squares
Date: 02/01/14 Time: 22:28
Sample: 2000Q1 2004Q4
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	104.8959	49.35643	2.125274	0.0477
PUB	1.298215	0.699986	1.854631	0.0801

R-squared	0.160434	Mean dependent var	192.8500
Adjusted R-squared	0.113792	S.D. dependent var	64.96661
S.E. of regression	61.15868	Akaike info criterion	11.15946
Sum squared resid	67326.91	Schwarz criterion	11.25903
Log likelihood	-109.5946	Hannan-Quinn criter.	11.17890
F-statistic	3.439657	Durbin-Watson stat	2.399407
Prob(F-statistic)	0.080106		

<u>Commentaire</u> : Il en découle que les « *dépenses de publicité* » n'exerce pas d'influence sur les ventes (au seuil de 5%).

*Représentation graphique des séries « vente et publicité » et commentaire : Sur **EViews**, faire : plot VENTE PUB*



<u>Commentaires</u> : Visiblement, la série des ventes subit un effet saisonnier (l'influence du troisième trimestre) accusant un creux très prononcé à chaque troisième trimestre, contrairement aux dépenses de publicité. Ce mouvement saisonnier est de nature à occulter l'estimation économétrique.

Spécification et estimation du modèle adéquat : le modèle adéquat est celui qui prend en compte/intègre l'effet saisonnier comme suit (l'on introduit « k-1 » variables indicatrices pour éviter la multi-colinéarité) :

$$VENTE_t = b_0 + b_1 PUB_t + b_2 T_{1t} + b_3 T_{2t} + b_4 T_{3t} + e_t \dots \dots [E5.2]$$

Les résultats se présentent comme suit : ls VENTE c PUB T1 T2 T3

Dependent Variable: VENTE
Method: Least Squares
Date: 02/02/14 Time: 00:07
Sample: 2000Q1 2004Q4
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	129.1014	27.31974	4.725571	0.0003
PUB	1.372444	0.344994	3.978172	0.0012
T1	-7.212271	19.03064	-0.378982	0.7100
T2	8.874489	18.95858	0.468099	0.6464
T3	-118.6000	18.95846	-6.255784	0.0000

R-squared	0.831925	Mean dependent var	192.8500
Adjusted R-squared	0.787105	S.D. dependent var	64.96661
S.E. of regression	29.97595	Akaike info criterion	9.850986
Sum squared resid	13478.36	Schwarz criterion	10.09992
Log likelihood	-93.50986	Hannan-Quinn criter.	9.899580
F-statistic	18.56147	Durbin-Watson stat	0.662664
Prob(F-statistic)	0.000011		

<u>Commentaires</u> : Il en découle que les dépenses en publicité améliorent significativement les ventes, et la saisonnalité de ces dernières est principalement liée aux creux du troisième trimestre.

VII. *Tests de stabilité structurelle et test de spécification de Ramsey (avec EViews 6) : Feuille « E6 »*

- ***TD*** : Sur base de données « **E6** », il est demandé de :
  - Tester un éventuel changement structurel, partant d'une estimation récursive (tester la stabilité des coefficients : tests de CUSUM)
  - *Vérifier la qualité/bonté de la spécification.*
- ***Solution*** :

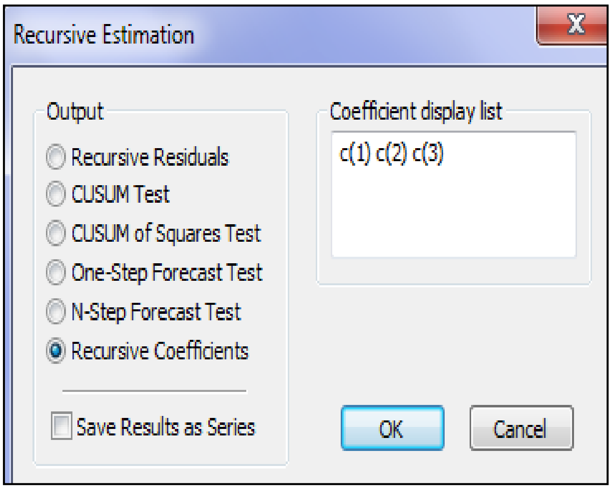
*Teste d'un éventuel changement structurel, partant d'une estimation récursive (test de la stabilité des coefficients : tests de CUSUM). Sur **EViews**, faire :*

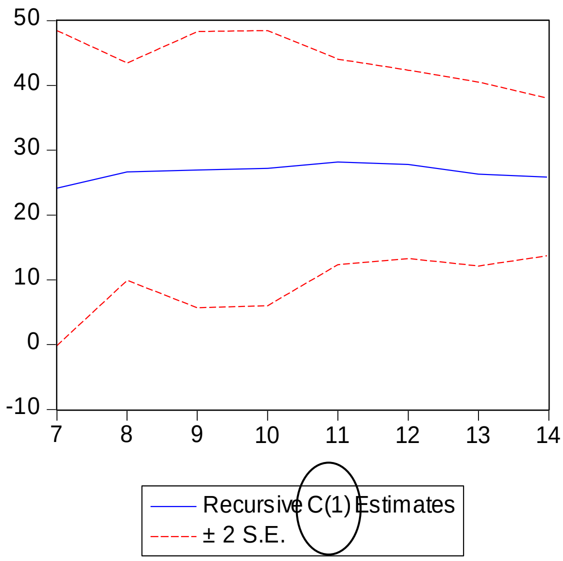
Ls Y c X1 X2
Dans l'output des résultats, suivre : View/Stability Tests/Recursive Estimates(OLS only)… : la figure à droite complète la procédure et, à gauche et en bas : les résultats des instructions données :

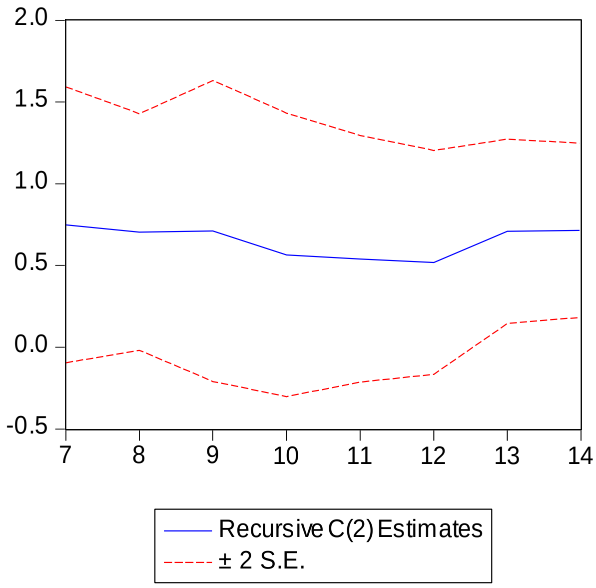
Dependent Variable: Y
Method: Least Squares
Date: 02/02/14 Time: 00:29
Sample: 1 14
Included observations: 14

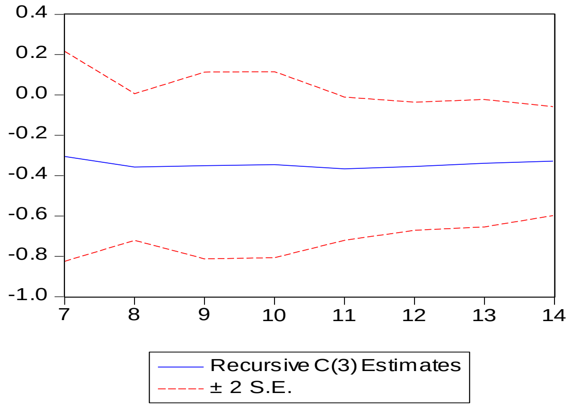
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	25.84214	6.064674	4.261092	0.0013
X1	0.714896	0.266264	2.684916	0.0212
X2	-0.328113	0.134561	-2.438392	0.0329

R-squared	0.687540	Mean dependent var	17.71429
Adjusted R-squared	0.630729	S.D. dependent var	4.177385
S.E. of regression	2.538501	Akaike info criterion	4.888434
Sum squared resid	70.88389	Schwarz criterion	5.025375
Log likelihood	-31.21904	Hannan-Quinn criter.	4.875758
F-statistic	12.10223	Durbin-Watson stat	3.078032
Prob(F-statistic)	0.001665		



