

Hétéroscédasticité des aléas : tests et estimation par MCG

Thème : Les dépenses de consommation financées par carte de crédit,

Fichiers à télécharger :

<http://perso.univ-rennes1.fr/isabelle.cadoret-david/econometrie/MCG/cartecredit.xls>

<http://perso.univ-rennes1.fr/isabelle.cadoret-david/econometrie/MCG/cartecreditrats.xls>

<http://perso.univ-rennes1.fr/isabelle.cadoret-david/econometrie/MCG/cartecredit.prg>

N.B. : Il vous est demandé de réaliser les estimations et les tests à l'aide d'EXCEL.

Les variables sont les suivantes (les observations sont reproduites dans l'**annexe 1**) :

- *Dépenses* : dépenses mensuelles effectuées avec la carte de crédit
- *Age* : âge de l'individu
- *Prop* : variable indicatrice (1 si l'individu est propriétaire de son logement, 0 sinon).
- *Revenu* : revenu de l'individu
- *Revenu2* : carré du revenu

Source : W.Greene, "Econometric Analysis" Prentice Hall International, 4ème édition, 2000

1. Estimation préalable par MCO et MV :

- i) Estimer avec la méthode des MCO les dépenses des individus en fonction des variables : *Age, Prop, Revenu et Revenu2* (préciser les lignes de commande RATS, cf. **annexe 2**)
- ii) Préciser comment sont calculés les coefficients de détermination, puis les résidus et leur variance (cf. aussi les résultats du programme RATS de l'**annexe 3**) . Commenter.
- iii) Tester la signification des paramètres et interpréter les résultats obtenus.
- iv) Représenter sur graphique le carré des résidus en fonction de la variable *revenu*. Conclure.

2. Analyse des résidus : tests d'hétéroscédasticité

- i) Conduire le test de White :
 - Ecrire l'équation de régression du carré des résidus.
 - Poser le problème de test.
 - Ecrire la statistique associée et en déduire sa distribution de probabilité.
 - Définir et représenter graphiquement la région critique pour un niveau $\alpha=5\%$.
 - Conclure.

- ii) Réaliser le test de White modifié en considérant que la variance des aléas est uniquement fonction de *Revenu, Revenu2 et Revrev2* (et d'un terme constant) sous l'hypothèse rivale.
- iii) Appliquer le test de Goldfeld-Quandt en utilisant les résultats des estimations présentés dans les tableaux 2 et 3 ci-après.
 - Trier l'échantillon en fonction de *Revenu* par ordre décroissant.
 - Ecrire l'équation de régression.
 - Retirer les 8 observations centrales, puis estimer le modèle complet sur les deux sous-échantillons ainsi obtenus.
 - Poser le problème de test.
 - Ecrire la statistique associée et en déduire sa distribution de probabilité sous l'hypothèse nulle.
 - Définir et représenter graphiquement la région critique pour un niveau $\alpha=5\%$.
 - Conclure.

3. Recherche d'estimateurs robustes

- i) Estimer le modèle avec la méthode des MCO tout en corrigeant le calcul de la matrice de variance covariance des paramètres (correction de White).
- ii) On suppose à présent que la variance des résidus proportionnelle à une variable explicative donnée.
 - Appliquer la méthode des MCG en admettant que la variance des résidus est fonction de la variable *Revenu*.
 - Réaliser l'estimation en supposant désormais que la variance est fonction de la variable *Revenu2*.
 - Commenter les résultats obtenus.

