

LESAUX Loïc
MAROT Gildas
TANGUY Brewal



Économétrie

-

Une Étude de la Création d'Entreprise entre 1994 et 2007

Charpentier Arthur
Cadoret Isabelle

Semestre 2 2008 Master 1

Introduction	3
Présentation des données	3
1) Créations et défaillances d'entreprises	3
2) Taux Longs.....	4
3) Taux de Chômagés par tranches d'âges	4
MODELE STATIQUE	5
1) Régression du modèle par les MCO.....	5
2) Test de la stabilité du modèle.....	6
3) Test de l'hétéroscédasticité	9
A) <i>Test de White</i>	9
B) <i>Test de White modifié</i>	10
C) <i>Test de Goldfeld-Quandt</i>	10
4) Autocorrélation des aléas	11
A) <i>Test de Durbin-Watson</i>	11
B) <i>Correction avec la méthode de Cochrane Orcutt</i>	12
5) Résultat du modèle statique.....	13
MODELE DYNAMIQUE.....	15
1) Régression par les MCO	15
2) Autocorrélation des aléas	15
3) Résultats du modèle dynamique.....	16
CONCLUSION.....	17

Introduction

Nous avons décidé d'étudier la variation du nombre de création d'entreprises en France chaque trimestre entre 1994 et 2007. C'est un sujet très intéressant car la plupart des entreprises créées sont de petites entreprises qui sont un facteur très important de croissance économique. Nous avons retenu cinq variables qui pourraient expliquer la variation du nombre de création d'entreprise.

En premier lieux le nombre de défaillance d'entreprise, si il y a beaucoup d'entreprise qui disparaissent on est tenté de penser que les autres entrepreneurs vont être plus frileux à lancer leur propre affaire la où les autres ont échoué.

Nous avons également retenu le taux d'intérêt car une création d'entreprise nécessite la plupart du temps un emprunt en plus d'un apport personnel. Il est plus tentant de créer son entreprise s'il est facile d'emprunter à un taux intéressant.

Enfin partant du constat que beaucoup d'entreprises sont créé par des chômeurs ne trouvant pas leur bonheur sur le marché du travail et décidant donc de lancer leur propre activité nous avons retenu comme variables le taux de chômage. Nous utiliserons les taux de chômeages des moins de 25 ans, des 25-49 ans et des plus de 50 ans afin de cerner quelle tranche d'âge est la plus créatrice d'entreprises.

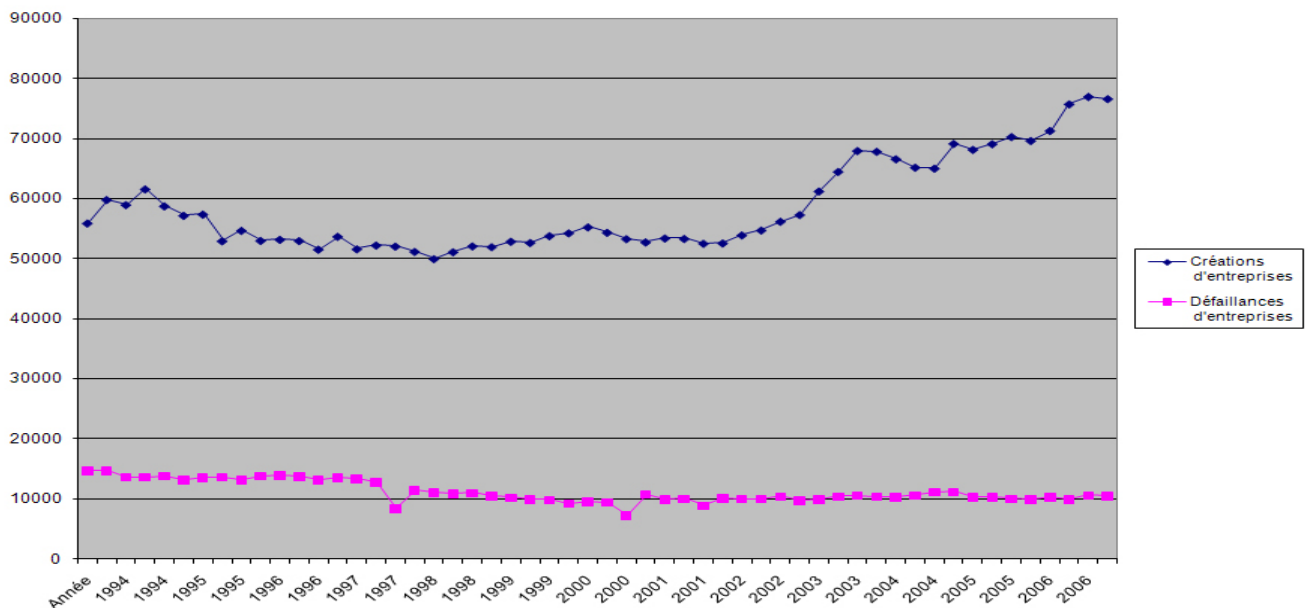
Toutes nos données sont issues du site de l'INSEE à l'exception du taux long qui est issue du site de la Banque Centrale Européenne.