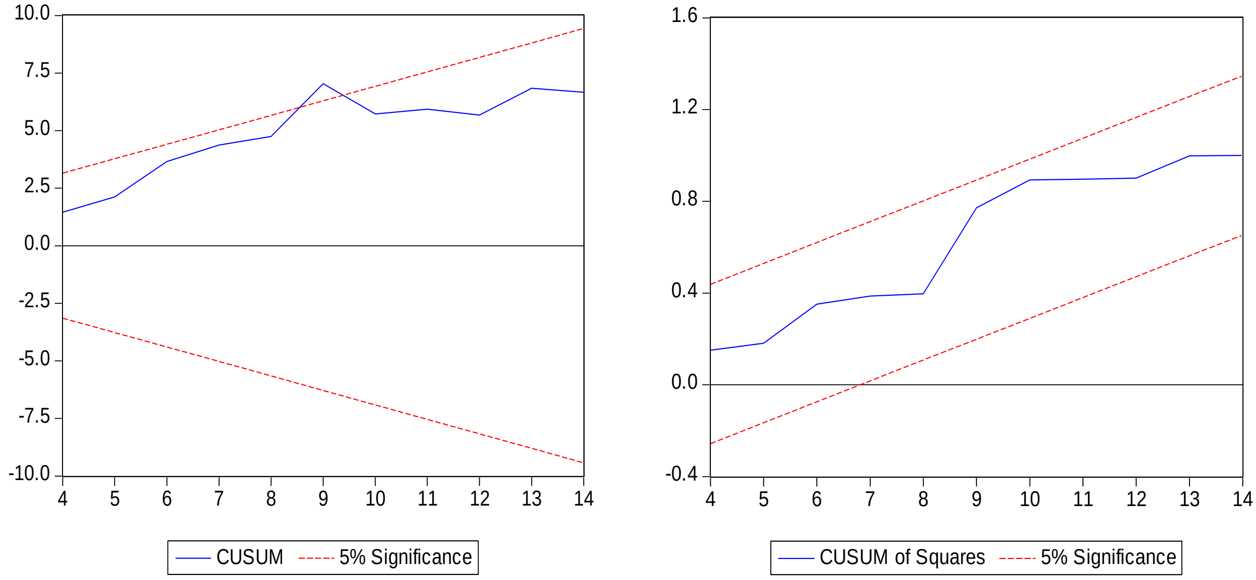






*Commentaires : Nos trois coefficients sont à l'intérieur de l'intervalle de confiance (±2*écart-type), ce qui amène à conclure en faveur de la stabilité des paramètres (on commence à la 7ème observation, parce que le degré de liberté est faible).*

Sélectionner « CUSUM Test » et « CUSUM of Squares Test » pour obtenir les graphiques qui suivent :



Commentaires : les statistiques CUSUM et CUSUM Squares évoluent à l'intérieur de l'intervalle de confiance, sauf un léger débordement pour la 9ème (Cfr CUSUM). Ainsi, nous concluons en faveur de l'absence d'un changement structurel.

*Test de Ramsey : Sur **EViews**, dans l'output des résultats, suivre : View/Stability Tests/Ramsey RESET Test… →Number of fitted : 1 :*

Ramsey RESET Test:

F-statistic	0.049866	Prob. F(1,10)	0.8278
Log likelihood ratio	0.069639	Prob. Chi-Square(1)	0.7919

Test Equation:
Dependent Variable: Y
Method: Least Squares
Date: 02/02/14 Time: 01:08
Sample: 1 14
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16.40322	42.74223	0.383771	0.7092
X1	0.326961	1.759415	0.185835	0.8563
X2	-0.150146	0.809298	-0.185526	0.8565
FITTED^2	0.015484	0.069342	0.223308	0.8278

Commentaires : le modèle est bien spécifié (la variable non linéaire du type quadratique est non significative et la probabilité associée à la statistique de Fisher calculée est >5%).

VIII. *Tests de multi-colinéarité (avec EViews 6) : Feuille « E7 »*

- ***TD*** : Sur base de données « **E7** », il est demandé (i) d'estimer la relation suivante :

$$y_t = b_0 + b_1x_{1t} + b_2x_{2t} + b_3x_{3t} + b_4x_{4t} + e_t \dots\dots [E7.1]$$

et d'effectuer les tests de multi-colinéarité ci-après :

- *Test de KLEIN ;*
- *Test de FARRAR-GLAUBER.*

- ***Solution*** :

*a) Résultats d'estimation : Sur **EViews**, faire : ls y c x1 x2 x3 x4*

Dependent Variable: Y
Method: Least Squares
Date: 02/02/14 Time: 09:56
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-13.51825	7.510065	-1.800018	0.1318
X1	0.097017	0.026480	3.663736	0.0145
X2	0.015012	0.049366	0.304089	0.7733
X3	-0.199241	0.090087	-2.211651	0.0779
X4	0.340042	0.149646	2.272306	0.0722

R-squared	0.998011	Mean dependent var	14.00000
Adjusted R-squared	0.996420	S.D. dependent var	4.301163
S.E. of regression	0.257334	Akaike info criterion	0.429971
Sum squared resid	0.331105	Schwarz criterion	0.581264
Log likelihood	2.850143	Hannan-Quinn criter.	0.264004
F-statistic	627.3275	Durbin-Watson stat	3.382574
Prob(F-statistic)	0.000001		

b) Test de KLEIN

Obtenir la matrice de corrélations simples, taper : ***cor x1 x2 x3 x4***

	X1	X2	X3	X4
X1	*1*	*0.9883175015361664*	*0.9803676239479733*	*0.9876720483315924*
X2	*0.9883175015361664*	*1*	*0.9699616465691321*	*0.9694773335463723*
X3	*0.9803676239479733*	*0.9699616465691321*	*1*	*0.9917959138067758*
X4	*0.9876720483315924*	*0.9694773335463723*	*0.9917959138067758*	*1*

Obtenir la matrice de corrélations partielles au carré, taper (Cfr Excel) :

	C1	*C2*	*C3*	*C4*
R1	*1*	*0.97677148384269*	*0.961120678085394*	*0.975496075055522*
R2	*0.97677148384269*	*1*	*0.940825595815102*	*0.939886300260183*
R3	*0.961120678085394*	*0.940825595815102*	*1*	*0.983659134643816*
R4	*0.975496075055522*	*0.939886300260183*	*0.983659134643816*	*1*

Commentaires : pas de risques de multi-colinéarité parce que les coefficients de corrélations simples élevés au carré sont tous inférieurs à « R^2=0.998 ». Mais, c'est inquiétant que ces coefficients soient aussi élevés. Que dit *Farrar-Glauber* ?

c) Test de FARRAR-GLAUBER

Calcul du déterminant de la matrice de corrélations simples ci-haut (*matrice nommée « A », et DA : son déterminant*) : Sur **EViews**, faire :

```
matrix (4,4) A
A(2,1)=0.988
A(3,1)=0.980
A(4,1)=0.987
…
scalar DA=@det(A)
scalar LDA=log(DA)
```

- *La statistique du test (Chi-carré) se calcule comme suit (sur 10 observations) :*

$$\chi_c^2 = -\left[n - 1 - \frac{1}{6}(2K+5)\right] \times \ln DA = -\left[10 - 1 - \frac{1}{6}(2 \times 5 + 5)\right] \times (-11.59) = 75.33$$

Théoriquement, χ_t^2 à « $\frac{1}{2}K(K-1) = \frac{20}{2} = 10$ » degrés de liberté et au seuil de 5% vaut :

$$\chi_t^2 = 18.31$$

- *Les hypothèses du test sont (NB : $DA = 9.219785 \times 10^{-6}$) :*

  H0 : $DA = 1$: Présomption d'absence de multi-colinéarité ($\chi_c^2 < \chi_t^2$) ;

  H1 : $DA \neq 1 (DA = 0)$: il y a présomption de multi-colinéarité ($\chi_c^2 > \chi_t^2$).