ANNEXE 1 : Echantillon d'observations

Isabelle Cadoret : Analyse des dépenses des ménages à partir des dépenses en carte de crédit

Source : Greene "Econometric Analysis"

	Depenses	Age	Prop	Revenu	Revenu2
1	124,98	38	1	4,52	20,4304
2	9,85	33	0	2,42	5,8564
3	15	34	1	4,5	20,25
4	137,87	31	0	2,54	6,4516
5	546,5	32	1	9,79	95,8441
6	92	23	0	2,5	6,25
7	40,83	28	0	3,96	15,6816
8	150,79	29	1	2,37	5,6169
9	777,82	37	1	3,8	14,44
10	52,58	28	0	3,2	10,24
11	256,66	31	1	3,95	15,6025
12	78,87	29	1	2,45	6,0025
13	42,62	35	1	1,91	3,6481
14	335,43	41	1	3,2	10,24
15	248,72	40	1	4	16
16	548,03	40	1	10	100
17	43,34	35	1	2,35	5,5225
18	218,52	34	1	2	4
19	170,64	36	0	4	16
20	37,58	43	1	5,14	26,4196
21	502,2	30	0	4,51	20,3401
22	73,18	22	0	1,5	2,25
23	1532,77	40	1	5,5	30,25
24	42,69	22	0	2,03	4,1209
25	417,83	29	0	3,2	10,24
26	552,72	21	1	2,47	6,1009
27	222,54	24	0	3	9
28	541,3	43	1	3,54	12,5316
29	568,77	37	1	5,7	32,49
30	344,47	27	0	3,5	12,25
31	405,35	28	1	4,6	21,16
32	310,94	26	1	3	9
33	53,65	23	0	2,59	6,7081
34	63,92	30	0	1,51	2,2801
35	165,85	30	0	1,85	3,4225
36	9,58	38	0	2,6	6,76
37	319,49	36	0	2	4
38	83,08	26	0	2,35	5,5225
39	644,83	28	1	7	49
40	93,2	24	0	2	4
41	105,04	21	0	1,7	2,89
42	34,13	24	0	2,8	7,84
43	41,19	26	0	2,4	5,76
44	169,89	33	0	3	9
45	1898,03	34	0	4,8	23,04
46	810,39	33	0	3,18	10,1124
47	32,78	21	0	1,5	2,25
48	95,8	25	0	3	9
49	27,78	27	0	2,28	5,1984
50	215,07	26	0	2,8	7,84
51	79,51	22	0	2,7	7,29
52	306,03	41	0	6	36
53	104,54	42	0	3,9	15,21
54	642,47	25	0	3,07	9,4249
55	308,05	31	1	2,46	6,0516
56	186,35	27	0	2	4
57	56,15	33	0	3,25	10,5625
58	129,37	37	0	2,72	7,3984
59	93,11	27	0	2,2	4,84
60	292,66	24	0	3,75	14,0625
61	98,46	25	0	2,88	8,2944
62	258,55	36	0	3,05	9,3025
63	101,68	33	0	2,55	6,5025
64	65,25	55	1	2,64	6,9696
65	108,61	20	0	1,65	2,7225
66	49,56	29	0	2,4	5,76
67	235,57	41	1	7,24	52,4176
68	68,38	43	0	2,4	5,76
69	474,15	33	1	6	36

ANNEXE 2 : PROGRAMME RATS

```

1.  **AUTEUR : Isabelle CADORET-DAVID**

2.  ALLOCATE 72
3.  OPEN DATA cartecreditrats.xls
4.  DATA (FORMAT=xls,ORG=obs) /
5.  PRINT /

6.  DISPLAY '***estimation avec la méthode des MCO***'
7.  LINREG Depenses / resid
8.  # CONSTANT age prop revenu revenu2
9.  *estimation MCO*

10. SET resid2 = resid**2
11. COMPUTE scr = %RSS
12. DISPLAY scr
13. *calcul du carré des résidus et la somme des carrés des résidus*

14. SCAT (STYLE=SYMBOL,HMAX=6000) 1
15. # resid2 revenu
16. *graphique du carré des résidus en fonction du revenu*

17. DISPLAY '***Test de White***'
18. SET age2 = age**2
19. SET revenu4 = revenu**2
20. SET ageprop = age*prop
21. SET agerev = age*revenu
22. SET agerev2 = age*revenu2
23. SET proprev = prop*revenu
24. SET proprev2 = prop*revenu2
25. SET revrev2 = revenu * revenu2

26. LINREG resid2
27. # CONSTANT age prop revenu revenu2 age2 revenu4 ageprop agerev agerev2 $
28. proprev proprev2 revrev2

29. compute TW = %nobs*%rsquared
30. DISPLAY 'statistique de White' TW

31. CDF CHISQ TW (%NREG-1)
32. *niveau de significativité du test*

33. DISPLAY '***Test de White modifié ***'
34. LINREG resid2 / RESIDS
35. # CONSTANT revenu revenu2 revrev2
36. COMPUTE lm = %nobs*%rsquared
37. DISPLAY 'statistique du multiplicateur de lagrange' lm
38. CDF CHISQ lm (%NREG-1)

39. DISPLAY '***Test de Goldfeld - Quandt***'
40. ORDER revenu / depenses age prop revenu2
41. print / revenu depenses
42. *range par ordre de croissant du revenu les séries*