Présentation des données

1) Créations et défaillances d'entreprises

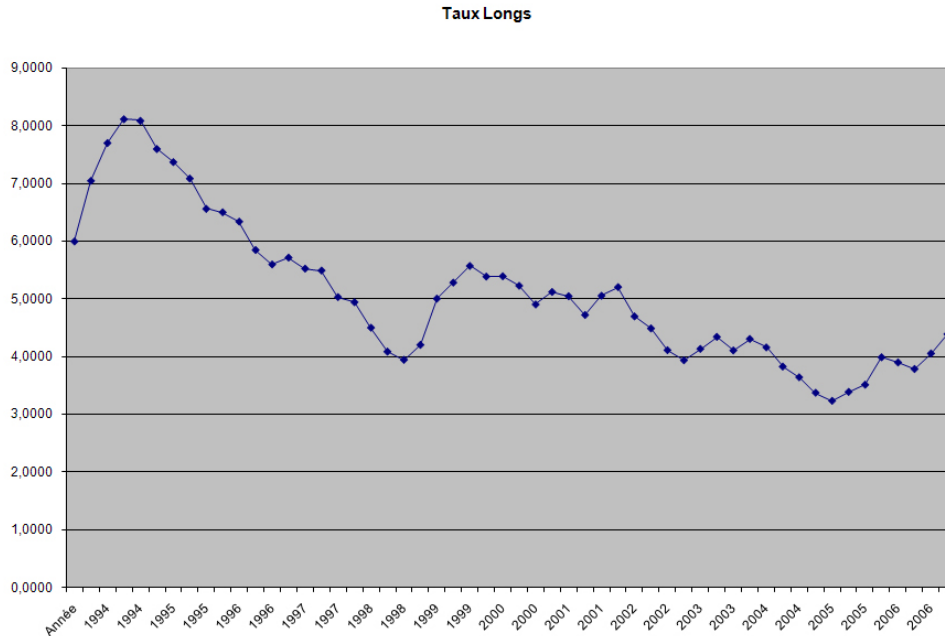
On observe que le nombre de création d'entreprise est relativement stable de 1994 à 2003, il reste compris entre 50000 et 60000 entreprise crée chaque trimestre. Au premier trimestre 2003 le nombre de création d'entreprise commence à s'accroître pour atteindre plus de 77000 création trimestrielle en 2006. Le nombre de défaillance d'entreprise observe une baisse sur la période étudié, d'environ 15000 créations par trimestre en 1994 à environ 10000 en 2007.

Créations et Défaillances d'Entreprises



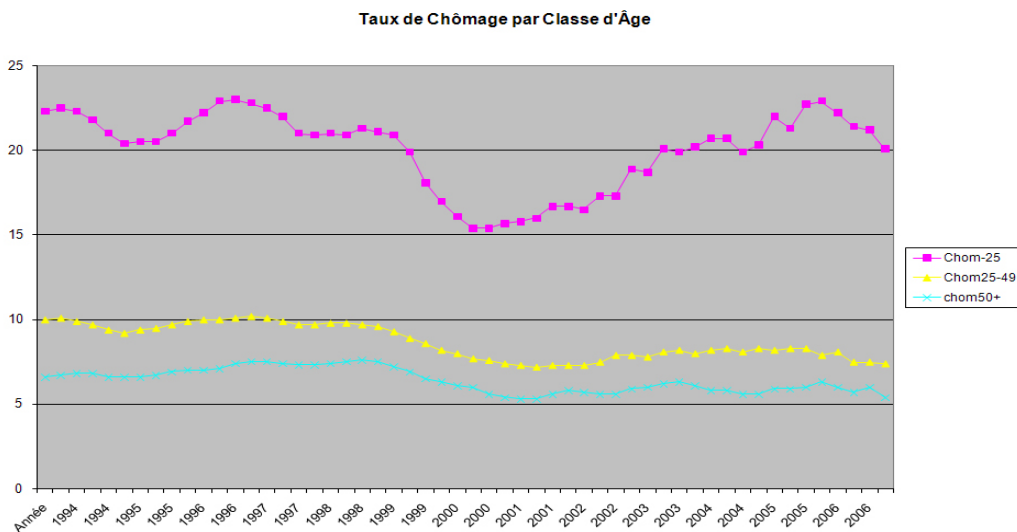
2) Taux Longs

Le taux d'intérêt augmente jusqu'à plus de 8% durant l'année 1994 pour baisser jusqu'à moins de 4% fin 1998 ou il réentame une hausse. On remarque ensuite une tendance à la baisse jusqu'en 2005 puis à la hausse jusqu'en 2007.