Décision : parce que « $\chi_c^2 > \chi_t^2$ », nous présumons qu'il y a multi-colinéarité (le test de Farrar-Glauber est privilégié du fait de son fondement théorique affirmé).

IX. Tests d'indépendance des erreurs (avec EViews 6) : Feuille « E8 »

- ***TD*** : Sur base de données « **E8** », il est demandé de :
  - *Estimer les coefficients du modèle ;*
  - *Analyser l'évolution graphique des résidus ;*
  - *Calculer la statistique de Durbin-Watson et tester l'autocorrélation des erreurs d'ordre 1 ;*
  - *Tester l'autocorrélation des erreurs d'ordre 2 (test de Breusch-Godfrey).*
- ***Solution*** :

a) Estimation du modèle « E8.1 »/MCO : $y_t = b_0 + b_1x_{1t} + b_2x_{2t} + b_3x_{3t} + e_t$

Dependent Variable: Y
Included observations: 20

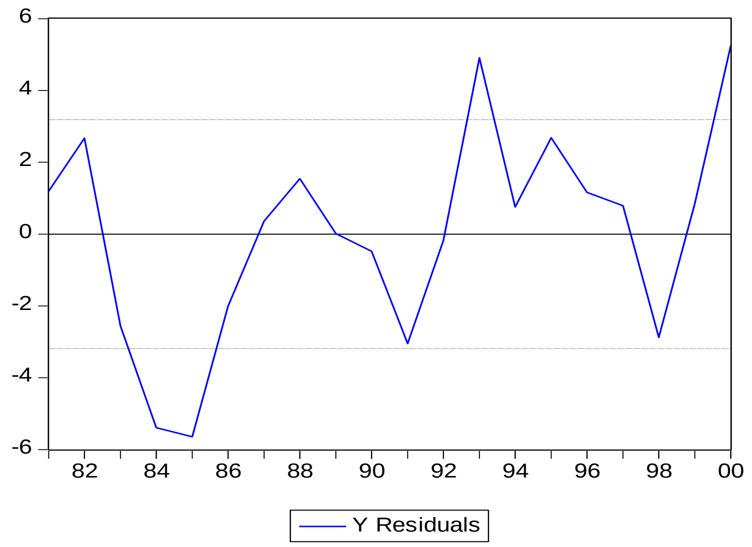
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-242.7951	26.79995	-9.059538	0.0000
X1	3.897408	0.400325	9.735616	0.0000
X2	0.404365	0.061351	6.590977	0.0000
X3	-0.878886	0.240216	-3.658727	0.0021

R-squared	0.938895	Mean dependent var	97.53500
Adjusted R-squared	0.927438	S.D. dependent var	11.83048
F-statistic	81.94784	Durbin-Watson stat	1.053794
Prob(F-statistic)	0.000000		

Note : *tous les paramètres sont statistiquement significatifs (prob<5%) et le modèle est globalement bon (Prob-Fisher<5%).*

b) Analyse graphique de l'évolution des résidus

*Sur **EViews**, dans l'output de l'estimation, suivre : View/Actual, Fitted, Residual/Residual Graph :*



Note : à cause de l'évolution cyclique des résidus, il y a présomption d'autocorrélation positive des résidus.

c) Calcul de la statistique de Durbin-Watson et teste d'autocorrélation des erreurs d'ordre 1

Il ressort des estimations que le coefficient de Durbin-Watson (DW) vaut : « DW=1.053794 ». Sur EViews, ce coefficient se calcule comme suit :

```
ls Y c X1 X2 X3
scalar SCR=@ssr
genr E=resid
genr DE=d(E)
genr DE2=DE^2
scalar SDE2=@sum(DE2)
scalar DW=SDE2/SCR
```

Sachant que :

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=2}^{T} {e_t}^2} = SDE2/SCR$$

Avec : $SDE2 = \sum_{t=2}^{T}(e_t - e_{t-1})^2$ *et*

$$SCR = \sum_{t=2}^{T} {e_t}^2$$

Pour n=20 et k=3 (d1=1,00 et d2=1,68), la valeur DW calculée est comprise entre « $d_1 < DW < d_2$ » : la zone d'incertitude (immédiatement proche de la zone de rejet). Nous présumons une dépendance positive des résidus.

d) Teste d'autocorrélation des erreurs d'ordre 2 (test de Breusch-Godfrey)

Sur ***EViews****, dans l'output de l'estimation, suivre : View/Residual Tests/Serial Correlation LM test… →lags to include : 2 →ok : pour effectuer le test de Breusch-Godfrey (ci-dessous les résultats).*

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.855850	Prob. F(2,14)	0.0912
Obs*R-squared	5.795238	Prob. Chi-Square(2)	0.0552

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 02/02/14 Time: 13:48
Sample: 1981 2000
Included observations: 20
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.934137	24.60858	-0.363050	0.7220
X1	0.102924	0.366821	0.280585	0.7831
X2	-0.026070	0.056469	-0.461671	0.6514
X3	0.011036	0.229070	0.048177	0.9623
RESID(-1)	0.662840	0.283063	2.341671	0.0345
RESID(-2)	-0.389645	0.290089	-1.343190	0.2006

R-squared	0.289762	Mean dependent var	-1.19E-14
Adjusted R-squared	0.036105	S.D. dependent var	2.924432
S.E. of regression	2.871152	Akaike info criterion	5.190629
Sum squared resid	115.4092	Schwarz criterion	5.489349
Log likelihood	-45.90629	Hannan-Quinn criter.	5.248942
F-statistic	1.142340	Durbin-Watson stat	2.006711
Prob(F-statistic)	0.383889		

Décision : Il y a absence d'autocorrélation d'ordre 2, mais présence d'une autocorrélation d'ordre 1 (les estimateurs sont sans biais, mais leurs variances ne sont pas minimales. D'où la correction, par une méthode adéquate, tient lieu avant d'interpréter les résultats).

X. Tests de détection d'une hétéroscédasticité et correction (avec EViews 6) : Feuille « E9 »

- **TD** : Sur base de données « **E9** », il est demandé d'effectuer les tests suivants (détection d'hétéroscédasticité) :
  - *Test d'égalité des variances ;*
  - *Test de Golfed-Quandt ;*
  - *Test de Gleisjer ;*
  - *Test de White ;*
  - *Si hétéroscédasticité il y en a, en corriger.*
- **Solution :**

a) Test d'égalité des variances

*La statistique du test « **Q** » (Q' amélioré) se calcul comme suit :*

$$Q = \frac{Q'}{C} \rightarrow \chi^2(m-1)$$

Avec : m=nombre de classes (6) ;

$$C = 1 + \frac{1}{3(m-1)}\left(\sum_{i=1}^{m} \frac{1}{v_i} - \frac{1}{v}\right)$$

$$Q' = v \ln \hat{\sigma}_T^2 - \sum_{i=1}^{m} v_i \ln \hat{\sigma}_i^2 \rightarrow \chi^2(m-1)$$

$$v_i = n_i - 1 \; et \; v = \sum_{i=1}^{m} v_i = \sum_{i=1}^{m} (n_i - 1)$$

Variance empirique pour chaque groupe « $\hat{\sigma}_i^2$ » :

$$\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{ni}\left(Y_{ij} - \bar{Y}_i\right)^2}{n_i - 1} = \frac{\sum_{j=1}^{5}\left(Y_{ij} - \bar{Y}_i\right)^2}{4}$$

Variance totale « $\hat{\sigma}_T^2$ » :

$$\hat{\sigma}_T^2 = \frac{\sum_{i=1}^{m}(n_i - 1)^2 \hat{\sigma}_i^2}{\sum_{i=1}^{m}(n_i - 1)} = \frac{\sum_{i=1}^{m} v_i \hat{\sigma}_i^2}{v}$$

Les hypothèses du test sont :

H0 : le modèle est homoscédastique $(Q < \chi^2_{0.95}(m) \; ; \; prob.F - stat > 5\%)$ *;*
H1 : le modèle est hétéroscédastique $(Q > \chi^2_{0.95}(m) \; ; \; prob.F - stat < 5\%)$.

Statistique calculée : Q'=13.26 ; C=1.097 ; Q=12.09 et $\chi^2_{0.95}(6) = 11.07$.