```

```

43. LINREG Depenses 1 32
44. # CONSTANT age prop revenu revenu2
45. COMPUTE res1 = %RSS
46. COMPUTE ndf1 = %NDF
47. *estime le modèle sur les 32 premières observations classées*

48. LINREG Depenses 41 72
49. # CONSTANT age prop revenu revenu2
50. COMPUTE res2 = %RSS
51. COMPUTE ndf2 = %NDF
52. *estime le modèle sur les 32 dernières observations classées*

53. COMPUTE Fstat = (res2/ndf2)/(res1/ndf1)
54. DISPLAY 'Fstat' Fstat
55. CDF FTEST Fstat ndf1 ndf2

56. DISPLAY '***estimation avec la méthode des MCG***'
57. DISPLAY '***premier cas : la variance des aléas est fonction de la variable
revenu***'
58. LINREG (SPREAD=revenu) Depenses
59. # CONSTANT age prop revenu revenu2

60. DISPLAY '***second cas : la variance des aléas est fonction de la variable
revenu2***'
61. LINREG (SPREAD=revenu2) Depenses
62. # CONSTANT age prop revenu revenu2

63. DISPLAY '***estimation avec la méthode des mco, mais avec calcul de la matrice
de variance covariance***'
64. LINREG (ROBUSTERRORS) Depenses
65. # CONSTANT age prop revenu revenu2

```

ANNEXE 3 : Résultats du programme RATS

estimation avec la méthode des MCO

Linear Regression - Estimation by Least Squares

Dependent Variable	Usable Observations	Degrees of Freedom
DEPENDENSES	72	67
Centered R**2	0.243578	R Bar **2 0.198418
Uncentered R**2	0.552669	T x R**2 39.792
Mean of Dependent Variable	262.53208333	
Std Error of Dependent Variable	318.04683131	
Standard Error of Estimate	284.75079529	
Sum of Squared Residuals	5432562.0332	
Regression F(4,67)	5.3937	
Significance Level of F	0.00079522	
Log Likelihood	-506.48876	
Durbin-Watson Statistic	1.640032	

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-237.1465136	199.3516649	-1.18959	0.23840661
2. AGE	-3.0818140	5.5147165	-0.55883	0.57813774
3. PROP	27.9409084	82.9223236	0.33695	0.73720572
4. REVENU	234.3470270	80.3659504	2.91600	0.00481866
5. REVENU2	-14.9968442	7.4693370	-2.00779	0.04870093

5432562.03315

Test de White

Linear Regression - Estimation by Least Squares

Dependent Variable	Usable Observations	Degrees of Freedom
RESID2	72	59
Centered R**2	0.199013	R Bar **2 0.036101
Uncentered R**2	0.254059	T x R**2 18.292
Mean of Dependent Variable	75452.25046	
Std Error of Dependent Variable	279705.52995	
Standard Error of Estimate	274610.34322	
Sum of Squared Residuals	4.44924e+12	
Regression F(12,59)	1.2216	
Significance Level of F	0.29051266	
Log Likelihood	-996.65875	
Durbin-Watson Statistic	1.820714	