3) Taux de Chômagés par tranches d'âges

Le chômage observe des fluctuations régulières. On dégage cependant une stagnation entre 1994 et 1999 ou se produit une importante baisse du taux de chômage jusqu'en 2000. La tendance est ensuite à la hausse jusqu'en 2006 ou commence une période de baisse du taux de chômage. Ces variations sont vraies pour nos trois tranches d'âges, cependant les moins de 25 ans subissent des variations de plus grande amplitude que les deux autres tranches. De plus si le chômage des plus de 50 ans est compris entre 5,3 et 7,6% entre 1994 et 2007, celui des 25-49 ans est compris entre 7,2 et 10,2% et celui des moins de 25 ans entre 15,4 et 23%. Nous avons donc une grande différence de proportion et d'amplitude des variations entre les taux de chômage des différentes tranches d'âges.



MODELE STATIQUE

1) Régression du modèle par les MCO

Par commodité et par soucis de clarté nous avons renommé les variables

Chômage des moins de 25 ans = ChomA

Chômage des 25-49 ans = ChomB

Chômage des plus de 50 ans = ChomC

Défaillance d'entreprise = Defent

Taux longs = TxLg

Modèle 1 :

```
Linear Regression - Estimation by Least Squares
Dependent Variable CREA
Quarterly Data From 1994:01 To 2007:02
Usable Observations 54 Degrees of Freedom 48
Centered R**2 0.844206 R Bar **2 0.827977
Uncentered R**2 0.997470 T x R**2 53.863
Mean of Dependent Variable 58635.203704
Std Error of Dependent Variable 7604.441891
Standard Error of Estimate 3153.989051
Sum of Squared Residuals 477487052.71
Regression F(5,48) 52.0198
Significance Level of F 0.00000000
Log Likelihood -508.48940
Durbin-Watson Statistic 0.879980
```

Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	66723.97094	4507.05728	14.80433	0.00000000
2. DEFENT	-0.62362	0.57117	-1.09184	0.28035363
3. CHOMA	3917.93166	339.06901	11.55497	0.00000000
4. CHOMB	-5760.42935	1740.47690	-3.30968	0.00177745
5. CHOMC	-5273.89003	1903.22919	-2.77102	0.00792638
6. TXLG	791.48679	593.79598	1.33293	0.18884821

⇒ On remarque que les significativités individuelles des défaillances d'entreprises et des taux longs sont supérieurs à 5%.

Il convient alors d'effectuer un test de contrainte $\beta_2 = \beta_6 = 0$

Nous allons donc appeler le premier modèle vu précédemment le modèle non contraint.

Modèle Contraint

Modèle 2 :

```
Linear Regression - Estimation by Least Squares
Dependent Variable CREA
Quarterly Data From 1994:01 To 2007:02
Usable Observations 54 Degrees of Freedom 50
Centered R**2 0.837922 R Bar **2 0.828197
Uncentered R**2 0.997368 T x R**2 53.858
Mean of Dependent Variable 58635.203704
Std Error of Dependent Variable 7604.441891
Standard Error of Estimate 3151.975898
Sum of Squared Residuals 496747603.05
Regression F(3,50) 86.1642
```

Significance Level of F	0.00000000			
Log Likelihood	-509.55711			
Durbin-Watson Statistic	0.750118			
Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	67557.55145	4368.57929	15.46442	0.00000000
2. CHOMA	3627.87435	266.71168	13.60223	0.00000000
3. CHOMB	-5907.75628	1158.52249	-5.09939	0.00000527
4. CHOMC	-4757.33989	1512.85681	-3.14461	0.00279804

Test de Contrainte

Soit $H_0 : \beta_2 = \beta_6 = 0$
 $H_1 : \beta_2 \text{ et } \beta_6 \neq 0$

$$F_{cal} = \left(\frac{(SCR_{contra int} - SCR_{non_contra int}) / J}{SCR_{non_contra int} / (N - K)} \right)$$

Nous obtenons :

$$F_{cal} = 0,9681$$

Or $F(2,48) = 3,23$

La statistique de Fisher calculé est inférieure à la statistique de Fisher à 2 contraintes et 48 degrés de liberté.

⇒ On accepte H_0 , les paramètres des variables ‘Défaillance d’entreprise’ et ‘Taux longs’ ne sont pas significativement différent de 0

Nous retenons alors dans le modèle les variables explicatives ChomA, ChomB, et ChomC. Les variables ‘défaillances d’entreprise’ et ‘taux longs’ ne sont pas pertinentes dans la régression linéaire car non significatives.

Remarquons que la qualité de la régression est bonne car 83,79 % des valeurs de l’endogène sont expliquées par le modèle.

2) Test de la stabilité du modèle

Nous avons observé une rupture dans le nombre de création d’entreprise au cours du second trimestre 2003. Afin de déterminer si le modèle est stable, il convient d’effectuer un test de Chow.