Décision : Parce que $Q > \chi^2_{0.95}$, *nous concluons que le modèle est hétéroscédastique.*

______***Test d'égalité des variances à l'automatique sur EViews*** *: faire (commandes) :*

Show X Y
Dans la base de données, suivre : View/Tests of Equality…→la boîte de dialogue à droite apparaît (elle complète la procédure)→ok :

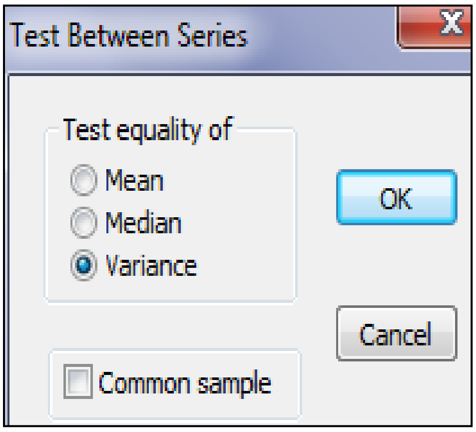
Test for Equality of Variances Between Series
Date: 02/03/14 Time: 00:52
Sample: 1 30
Included observations: 30

Method	df	Value	Probability
F-test	(29, 29)	44.03433	0.0000
Siegel-Tukey		0.034570	0.9724
Bartlett	1	69.66292	0.0000
Levene	(1, 58)	32.91994	0.0000
Brown-Forsythe	(1, 58)	29.37368	0.0000

Category Statistics

Variable	Count	Std. Dev.	Mean Abs. Mean Diff.	Mean Abs. Median Diff.	Mean Tukey-Siegel Rank
X	30	1.293774	1.111111	1.083333	30.59444
Y	30	8.585272	6.666667	6.566667	30.40556
All	60	9.100331	3.888889	3.825000	30.50000

Bartlett weighted standard deviation: 6.139249



b) Test de Golfed-Quandt

______**Conditions du test** : la variable à la base de l'hétéroscédasticité est connue et la taille de l'échantillon est élevée. Les étapes sont : (i) *classer les observations en ordre croissant ou décroissant de la variable expliquée ou du régresseur à la base de l'hétéroscédasticité ; (ii) omettre « c » observations centrales ; et (iii) faire une régression sur deux sous échantillons et effectuer le test.*

______**Application des étapes** :

(i) *Classement de données* : de par leur présentation, les données sont déjà classées suivant les valeurs décroissantes de « X ». *Si tel n'était pas le cas, pour un classement en ordre croissant suivant les valeurs de X, sur* **EViews**, *l'on aurait du faire (dans le workfile) : Proc/Sort Current Page…→yes→X→ascending→ok ;*

(ii) *Omission de « C » observations centrales : C constitue de valeurs centrales à choisir arbitrairement. Bien souvent, $C = int(T/4)$, avec « int : partie entière ». Dans le cas d'espèce, $C = int(30/4) = 7.5 \approx 8$.*

(iii) Régression sur deux sous-échantillons en omettant « C=8 » observations : 30-8=22/2=11 :

- *1er Sous-échantillon : j=1,..,11 ;*
- *2ème Sous-échantillon : j= 20,…,30.*

Ci-dessous, les commandes **EViews** *suivies des résultats des estimations par sous-échantillons (à gauche : le 1er sous-échantillon et, à droite : le 2ème sous-échantillon) :*

```
create u 1 30
data Y X
smpl 1 11
ls Y C X
smpl 20 30
ls Y C X
```

Dependent Variable: Y
Method: Least Squares
Date: 02/02/14 Time: 16:24
Sample: 1 11
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16.93333	8.418293	2.011493	0.0751
X	-2.133333	2.316660	-0.920866	0.3811

R-squared	0.086108	Mean dependent var	9.272727
Adjusted R-squared	-0.015435	S.D. dependent var	4.244783
S.E. of regression	4.277417	Akaike info criterion	5.907541
Sum squared resid	164.6667	Schwarz criterion	5.979886
Log likelihood	-30.49148	Hannan-Quinn criter.	5.861938
F-statistic	0.847994	Durbin-Watson stat	1.080027
Prob(F-statistic)	0.381141		

Dependent Variable: Y
Method: Least Squares
Date: 02/02/14 Time: 16:25
Sample: 20 30
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23.08000	8.117016	2.843409	0.0193
X	-1.320000	9.233860	-0.142952	0.8895

R-squared	0.002265	Mean dependent var	22.00000
Adjusted R-squared	-0.108594	S.D. dependent var	9.348797
S.E. of regression	9.843328	Akaike info criterion	7.574431
Sum squared resid	872.0200	Schwarz criterion	7.646775
Log likelihood	-39.65937	Hannan-Quinn criter.	7.528827
F-statistic	0.020435	Durbin-Watson stat	1.748023
Prob(F-statistic)	0.889477		

Statistique du test : la statistique du test est un « Fisher » construit/calculé comme suit (*au numérateur, la « SCR » la plus élevée*) :

$$F_c = \frac{SCR_2/dl_2}{SCR_1/dl_1} = \frac{872.02}{164.66} = 5.29 \rightarrow F_{9,9}^{0.05} = 3.18$$

Les hypothèses du test :

H0 : le modèle est homoscédastique $(F_c < F_{9,9}^{0.05})$ *;*
H1 : le modèle est hétéroscédastique $(F_c > F_{9,9}^{0.05})$.

Décision : *Rejet de H0, le modèle est hétéroscédastique* $(F_c > F_{9,9}^{0.05})$.

c) *Test de Gleisjer*

Avantages et principes : détection de l'hétéroscédasticité et sa forme. On régresse les résidus de l'estimation de base sur la variable responsable d'hétéroscédasticité (*régresseur significatif = hétéroscédasticité*) ;

Test de Gleisjer à l'automatique *: Dans l'output de l'estimation, suivre : View/Residual Tests/Heteroskedasticity Tests…→la boîte de dialogue à droite s'affiche (elle complète la procédure) et, à gauche : les résultats du test :*

NB : par défaut, la forme choisie est linéaire, càd : $|e_j| = f(X_j)$, *avec l'hétéroscédasticité du type «* $\hat{\sigma}^2_{uj} = k^2 X_j^2 \ (k = constante)$ *». Si non, Gleisjer suggère de tester aussi les différentes formes suivantes :*

- $|e_j| = f\left(X_j^{1/2}\right) \rightarrow \hat{\sigma}^2_{uj} = k^2 X_j \ (k = constante)$
- $|e_j| = f(X_j^{-1}) \rightarrow \hat{\sigma}^2_{uj} = k^2 X_j^{-2} (k = constante)$

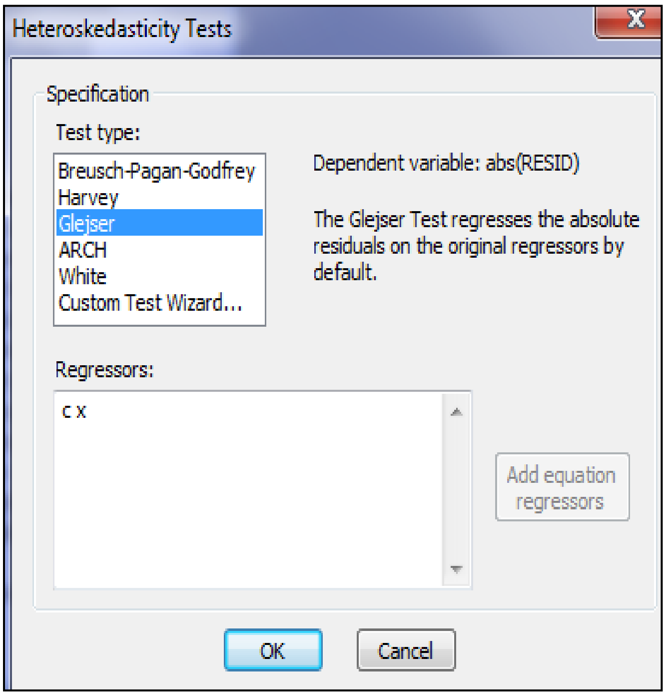
Heteroskedasticity Test: Glejser

F-statistic	6.548031	Prob. F(1,28)	0.0162
Obs*R-squared	5.686024	Prob. Chi-Square(1)	0.0171
Scaled explained SS	6.140230	Prob. Chi-Square(1)	0.0132

Test Equation:
Dependent Variable: ARESID
Method: Least Squares
Date: 02/02/14 Time: 16:59
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.090138	1.399626	5.780215	0.0000
X	-1.467253	0.573389	-2.558912	0.0162

R-squared	0.189534	Mean dependent var	5.033362
Adjusted R-squared	0.160589	S.D. dependent var	4.360332
S.E. of regression	3.994909	Akaike info criterion	5.672259
Sum squared resid	446.8603	Schwarz criterion	5.765672
Log likelihood	-83.08389	Hannan-Quinn criter.	5.702143
F-statistic	6.548031	Durbin-Watson stat	1.500031
Prob(F-statistic)	0.016193		



Les hypothèses du test :

$H0$: *le modèle est homoscédastique* $(régresseur\ non\ significatif - prob > 5\%)$;
$H1$: *le modèle est hétéroscédastique* $(régresseur\ significatif - prob < 5\%)$.

Décision : nous rejetons l'hypothèse nulle et concluons à la présence d'hétéroscédasticité $(régresseur\ significatif - prob < 5\%)$.

d) Test de White

Principes : test proche du précédent, il consiste à régresser le carré des résidus de l'estimation de base sur une ou plusieurs variables explicatives en niveau et élevées au carré, dans une même équation (les termes croisés peuvent être intégrés s'ils sont source d'hétéroscédasticité. Aussi, *régresseur significatif = hétéroscédasticité*).

Les hypothèses du test :

$H0$: *le modèle est homoscédastique* $[(n \times R_a^2) < \chi^2(p) - prob > 5\%]$;
$H1$: *le modèle est hétéroscédastique* $[(n \times R_a^2) > \chi^2(p) - prob < 5\%]$

Avec « $p = 2k$ » et « R_a^2 » : coefficient de détermination du modèle faisant l'objet du test (modèle associé).

Décision : *Sur EViews, Cfr procédure de « Gleisjer », cocher « **White** » pour obtenir les résultats ci-dessous :*

Heteroskedasticity Test: White				
F-statistic	3.956388	Prob. F(2,27)		**0.0311**
Obs*R-squared	6.799324	Prob. Chi-Square(2)		**0.0334**
Scaled explained SS	7.172065	Prob. Chi-Square(2)		**0.0277**
Test Equation:				
Dependent Variable: RESID^2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	136.0182	43.70370	3.112280	0.0044
X	-78.58151	48.29367	-1.627160	**0.1153**
X^2	11.98436	10.25804	1.168289	**0.2529**

Nous rejetons l'hypothèse nulle : les erreurs sont hétéroscédastiques $(prob < 5\%)$.

e) Correction de l'hétéroscédasticité

Sous l'hypothèse que les erreurs sont hétéroscédastique du type : $\hat{\sigma}_{uj}^2 = k^2 X_j$ $(k = constante)$, *le modèle homoscédastique est celui transformé comme suit :*

$$\frac{Y_j}{\sqrt{X_j}} = \frac{a_0}{\sqrt{X_j}} + a_1 \frac{X_j}{\sqrt{X_j}} + \frac{u_j}{\sqrt{X_j}} \text{ , sachant que } f\left(\frac{u_j}{\sqrt{X_j}}\right)^2 = \frac{1}{\sqrt{X_j}} = k^2$$

Il peut s'écrire :

$$Q_j = b_1 P_{1j} + b_2 P_{2j} + e_j$$

Avec : $Q_j = \frac{Y_j}{\sqrt{X_j}}$; $P_{1j} = \frac{1}{\sqrt{X_j}}$ et $P_{2j} = \frac{X_j}{\sqrt{X_j}}$

Le modèle de départ s'écrit :

$$Y_j = a_0 + a_1 X_j + e_j$$

Sur EViews, *taper les commandes suivantes et obtenir les résultats ci-dessous :*

```
genr RX=sqr(X)
genr Q=Y/RX
genr P1=1/RX
genr P2=X/RX
ls Q P1 P2
```

Dependent Variable: Q
Method: Least Squares
Date: 02/02/14 Time: 17:56
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
P1	24.94147	2.503469	9.962763	0.0000
P2	-4.531905	1.535487	-2.951445	0.0063

R-squared	0.643778	Mean dependent var	14.42003
Adjusted R-squared	0.631055	S.D. dependent var	12.44131
S.E. of regression	7.556954	Akaike info criterion	6.947154
Sum squared resid	1599.012	Schwarz criterion	7.040567
Log likelihood	-102.2073	Hannan-Quinn criter.	6.977038
Durbin-Watson stat	1.723276		

La correspondance entre les coefficients des modèles se fait comme suit :

$$Q_j = b_1 P_{1j} + b_2 P_{2j} + e_j$$
$$Y_j = a_0 + a_1 X_j + e_j$$

$\hat{a}_0 = \hat{b}_1 = 24.96$ et $\hat{a}_1 = \hat{b}_2 = -4.53$. Le coefficient de détermination n'est pas significatif (modèle sans terme constant).

***Commentaire** : le temps passé à la vérification d'un véhicule par le chef d'atelier influe significativement sur le nombre de défauts détectés : 1 heure de vérification permet de supprimer 4,5 défauts en moyenne.*

XI. Estimation d'une fonction de production « de type Cobb-Douglass » (avec EViews 6) : Feuille « E11 »

Soit le modèle/fonction de production de type Cobb-Douglass spécifié comme suit :

$$Q_i = a_0 K_i^{a_1} L_i^{a_2} e^{\varepsilon} \dots\dots\dots\dots [E11.1]$$

Sous forme linéaire :

$$LQ_i = a_0 + a_1 LK_i + a_1 LL_i + \varepsilon \dots\dots\dots . [E11.2]$$

Avec :

- i : 25 entreprises ;
- Q_i : la production, en millions d'Euros, pour l'entreprise « i » (LQ_i : *la production transformée en logarithme*) ;
- K_i : le facteur capital, en millions d'Euros, pour l'entreprise « i » (LK_i : *le capital transformé en logarithme*) ;
- L_i : le facteur travail, en millions d'Euros, pour l'entreprise « i » (LL_i : *le travail transformé en logarithme*).
- a_0 : la technologie (paramètre à estimer) ;
- « a_1 *et* a_2 » : les rendements (paramètres à estimer).

TD : Sur base de données « **E11** », il est demandé de :

- *Estimer le modèle « E11.1 » ci-dessus et donner une interprétation économique aux coefficients estimés « $\hat{a}_1$ et $\hat{a}_2$ » (commenter les résultats) ;*
- *Calculer le risque « α » de première espèce sur l'hypothèse des rendements d'échelle constants.*

Solution :

a) Estimation du modèle « E11.1 » par les MCO et interprétation des résultats

Sur **EViews**, taper les commandes suivantes (à gauche : *le modèle brut « E11.1 » estimé* et, à droite : *le modèle transformé « E11.2 » estimé*) :

```
create u 1 25
data Qi Ki Li
equation modelbrut.ls Qi c Ki Li
genr LQi=log(Qi)
genr LKi=log(Ki)
genr LLi=log(Li)
equation modeltransforme.ls LQi c LKi LLi
```

Dependent Variable: QI
Method: Least Squares
Date: 02/08/14 Time: 20:42
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.155253	9.014401	0.239090	0.8132
KI	11.64477	1.135392	10.25617	0.0000
LI	0.476314	0.092150	5.168891	0.0000

R-squared	0.830989	Mean dependent var	84.56080
Adjusted R-squared	0.815624	S.D. dependent var	29.70348
S.E. of regression	12.75438	Akaike info criterion	8.041794
Sum squared resid	3578.834	Schwarz criterion	8.188059
Log likelihood	-97.52242	Hannan-Quinn criter.	8.082361
F-statistic	54.08442	Durbin-Watson stat	2.397522
Prob(F-statistic)	0.000000		