Il s’agit ici de tester une forme différente de contrainte sur les paramètres. Le test de Chow permet de déterminer si les paramètres du modèle se sont modifiés au cours du temps. Lorsqu’il est appliqué sur des données temporelles, il est nécessaire de connaître a priori une date de rupture. Le test indique alors si les paramètres sont identiques avant et après la date de rupture.

$$H_0 = \begin{cases} \beta_1^1 = \beta_1^2 = \beta_1 \\ \beta_2^1 = \beta_2^2 = \beta_2 \\ \beta_3^1 = \beta_3^2 = \beta_3 \\ \beta_4^1 = \beta_4^2 = \beta_4 \end{cases}$$

L’hypothèse H_0 correspond à l’hypothèse de stabilité des paramètres sur les deux sous-périodes.

$$H_1 = \begin{cases} \beta_1^1 \neq \beta_1^2 \neq \beta_1 \\ \beta_2^1 \neq \beta_2^2 \neq \beta_2 \\ \beta_3^1 \neq \beta_3^2 \neq \beta_3 \\ \beta_4^1 \neq \beta_4^2 \neq \beta_4 \end{cases}$$

Sous l'hypothèse H_1 , les paramètres du modèle sont différents.

La première sous-période prise en compte est ici celle allant du premier trimestre 1994 au premier trimestre 2003, prenant en compte 37 observations. Concernant la dernière sous-période, les observations sont au nombre de 17, du second trimestre 2003 au second trimestre 2007.

L'hypothèse d'égalité des paramètres sur les deux sous-périodes sera refusée si la statistique de Fisher calculée est supérieure à celle lue dans la table, à savoir ici $F(4,46)$. Cette statistique est distribuée suivant une loi de Fisher-Snédecor à K et $N-2K$ degré de liberté ($F(K,N-2K)$). Dans le cas où H_0 serait refusée, il faudrait estimer le modèle sur les deux sous-périodes. Pour ce faire, la statistique à calculer est la suivante :

$$F_{cal} = \frac{\frac{\hat{\varepsilon}^{*1} \hat{\varepsilon}^{*2} - (\hat{\varepsilon}^{1'} \hat{\varepsilon}^{1'} + \hat{\varepsilon}^{2'} \hat{\varepsilon}^{2'})}{K}}{\frac{(\hat{\varepsilon}^{1'} \hat{\varepsilon}^{1'} + \hat{\varepsilon}^{2'} \hat{\varepsilon}^{2'})}{T-2K}}$$

Une fois le test effectué, Rats nous indique

Fstat	9.27637
F(4,46)=	9.27637 with Significance Level 0.00001393

Ainsi l'hypothèse H_0 selon laquelle le modèle est stable, à un seuil de 5%, a une probabilité de 0,00001393 de se réaliser.

⇒ On rejette l'hypothèse de stabilité du modèle, il y a donc bien une rupture au second trimestre 2003.

Il convient alors d'introduire une correction à l'aide de Dummy noté 0 du premier trimestre 1994 au premier trimestre 2003, et noté 1 du second trimestre 2003 au second trimestre 2007.

On va multiplier les variables explicatrices par la Dummy, que nous allons noter

ChomA*Dummy = ChomADum

ChomB*Dummy = ChomBDum

ChomC*Dummy = ChomCDum

Modèle 3

Linear Regression - Estimation by Least Squares			
Dependent Variable CREA			
Quarterly Data From 1994:01 To 2007:02			
Usable Observations	54	Degrees of Freedom	46
Centered R**2	0.910287	R Bar **2	0.896635
Uncentered R**2	0.998543	T x R**2	53.921
Mean of Dependent Variable	58635.203704		
Std Error of Dependent Variable	7604.441891		
Standard Error of Estimate	2444.854804		
Sum of Squared Residuals	274956490.62		
Regression F(7,46)	66.6784		

Significance Level of F	0.00000000			
Log Likelihood	-493.58744			
Durbin-Watson Statistic	1.194937			
Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	59024.14424	3762.70590	15.68662	0.00000000
2. CHOMA	127.68089	677.23000	0.18853	0.85128737
3. CHOMB	3504.52551	1976.44954	1.77314	0.08282627
4. CHOMC	-5890.35963	1409.17999	-4.17999	0.00012913
5. DUMMY	50201.37886	21122.72782	2.37665	0.02168786
6. CHOMADUM	2776.23522	853.42866	3.25304	0.00214245
7. CHOMBDUM	-12252.46550	2893.26007	-4.23483	0.00010841
8. CHOMCDUM	636.24592	2939.98902	0.21641	0.82962428

⇒ Les significativités individuelles de ChomA et ChomCDum sont supérieur à 5%, on les enlève donc et nous pouvons refaire une régression.

Modèle 4

Linear Regression - Estimation by Least Squares				
Dependent Variable CREA				
Quarterly Data From 1994:01 To 2007:02				
Usable Observations	54	Degrees of Freedom	48	
Centered R**2	0.910113	R Bar **2	0.900750	
Uncentered R**2	0.998540	T x R**2	53.921	
Mean of Dependent Variable	58635.203704			
Std Error of Dependent Variable	7604.441891			
Standard Error of Estimate	2395.706136			
Sum of Squared Residuals	275491578.75			
Regression F(5,48)	97.2005			
Significance Level of F	0.00000000			
Log Likelihood	-493.63994			
Durbin-Watson Statistic	1.194727			
Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	58744.99639	3570.21768	16.45418	0.00000000
2. CHOMB	3735.99057	835.96272	4.46909	0.00004785
3. CHOMC	-5781.93350	1198.59855	-4.82391	0.00001465
4. DUMMY	52507.16216	18535.06932	2.83285	0.00672564
5. CHOMADUM	2922.79935	502.61523	5.81518	0.00000048
6. CHOMBDUM	-12396.32202	2133.36275	-5.81070	0.00000049

Le test de Chow sur le nouveau modèle nous donne alors

Fstat	2.06760
F(3,45)=	2.06760 with Significance Level 0.11789390

Le modèle est alors corrigé par les dummy et donc stable.