Dependent Variable: LQI
Method: Least Squares
Date: 02/08/14 Time: 20:44
Sample: 1 25
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.481079	0.128619	19.29016	0.0000
LKI	0.640111	0.034731	18.43063	0.0000
LLI	0.257335	0.026959	9.545404	0.0000

R-squared	0.941558	Mean dependent var	4.375734
Adjusted R-squared	0.936245	S.D. dependent var	0.365243
S.E. of regression	0.092223	Akaike info criterion	-1.817045
Sum squared resid	0.187112	Schwarz criterion	-1.670780
Log likelihood	25.71306	Hannan-Quinn criter.	-1.776477
F-statistic	177.2195	Durbin-Watson stat	2.544657
Prob(F-statistic)	0.000000		

Commentaires **:**

- Au regard des critères « AIC et SIC » (minimum) et des mesures « R^2 et F-stat » (maximum), pour ne s'intéresser qu'à ces éléments là, le modèle transformé « E11.2 » est plus optimal. Il s'en suit que ce modèle est globalement bon (*Prob F-stat*<5%) et tous les paramètres estimés sont statistiquement significatifs (tc>tt, avec tt=$t_{22}^{0.05} = 2.07$) ;
- Pour le modèle spécifié en coupe instantané/données individuelles (dans le cas d'espèce), la statistique de Durbin-Watson est de moindre importance ;
- L'élasticité de la production par rapport au capital est de « 0.64 » : 10% de changements intervenus dans le capital font varier la production de 6.4% dans le même sens) ;

- L'élasticité travail de la production est de « 0.26 » : la production augmente de 2.6% si l'on fait varier le travail de 10% dans le sens positif ;
- La technologie (paramètre « a_0 ») se détermine comme suit : $a_0 = 10^{2{,}48} = 301{,}995$. L'expression « E11.1 » estimé peut s'écrire :

$$\boldsymbol{\hat{Q}_i = 301.995 K_i^{0.64} L_i^{0.26}}$$

b) *Test sur l'hypothèse des rendements constants*

Rappels *(rendement : sensibilité de la production à la variation conjointe de tous ses facteurs) :*

- *Si $a_1 + a_2 > 1$: les rendements sont dits croissants (la production croit plus vite/plus que proportionnellement que les facteurs de production : Zone de démarrage/Zone 1) ;*
- *Si $a_1 + a_2 = 1$: les rendements sont dits constants (la production croit au même rythme/proportionnellement aux facteurs de production : Zone de profit maximum/Zone 2) ;*
- *Si $a_1 + a_2 < 1$: les rendements sont dits décroissants (la production croit moins vite/moins que proportionnellement que les facteurs de production : Zone de fermeture/Zone 3).*

Analyse : dans notre cas, « $a_1 + a_2 = 0.9$ » : rendements décroissants si et seulement si cette valeur est significativement inférieure à 1 ; Si non, les rendements seront dits constants. Pour vérifier la significativité de cette valeur, les hypothèses à tester sont :

H0 : $a_1 + a_2 = 1$: les rendements sont constants (tc<tt ; prob>5%) ;
H1 : $a_1 + a_2 < 1$: les rendements sont décroissants (tc>tt ; prob<5%)

Avec « t de student calculé – tc exprimé comme suit :

$$t_c = \frac{\hat{a}_1 + \hat{a}_2 - 1}{\hat{\sigma}_{a1+a2}}$$

NB : $\hat{\sigma}^2_{a1+a2} = \hat{\sigma}^2_{a1} + \hat{\sigma}^2_{a2} + 2\,cov(\hat{a}_1, \hat{a}_2) = 0.001206 + 0.000727 + 2 * 0.000300$
$\hat{\sigma}^2_{a1+a2} = 0.002533$ (Cfr matrice de variances-covariances ci-dessous). Partant, $\hat{\sigma}_{a1+a2} = \sqrt{\hat{\sigma}^2_{a1+a2}} = \sqrt{0.002533} = 0{,}05033$

Ainsi, « tc » se calcul comme suit :

$$t_c = \frac{\hat{a}_1 + \hat{a}_2 - 1}{\hat{\sigma}_{a1+a2}} = \frac{0.640111 + 0.257335 - 1}{0.050} = \frac{0.897446 - 1}{0.050} = \frac{-0.103}{0.050}$$

$$t_c = -2.06 \approx t_{22}^{0.020}$$

Conclusion : *Nous acceptons H1 avec 2% de risque de se tromper (rejeter H0 à tort : erreur de 1ère espèce) : les rendements sont décroissants ($a_1 + a_2 < 1$).*

NB : Sur **EViews**, dans l'output des résultats, suivre : *View/Covariance Matrix* : obtenir la matrice de variances-covariances ci-dessous :

Coefficient Covariance Matrix			
	C	LKI	LLI
C	0.016543	-0.002869	-0.003187
LKI	-0.002869	0.001206	0.000300
LLI	-0.003187	0.000300	0.000727

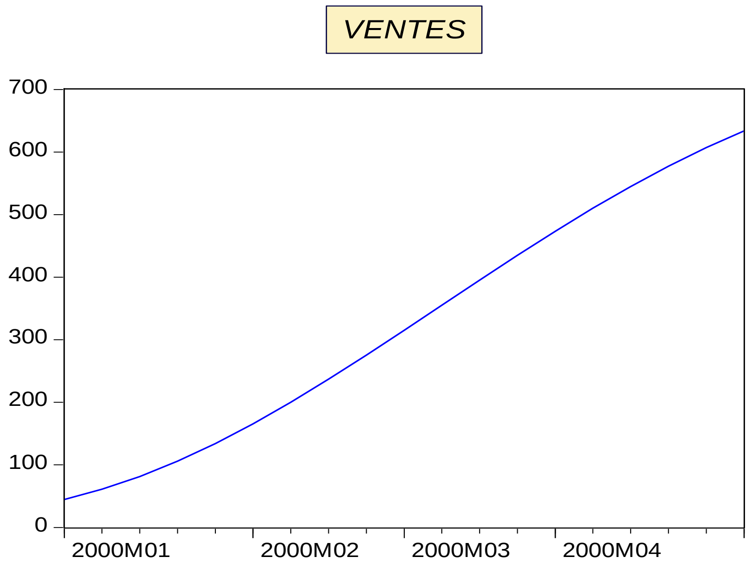
XII. Estimation d'un modèle de diffusion de type Logistique (avec EViews 6) : Feuille « E12 »

- ***TD*** : Sur base de données « **E12** », il est demandé de représenter graphiquement l'évolution des ventes cumulées sur 19 semaines (commenter) et d'estimer un modèle de type logistique y relatif suivant les approches ci-dessous :
  - *Par analogie historique : se référer aux précédentes ventes de CD du même artiste et faire l'hypothèse d'un seuil de saturation (ventes maximums) à 800 mille unités ;*
  - *Par la méthode de Balayage ;*
  - ***A l'aide de la fonction « SOLVER » du tableur Excel (not yet).***
- ***Solution :***

a) Evolution graphique de ventes cumulées sur 19 semaines et commentaire

Commandes EViews → *create w* ***1/1/2000 5/3/2000*** / *data VENTES* / *plot VENTES*

Résultats dans le Workfile :
*1/1/2000 5/**12**/2000*



Commentaire : A la lecture de ce graphique, visiblement les ventes semblent évoluer suivant un modèle de diffusion de type Logistique.

b) Estimation d'un modèle de type logistique des ventes par analogie « historique »

Hypothèse : le seuil de saturation/ventes maximales est à 800 mille unités ;

Modèle : Soit le modèle de diffusion de type logistique, expliquant les ventes, spécifié comme suit :

$$v_t = \frac{v_{max}}{1 + br^t} \quad \ldots \ldots . \ldots \ldots . . [E12.1]$$

Avec :

- v_{max} : Seuil de saturation/ventes maximales (soit, $v_{max} = 800$) ;
- b et r : les paramètres du modèle ($0 < r < 1$) ;
- r : vitesse de diffusion ($r \to 0$: plus vite on atteindra l'asymptote) ;
- b : ordonnée à l'origine ;
- t : temps (en nombre de semaines) ;
- v_t : les ventes à la semaine t.

Pour obtenir la forme linéaire, procédons comme suit :

$$\frac{v_{max}}{V_t} = 1 + br^t \to \frac{v_{max}}{v_t} - 1 = br^t \to \ln\left(\frac{v_{max}}{v_t} - 1\right) = \ln(b) + t\ln(r)$$

En posant (à calculer) : $V_t = \ln\left(\frac{v_{max}}{v_t} - 1\right) = \ln\left(\frac{800}{v_t} - 1\right)$, le modèle linéaire obtenu est le suivant :

$$\boldsymbol{V_t = a_0 + a_1 t + \varepsilon_t \ldots \ldots \ldots \ldots [E12.2]}$$

Avec : $\ln(b) = a_0 \to \hat{b} = e^{a_0}$ et $\ln(r) = a_1 \to \hat{r} = e^{a_1}$.

Résultats d'estimation du modèle linéaire « E12.2 » par analogie :

Dependent Variable: V
Method: Least Squares
Date: 02/09/14 Time: 09:23
Sample: 1/01/2000 5/06/2000
Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.788297	0.054746	50.93144	0.0000
T	-0.224566	0.004802	-46.76960	0.0000

R-squared	0.992288	Mean dependent var	0.542632
Adjusted R-squared	0.991834	S.D. dependent var	1.268607
S.E. of regression	0.114635	Akaike info criterion	-1.394819
Sum squared resid	0.223402	Schwarz criterion	-1.295405
Log likelihood	15.25079	Hannan-Quinn criter.	-1.377995
F-statistic	2187.395	Durbin-Watson stat	0.166365
Prob(F-statistic)	0.000000		

<u>Commandes EViews</u>

genr v1=(800/vt)-1
genr V=log(v1)
genr t=@trend+1
equation eqlogistique.ls V c t

- *Ecrivons le modèle estimé comme suit (sous forme linéaire) :* $\boldsymbol{\hat{V}_t = 2.79 - 0.22t}$

- Sous sa forme de départ, le modèle estimé s'écrira :

$$v_t = \frac{v_{max}}{1 + br^t} = \frac{800}{1 + 16.25 \times 0.7988^t}$$

Avec :

- $\ln(b) = a_0 \rightarrow \hat{b} = e^{a_0} = e^{2.788297} = 16.25$; et
- $\ln(r) = a_1 \rightarrow \hat{r} = e^{a_1} = e^{-0.224566} = 0.7988.$

c) *Estimation d'un modèle de type logistique des ventes par la méthode du Balayage*

Rappels : lorsqu'on recourt à « *la méthode du Balayage* », les étapes à suivre sont :

(i) Déterminer/définir un intervalle de saturation probable « $[v_{min} ; v_{max}]$. Dans le cas d'espèce, nous avons défini : $\boldsymbol{v_{min} = 680}$ ***et*** $\boldsymbol{v_{max} = 990}$.

(ii) Déterminer un pas « r » qui oriente les itérations de v_{min} à v_{max}. Généralement $r = [(v_{max} - v_{min})/10]$. Dans notre cas, $\boldsymbol{r = 10}$;

(iii) Ensuite, procéder à des estimations du modèle linéarisé/régression sur le temps « E12.2 » (de façon itérative : 10 itérations dans notre cas), comparer les sommes des carrés des résidus/SCR des différents modèles estimés, et retenir « *le seuil de saturation optimal* » relatif au modèle estimé qui minimise la SCR. Ayant procédé ainsi, nous avons trouvé comme seuil optimal : $v_{max} = 710$ (le modèle optimal associé offre une $SCR = 0.0889$).

Résultats d'estimation du modèle linéaire « E12.2 » par la méthode du Balayage, avec $v_{max} = 710$:

Dependent Variable: VB
Method: Least Squares
Date: 02/09/14 Time: 11:12
Sample: 1/01/2000 5/06/2000
Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.798395	0.034540	81.01991	0.0000
T	-0.253858	0.003029	-83.80036	0.0000

R-squared	0.997585	Mean dependent var	0.259811
Adjusted R-squared	0.997443	S.D. dependent var	1.430269
S.E. of regression	0.072324	Akaike info criterion	-2.316018
Sum squared resid	0.088923	Schwarz criterion	-2.216604
Log likelihood	24.00217	Hannan-Quinn criter.	-2.299193
F-statistic	7022.501	Durbin-Watson stat	0.331189
Prob(F-statistic)	0.000000		

Commandes EViews

genr v2=(710/vt)-1
genr VB=log(v2)
equation eqbalayage.ls VB c t

- *Le modèle estimé s'écrit (sous forme linéaire) :* $\boldsymbol{\widehat{VB}_t = 2.798 - 0.25t}$
- Sous sa forme de départ, le modèle estimé s'écrira :

$$v_t = \frac{v_{max}}{1 + br^t} = \frac{710}{1 + e^{2.798} \times e^{-0.25t}} = \frac{710}{1 + 16.42 \times 0.776^t}$$

XIII. Estimation des paramètres d'un modèle de GOMPERTZ par analogie et par l'ALGORITHME DE GAUSS-NEWTON (avec EViews 6) : Feuille « E12 Bis»

- ***TD*** : Sur base de données « **E12 Bis** », il est demandé d'estimer un *modèle du type Gompertz* suivant les approches ci-dessous :
  - *Par analogie historique : se référer aux précédentes ventes de CD du même artiste et faire l'hypothèse du volume total des ventes pour 800 mille unités ;*
  - *Par une méthode d'optimisation non linéaire.*
- ***Solution*** :

Modèle de Gompertz : soit le modèle de diffusion de type Gompertz, expliquant les ventes, spécifié comme suit :

$$v_t = e^{br^t+a} \quad \dots \dots . \dots \dots . . [E12.3]$$

Avec :

- e : base du logarithme népérien (exponentiel) ;
- b et r : les paramètres du modèle ($b < 0$ et $0 < r < 1$) ;
- $e^a = v_{max}$: seuil de saturation (dans notre cas/total des ventes : 800) ;
- t : temps (en nombre de semaines) ;
- v_t : le volume total des ventes à la semaine t.

Pour obtenir la forme linéaire, procédons comme suit :

$$\frac{e^a}{v_t} = e^{-br^t} \rightarrow \ln\left(\frac{e^a}{v_t}\right) = -br^t \rightarrow \ln\left[\ln\left(\frac{e^a}{v_t}\right)\right] = \ln(-b) + t\ln(r)$$

En posant (à calculer) : $V_t = \ln\left[\ln\left(\frac{e^a}{v_t}\right)\right] = \ln\left[\ln\left(\frac{800}{v_t}\right)\right]$; $\ln(-b) = a_0 \rightarrow \hat{b} = -e^{a_0}$ et $\ln(r) = a_1 \rightarrow \hat{r} = e^{a_1}$, le modèle linéaire simple de régression sur le temps obtenu est le suivant :

$$\boldsymbol{V_t = a_0 + a_1 t + \varepsilon_t \dots \dots \dots \dots [E12.4]}$$

Résultats d'estimation du modèle linéaire « E12.4 » par analogie :

Dependent Variable: VG
Method: Least Squares
Date: 02/09/14 Time: 13:17
Sample: 1/01/2000 5/06/2000
Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.272661	0.021490	59.22169	0.0000
T	-0.139005	0.001885	-73.75118	0.0000

R-squared	0.996884	Mean dependent var	-0.117385
Adjusted R-squared	0.996701	S.D. dependent var	0.783444
S.E. of regression	0.044998	Akaike info criterion	-3.265076
Sum squared resid	0.034423	Schwarz criterion	-3.165661
Log likelihood	33.01822	Hannan-Quinn criter.	-3.248251
F-statistic	5439.236	Durbin-Watson stat	0.148757
Prob(F-statistic)	0.000000		

Commandes EViews

genr VG=log(log(800/vt))
equation eqgompertz.ls VG c t

- *Le modèle estimé s'écrit (sous forme linéaire) :* $\widehat{\boldsymbol{VG}}_t = \mathbf{1.27 - 0.139}t$
- Sous sa forme de départ, le modèle estimé s'écrira :

$$v_t = e^{br^t + a} \to v_t = e^{(-3.57 \times 0.87^t + 2.90)}$$

Avec :

- $\ln(-b) = a_0 \to \hat{b} = -e^{a_0} = -e^{1.272661} = -3.57$; et
- $\ln(r) = a_1 \to \hat{r} = e^{a_1} = e^{-0.139005} = 0.87$;
- $e^a = 800 \to a = \log(800) = 2.90$.