Interprétations du test de Chow :

Il y a donc une rupture au 2nd trimestre 2003 au niveau du chômage des moins de 25 ans et du chômage des 25 à 49 ans. D'ailleurs le taux de chômage des moins de 25 ans n'entre dans le modèle qu'à partir de la date de rupture.

Par ailleurs la qualité de la régression est meilleure avec les Dummy car 91% des variables endogènes sont expliquées par le modèle retenu.

3) Test de l'hétéroscédasticité

A) Test de White

Le test de White est un test d'homoscédasticité fondé sur l'existence d'une relation entre le carré du résidu et une ou plusieurs variables explicatives en niveau ou au carré. On va pouvoir voir ici s'il y a un problème d'hétéroscédasticité des aléas. On estime alors le modèle suivant avec les MCO :

$$e_i^2 = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_2^2 + \beta_5 x_3^2 + \beta_6 x_2 x_3 + \varepsilon_i$$

où e_i sont les résidus du modèle, et x_2 et x_3 sont les variables explicatives du modèle.

On va estimer ici un modèle dans lequel la variable endogène est e_i^2 . Les variables explicatives vont être égales aux variables explicatives du modèle de base, additionnés du carré des mêmes variables, et des produits croisés des variables.

Si au moins un des coefficients de régression est significatif alors on rejette l'hypothèse nulle d'homoscédasticité en faveur de l'hypothèse alternative d'hétéroscédasticité.

Pour effectuer ce test, on utilise la statistique LM TR^2 , où T est le nombre d'observation et R^2 est le coefficient de détermination de la régression précédente.

Sous l'hypothèse nulle d'homoscédasticité, on a :

$$TR^2 \sim \chi^2_{(1-\alpha)}(2p)$$

où p est le nombre de régresseurs excluant la constante.

La règle de décision est la suivante :

- Si $TR^2 < \chi^2_{0,95}(2p)$, alors l'hypothèse nulle d'homoscédasticité est acceptée au seuil de 5%.
- Si la p-value est inférieure à 0.05, alors on rejette H_0 au seuil de 5% et donc les résidus sont hétéroscédastiques.

L'inconvénient de ce test est qu'il ne propose pas de solutions en cas de détection d'un problème.

Nous devons alors créer de nouvelles variables

```
SET resid2 = resid**2
SET ChomB2 = ChomB**2
SET ChomC2 = ChomC**2
SET Dummy2 = Dummy**2
SET ChomADum2 = ChomADum**2
SET ChomBDum2 = ChomBDum**2
SET ChomBC = ChomB*ChomC
SET ChomADumB = ChomADum*ChomB
SET ChomADumC = ChomADum*ChomC
SET ChomADumDum = ChomADum*Dummy
SET ChomADumBDum = ChomADum*ChomBDum
SET ChomBDumB = ChomBDum*ChomB
SET ChomBDumC = ChomBDum*ChomC
SET ChomBDumDum = ChomBDum*Dummy
```

Nous pouvons alors effectuer une régression par la méthode des MCO :

Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	174048010	185513803	0.93819	0.35363878
2. CHOMB	-163064829	60219358	-2.70785	0.00983173
3. CHOMC	172910535	73317446	2.35838	0.02320424
4. DUMMY	7718410526	3653622537	2.11254	0.04077534

5. CHOMADUM	-10169550	105346022	-0.09653	0.92356623
6. CHOMBDUM	-1941170752	759452196	-2.55601	0.01439082
7. CHOMB2	3690591	7526994	0.49031	0.62652547
8. CHOMC2	-22011879	16654110	-1.32171	0.19359297
9. DUMMY2	0	0	0.00000	0.00000000
10. CHOMBC	13774930	20754163	0.66372	0.51058542
11. CHOMADUMB	-7487932	9752289	-0.76781	0.44699816
12. CHOMADUMC	10844682	9910560	1.09426	0.28023022
13. CHOMADUMDUM	0	0	0.00000	0.00000000
14. CHOMADUMBDUM	0	0	0.00000	0.00000000
15. CHOMBDUMB	142696382	50584649	2.82094	0.00734768
16. CHOMBDUMC	-25545411	25003686	-1.02167	0.31292986
17. CHOMBDUMDUM	0	0	0.00000	0.00000000
Statistique de White	25.05128			
Chi-Squared(16)=	25.051276	with Significance Level	0.06892661	

Ce test de White, avec un niveau de significativité à 0,0689 nous indique que l'hypothèse H0 d'homoscédasticité des aléas ne peut pas être rejetée. Il faut alors passer au test de White modifié, en gardant les variables significatives, mais sans inclure les dummy, car si on les garde pour ordonner le test de Goldfeld-Quandt, une grande partie des données sont égales à 0, donc il n'est pas intéressant de les garder.

Nous garderons donc pour le test de White modifié les variables ChomB et ChomC.

B) Test de White modifié

Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	43602389.60	12807285.80	3.40450	0.00129887
2. CHOMB	-7211859.23	3353059.32	-2.15083	0.03625102
3. CHOMC	4416513.75	4848830.61	0.91084	0.36666426
Statistique de White	9.72481			
Chi-Squared(2)=	9.724806	with Significance Level	0.00773188	

Le test de White modifié nous indique par contre la présence d'aléas hétéroscédastique. Pour vérifier ce fait nous allons devoir procéder au test de Goldfeld-Quandt.

C) Test de Goldfeld-Quandt

Une forme plus ancienne des tests d'hétéroscédasticité est donnée par le test de Goldfeld-Quandt. Nous allons donc tester ici d'une manière différente l'hétéroscédasticité des aléas. L'hypothèse H₀ de ce test est identique à celle du test de white, seule H₁ change (ici var(ε_i) = σ²). Dans un premier lieu, il s'agira de trier les observations d'une variable par ordre croissant. Par la suite on sépare l'échantillon en deux parties, en enlevant les observations centrales. Il reste alors pour notre modèle deux échantillons, chacun ayant (N-C)/2 observations. On réalise ensuite un test d'égalité de variance des deux échantillons.

On se retrouve alors face à deux possibilités :

- Si $\hat{\sigma}_2^2 > \hat{\sigma}_1^2$, alors $F_{cal} = \hat{\sigma}_2^2 / \hat{\sigma}_1^2 \sim F((N-C)/2 - K, (N-C)/2 - K)$
- Si $\hat{\sigma}_2^2 < \hat{\sigma}_1^2$, alors $F_{cal} = \hat{\sigma}_1^2 / \hat{\sigma}_2^2 \sim F((N-C)/2 - K, (N-C)/2 - K)$

La règle de décision est alors la suivante : Si $F_{cal} < F_{tabulé}$, alors on accepte H_0 , il n'y a pas d'hétéroscédasticité, sinon la variance des aléas est une fonction croissante de la variable en question.

Nous allons donc ici ordonner nos variables en fonction des variables retenues dans le test de White modifié.

- On ordonne en fonction de ChomB

Fstat	0.88792
F(17,20)=	0.88792 with Significance Level 0.59414111

⇒ L'hypothèse H_0 d'homoscédasticité des aléas est donc acceptée.

- On ordonne en fonction de ChomC

Fstat	0.88692
F(17,20)=	0.88692 with Significance Level 0.59506124

⇒ L'hypothèse H_0 d'homoscédasticité des aléas est donc acceptée.

En conclusion de ces tests nous pouvons alors affirmer que dans le modèle retenu les aléas sont homoscédastiques. Il n'y a donc pas besoin de corriger le modèle par la méthode des Moindres Carrés Généralisés (MCG).

Il reste tout de même à déterminer si dans ce modèle il y a une auto corrélation d'ordre 1 des aléas.

4) Autocorrélation des aléas

A) Test de Durbin-Watson

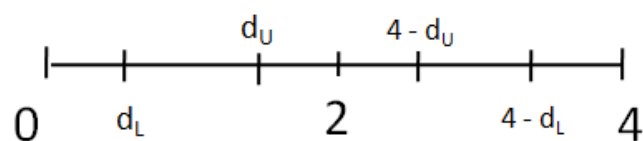
Le test de Durbin-Watson est un test très populaire qui permet de détecter la présence d'auto corrélation positive entre les résidus.

Pour ce test, nous avons ici la statistique suivante :

$$DW = \frac{\sum_{t=2}^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^T \hat{\varepsilon}_t^2}$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \quad \text{avec } u_t \sim N(0, \sigma_u^2)$$

Le Durbin Watson a une distribution qui oscille entre deux distributions, d_L et d_U , et est compris entre 0 et 4. Il se retrouve dans la représentation suivante :



Si la statistique DW est proche de 2 alors on accepte l'hypothèse de non autocorrélation d'ordre 1.

Afin d'être plus précis, la règle de décision est la suivante :

- Quand d_w est compris entre
- 0 et d_L , alors l'autocorrélation des résidus est positive.
 - d_U et $4-d_U$, il n'existe pas d'autocorrélation.
 - $4-d_L$ et 4, alors l'autocorrélation est négative.

En revanche, s'il se trouve entre d_L et d_U , ou encore entre $4-d_U$ et $4-d_L$, alors on est dans une zone d'indécision.

Par conséquent, s'il n'existe pas d'autocorrélation, on applique les MCO. Dans le cas contraire, ce sera les MCG. Si on se trouve dans la zone d'indécision, il faudra comparer les MCO et MCG, par une représentation graphique.