XIV. Estimation des paramètres d'une fonction de production du type CES/Constant Elasticity of Substitution (avec EViews 6) : Feuille « E11 »

- ***TD*** : Sur base de données « **E11** », il est demandé ce qui suit :
  - *(i) Obtenir la forme linéaire de la fonction CES en passant par le logarithme ;*
  - *(ii) Estimer les paramètres du modèle linéaire obtenu par les MCO ;*
  - *(iii) Comparer les résultats obtenus sur CES avec ceux obtenus sur Cobb-Douglass.*
- ***Solution*** :

a) *Rappels*

- Elasticité de substitution : elle permet de saisir l'impact des variations relatives des prix des facteurs sur la technologie (combinaison productive ou rapport « K/L »). La fonction CES a été introduite par Arrow K. et al. (1961) ;
- Le modèle CES se présente sous la forme ci-après :

$$Q = b_1[b_2K^{b_3} + (1 - b_2)L^{b_3}]^{b_4} \dots \dots \dots \dots [E11.3]$$

Autrement, il peut s'écrire :

$$Q = b_1[b_2K^{b_3} + (1 - b_2)L^{b_3}]^{v/b_4} \dots \dots \dots \dots [E11.3]^*$$

Avec :

- $\rho = 1/(1 - b_3)$: l'élasticité de substitution ;
- v : les rendements d'échelle ($\hat{v} = \hat{b}_3 \times \hat{b}_4$) ;
- $b_3 < 0 \;\; et \;\; b_4 < 0$;
- b_1 ; b_2; $b_3 \; et \; b_4$: les paramètres à estimer ;
- Q : production ; K : facteur capital et L : facteur travail.

b) *Transformation/Linéarisation en passant par le logarithme*

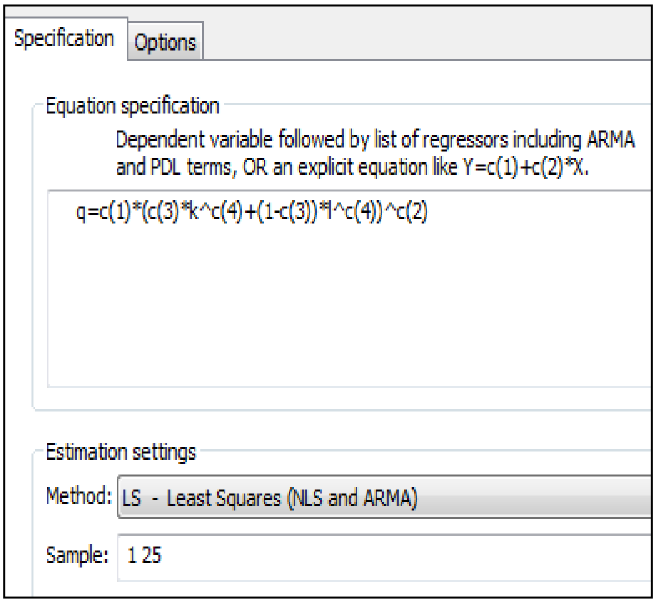
Astuce : Introduisons le logarithme dans « E11.3 », on obtient :

$$\log Q = \log(b_1) + b_4 \log[b_2K^{b_3} + (1 - b_2)L^{b_3}] \dots \dots \dots [E11.4]$$

Constat : le modèle « E11.4 » obtenu n'est pas linéaire malgré le passage au logarithme. Nous devons recourir à une méthode d'estimation des modèles non linéaires.

c) *Estimation des paramètres du modèle « E11.3 »*

Sur ***EViews***, taper (pour initialiser les 4 paramètres) : *param 1 12. 2 -1. 3 0.5 4 -0.5 ;* Ensuite, suivre : *Quick/Estimate Equation…* →la boîte de dialogue à gauche apparaît (*elle complète la procédure*) :



Dependent Variable: Q
Method: Least Squares
Date: 02/09/14 Time: 15:01
Sample: 1 25
Included observations: 25
Convergence achieved after 1 iteration
Q=C(1)*(C(3)*K^C(4)+(1-C(3))*L^C(4))^C(2)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	11.21349	1.425953	7.863854	0.0000
C(3)	0.405279	0.143586	2.822545	0.0102
C(4)	-0.596294	0.301700	-1.976445	0.0614
C(2)	-1.387204	0.765904	-1.811199	0.0844

R-squared	0.953721	Mean dependent var	84.56080
Adjusted R-squared	0.947110	S.D. dependent var	29.70348
S.E. of regression	6.831154	Akaike info criterion	6.826511
Sum squared resid	979.9581	Schwarz criterion	7.021531
Log likelihood	-81.33139	Hannan-Quinn criter.	6.880601
Durbin-Watson stat	2.555502		

Soit, dans la barre des commandes ***EViews****, taper (obtenir le même résultat) :*

param 1 12. 2 -1. 3 0.5 4 -0.5
nls Q=c(1)(c(3)*K^c(4)+(1-c(3))*L^c(4))^c(2)*

Le modèle E11.3 « $Q = b_1[b_2K^{b_3} + (1 - b_2)L^{b_3}]^{b_4}$ » estimé s'écrit alors :

$$\hat{Q} = 11.21[0.405K^{-0.596} + (1 - 0.405)L^{-0.596}]^{-1.38}$$

Ce qui revient à écrire :

$$\boldsymbol{\hat{Q} = 11.21[0.405K^{-0.596} + 0.595L^{-0.596}]^{-1.387}}$$

d) *Comparaison des résultats obtenus sur CES avec ceux obtenus sur Cobb-Douglass*

Notons ce qui suit :

- *Les résultats d'estimation de la fonction Cobb-Douglass nous ont renseignés que les rendements d'échelle étaient significativement décroissants (soit : $a_1 + a_2 = 0.9 < 1$). Avec la fonction CES, nous aboutissons à la même conclusion, du fait que : $\hat{v} = \hat{b}_3 \times \hat{b}_4 = -0.596 \times -1.387 = 0.83 < 1$;*
- *L'élasticité de substitution constante (Cfr CES) est égale à :*

$$\hat{\rho} = 1/\left(1 - \hat{b}_3\right) = 1/[1 - (-0.596)] = 1/1.596 = 0.63$$

- *L'on notera que la fonction Cobb-Douglass est commode pour l'étude du secteur secondaire, pendant que la fonction CES traduit mieux les réalités du* secteur tertiaire (en passant, le secteur primaire est souvent étudié à travers une fonction de Leontief).

***Annexes* : Programmes pour l'estimation des modèles non linéaires**

```
smpl 1979 2010
genr tend = @trend(1969)
smpl 1979 1997
scalar ymax = 800
genr Y = log(ymax/taux-1)
equation EQ1.ls Y c tend
scalar b2 = exp(c(1))
scalar r2 = exp(c(2))
'
' Question 3
'
scalar somcr=9999999
for !i= 650 to 980 step 10
  scalar ymax=!i
  genr Y=log(ymax/taux-1)
  equation EQ.ls Y c tend
  if @ssr<somcr then scalar somcr = @ssr
                                    scalar ymaxf = ymax
 endif
next
genr Y = log(ymaxf/taux-1)
equation balay.ls Y c tend
scalar b3=exp(C(1))
scalar r3=exp(C(2))
genr yprev = ymaxf/(1+b3*r3^tend)
'
' Méthode d'estimation non linéaire et prévsions avec intervalle de confiance
'
PARAM 1 600. 2 5. 3 0.5
equation EQNL.LS taux=C(1)/(1+c(2)*c(3)^tend)
EQNL.FORCST TAUXF ET
GENR IC1 = TAUXF - 2.3* ET
GENR IC2 = TAUXF + 2.3* ET
' Modèle de Gompertz
param 1 5. 2 -3. 3 0.5
equation EQGOM.ls taux = exp(c(1)+c(2)*c(3)^tend)
scalar YMAXG = EXP(C(1))
```

Source : Bourbonnais (Programmes téléchargeables).