Nous avons alors retenu le modèle 4, et Rats nous donne la statistique de Durbin-Watson

$$D_w = 1,1947$$

Il faut alors nous reporter à la table de Durbin-Watson sachant que nous avons :

$$K' = 5 \text{ \& } N = 54$$

	dl	du	4-du	4-dl
N=50	1.34	1.77	2.33	2.66
N=55	1.38	1.77	2.33	2.62

La statistique de Durbin-Watson est donc inférieure à 'dl'

⇒ Autocorrélation positive des aléas

B) Correction avec la méthode de Cochrane Orcutt

Le principe de cette méthode est ici de rechercher $\hat{\rho}$ qui minimise la somme des carrés des résidus.

Après avoir estimé le modèle avec la méthode des MCO, on cherche $\hat{\rho}$ à l'aide de la formule suivante :

$$\hat{\rho} = \frac{\sum_{t=2}^T e_{t-1} \cdot e_t}{\sum_{t=2}^T e_{t-1}^2}$$

Par la suite, il suffit d'estimer le modèle transformé, et on procède par une procédure itérative, au calcul des résidus du modèle. Une fois que $\hat{\rho}$, ne varie plus, on s'arrête. On a alors minimisé la somme des carrés des résidus.

Modèle 5

Regression with AR1 - Estimation by Cochrane-Orcutt	
Dependent Variable CREA	
Quarterly Data From 1994:02 To 2007:02	
Usable Observations	53
Degrees of Freedom	46
Centered R**2	0.912278
R Bar **2	0.900836
Uncentered R**2	0.998566
T x R**2	52.924
Mean of Dependent Variable	58732.433962
Std Error of Dependent Variable	7643.251016
Standard Error of Estimate	2406.884424
Sum of Squared Residuals	266482261.01
Regression F(6,46)	79.7306

Significance Level of F	0.0000000			
Log Likelihood	-484.11269			
Durbin-Watson Statistic	2.018035			
Q(13-1)	9.427963			
Significance Level of Q	0.66600706			
Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	59671.83205	4237.38936	14.08222	0.00000000
2. CHOMB	3702.78103	882.58408	4.19539	0.00012295
3. CHOMC	-5869.82274	1319.84392	-4.44736	0.00005464
4. DUMMY	50580.63175	18116.85751	2.79191	0.00760682
5. CHOMADUM	2837.67113	492.92610	5.75679	0.00000067
6. CHOMBDUM	-11945.69370	2132.85050	-5.60081	0.00000114

7. RHO	0.14005	0.14785	0.94726	0.34845469

Le modèle 5 présent nous donne donc la régression de notre modèle classique corrigé des erreurs, ici l'instabilité et l'autocorrélation des aléas.

5) Résultat du modèle statique

Pour pouvoir interpréter les résultats, il faut alors se baser sur les résultats du modèle estimé par la méthode de Cochrane Orcutt, qui est le modèle corrigé des erreurs.

Nous obtenons alors le modèle suivant

(Les chiffres entre parenthèses sont les statistiques de Student)

$$\begin{aligned}
 Crea = & 59671.83 + 3702.78 \cdot c \text{ hom } B - 5869.82 c \text{ hom } C + 50580.63 \cdot Dummy \\
 & (14.08222) \quad (4.19539) \quad (-4.44736) \quad (2.79191) \\
 & + 2837.67 \cdot c \text{ hom } Adum - 11945.69 \cdot c \text{ hom } Bdum \\
 & (5.75679) \quad (-5.60081)
 \end{aligned}$$

Il est alors intéressant d'observer plusieurs points.

- Tout d'abord la qualité de la régression, notre R^2 est ici de 0,91, c'est-à-dire que 91% des créations d'entreprises sont expliquées par notre modèle.
- Le chômage des moins de 25 ans (ChomA) n'intervient qu'à partir de la date de rupture et de façon positive. En effet, la hausse d'un point de pourcentage du taux de chômage des jeunes entraînerait la création de 2837 entreprises supplémentaires avec un écart-type de 492.
- Le Chômage des plus de 50 ans (ChomC), intervient négativement sur notre endogène. La hausse d'un point de pourcentage du taux de chômage des seniors aurait pour effet une diminution de création d'entreprise de 5869 unités avec un écart-type de 1319.
- Le chômage des 25-49 ans (ChomB) est plus ambigu. En effet bien qu'il agisse positivement sur la création d'entreprises, avec une création de 3702 entreprises supplémentaires avec l'augmentation d'un point de pourcentage, avec un écart-type de 882, le taux de chômage agit de manière différente à partir de la rupture, car la hausse d'un point de pourcentage entraînerait la diminution de création d'entreprise de 11945 unités avec un écart-type de 2132.

De manière générale, seuls les taux de chômage rentrent en compte dans le modèle statique. On sait effectivement que la plus grande majorité des entreprises sont de très petites entreprises, souvent créées par des personnes au chômage.

De plus des recherches nous ont peut-être permis d'identifier la cause de la rupture. En effet la loi pour l'initiative économique du 1^{er} août 2003 permet la création d'une SARL sans capital minimum (plus connu comme la loi permettant la création de « l'entreprise à 1€ »). On peut donc penser que la stagnation de création légèrement avant, et la montée du nombre de création après cette date traduisent une attente de la part des futurs créateurs d'entreprise, car l'introduction de la Dummy agit positivement sur la création.

Il nous semble alors intéressant de constater que cela a permis aux plus jeunes chômeurs de se lancer dans une activité, sans apport de capital, conjugué à des aides comme l'ACCRE (Aide à la Création ou à la Reprise d'Entreprise). Cela permet d'expliquer la pertinence de ce taux de chômage dans notre modèle à partir de cette rupture, son niveau de significativité étant très faible (0.00000067).

Dans le cas des chômeurs de plus de 50 ans, cela peut s'expliquer par le peu de créateurs d'entreprises qui ont plus de la cinquantaine.

Le plus troublant concerne le chômage des 25-49 ans, qui agit de manière positive, mais lorsque l'on prend en compte la rupture, il agit de manière négative. Une classification plus fine aurait sans doute pu permettre d'identifier quelle tranche d'âge agit plutôt positivement et laquelle agit négativement.

MODELE DYNAMIQUE

1) Régression par les MCO

Nous intégrons donc dans le modèle dynamique dans les variables explicatrices, la variable endogène de la période t_1 , que nous nommons 'Crea_1'

Modèle 6 :

Variable	Coeff	Std Error	T-Stat	Signif

1. Constant	61231.18690	3699.98468	16.54904	0.00000000
2. CHOMB	3268.81447	847.68434	3.85617	0.00034919
3. CHOMC	-4271.90982	1401.03007	-3.04912	0.00376273
4. DUMMY	53560.31281	18030.32034	2.97057	0.00467116
5. CHOMADUM	2867.27756	489.54238	5.85706	0.00000044
6. CHOMBDUM	-12436.14717	2074.42918	-5.99497	0.00000027
7. CREA_1	-0.13651	0.07030	-1.94191	0.05815496

⇒ toutes les variables sont significatives.

2) Autocorrélation des aléas

Dans le modèle dynamique, Rats nous donne la statistique de Durbin-Watson
 $D_w = 1.945909$

Il faut alors nous reporter à la table de Durbin-Watson sachant que nous avons :
 $K' = 6$ & $N = 54$

	dl	du	4-du	4-dl
N=50	1.29	1.82	2.18	2.71
N=55	1.33	1.81	2.19	2.67

La statistique de Durbin-Watson est donc comprise entre du et $4 - du$

⇒ Il n'y a pas d'autocorrélation des aléas.

Par conséquent, il n'est pas nécessaire d'utiliser les variables instrumentales et le test d'Hausman.

3) Résultats du modèle dynamique

Nous obtenons ici le modèle suivant :

$$\begin{aligned} Crea = & 61231.18 + 3268.81 \cdot c \text{ hom } B - 4271.90 c \text{ hom } C + 53560.31 \cdot Dummy \\ & (16.54904) \quad (3.85617) \quad (-3.04912) \quad (2.97057) \\ + & 2867.27 \cdot c \text{ hom } Adum - 12436.14 \cdot c \text{ hom } Bdum - 0.13651 \cdot Crea_1 \\ & (5.85706) \quad (-5.99497) \quad (-1.94191) \end{aligned}$$

Nous obtenons donc quasiment le même modèle qu'en statique, avec la variable endogène de la période précédente en sus.

Nous pouvons déjà remarquer que les coefficients sont quasi identiques que ceux du modèle statique, il convient alors de s'intéresser à cette nouvelle variable « Crea_1 ».

La création d'une entreprise au trimestre précédent entraîne la diminution de la création d'entreprise au trimestre présent de 0.13 unité, avec un écart-type 0.0703. Ce coefficient est assez proche de 0, et de plus son niveau de significativité est de 5,8%, ce qui est légèrement supérieur à 5%. La statistique de Student de ce paramètre est, en valeur absolue, de 1,94191, et la statique de Student, à 5%, pour N=54, est égale à 1,95966.

On a alors $T_{cal} < T_{stat}$, le paramètre n'est donc pas significativement différent de 0.

⇒ Le modèle dynamique n'apporte donc rien de plus que le modèle statique.

CONCLUSION

L'étude économétrique du nombre de créations d'entreprise nous a permis d'identifier ses composantes, qui sont les différents taux de chômage par tranches d'âges. Nous avons donc écarté le nombre de défaillances d'entreprise et les taux longs, contre nos attentes ils n'étaient pas significatifs dans le modèle retenu.

Nous avons aussi pu localiser un choc, et nous avons tenté d'en identifier la cause. Nos recherches nous nous alors amenées à une loi sur l'initiative économique de 2003, ce dont nous pensons être la cause de cette rupture dans notre modèle.

Il pourra être assez intéressant de continuer ce sujet en séparant en plusieurs parties le chômage des 25-49 ans, pour affiner notre modélisation. Il en est de même de l'utilisation de nombres de chômeurs, et non des taux, afin de prendre en compte l'effet taille de ces variables.