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	1637390.379	1290979.665	1.26833	0.20966142
2. AGE	5366.243	48893.837	0.10975	0.91297777
3. PROP	812036.761	991630.194	0.81889	0.41614423
4. REVENU	-2021697.610	1053559.103	-1.91892	0.05983568
5. REVENU2	669055.329	365666.731	1.82969	0.07235063
6. AGE2	-424.086	627.459	-0.67588	0.50175906
7. REVENU4	3762.737	2277.358	1.65224	0.10379877
8. AGEPROP	4661.704	14424.595	0.32318	0.74770349
9. AGEREV	11499.862	15614.253	0.73650	0.46434630
10. AGEREV2	-1093.319	1568.053	-0.69725	0.48838710
11. PROPREV	-510192.297	469792.619	-1.08599	0.28189858
12. PROPREV2	51835.098	61799.799	0.83876	0.40499043
13. REVREV2	-86805.325	51162.582	-1.69666	0.09503205

statistique de White 14.32895
 Chi-Squared(12)= 14.328953 with Significance Level 0.28019704

Test de White modifie

Linear Regression - Estimation by Least Squares

Dependent Variable RESID2
 Usable Observations 72 Degrees of Freedom 68
 Centered R**2 0.090918 R Bar **2 0.050811
 Uncentered R**2 0.153392 T x R**2 11.044
 Mean of Dependent Variable 75452.25046
 Std Error of Dependent Variable 279705.52995
 Standard Error of Estimate 272506.80240
 Sum of Squared Residuals 5.04968e+12
 Regression F(3,68) 2.2669
 Significance Level of F 0.08853612
 Log Likelihood -1001.21603
 Durbin-Watson Statistic 1.773970

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-100970.1719	366046.4255	-0.27584	0.78350767
2. REVENU	16525.0703	262359.4976	0.06299	0.94996215
3. REVENU2	19114.3695	54905.5662	0.34813	0.72881670
4. REVREV2	-2035.2632	3331.6358	-0.61089	0.54330900

statistique du multiplicateur de Lagrange 6.54608
 Chi-Squared(3)= 6.546082 with Significance Level 0.08786277

Test de Goldfeld - Quandt

ENTRY	REVENU	DEPENSES
1	1.50	73.18
2	1.50	32.78
3	1.51	63.92
4	1.65	108.61
5	1.70	105.04
6	1.85	165.85
7	1.91	42.62
8	2.00	218.52
9	2.00	319.49
10	2.00	93.20
11	2.00	186.35
12	2.03	42.69
13	2.20	93.11
14	2.28	27.78
15	2.35	43.34
16	2.35	83.08
17	2.37	150.79
18	2.40	41.19
19	2.40	49.56
20	2.40	68.38
21	2.42	9.85
22	2.45	78.87
23	2.46	308.05
24	2.47	552.72
25	2.50	92.00
26	2.54	137.87
27	2.55	101.68
28	2.59	53.65
29	2.60	9.58
30	2.64	65.25
31	2.70	79.51

32	2.72	129.37
33	2.80	34.13
34	2.80	215.07
35	2.88	98.46
36	3.00	222.54
37	3.00	310.94
38	3.00	169.89
39	3.00	95.80
40	3.05	258.55
41	3.07	642.47
42	3.18	810.39
43	3.20	52.58
44	3.20	335.43
45	3.20	417.83
46	3.25	56.15
47	3.50	344.47
48	3.54	541.30
49	3.60	234.05
50	3.75	292.66
51	3.80	777.82
52	3.90	104.54
53	3.95	256.66
54	3.96	40.83
55	4.00	248.72
56	4.00	170.64
57	4.50	15.00
58	4.51	502.20
59	4.52	124.98
60	4.60	405.35
61	4.80	1898.03
62	5.00	451.20
63	5.14	37.58
64	5.50	1532.77
65	5.50	251.52
66	5.70	568.77
67	6.00	306.03
68	6.00	474.15
69	7.00	644.83
70	7.24	235.57
71	9.79	546.50
72	10.00	548.03

Linear Regression - Estimation by Least Squares

Dependent Variable DEPENSES
 Usable Observations 32 Degrees of Freedom 27
 Centered R**2 0.202662 R Bar **2 0.084538
 Uncentered R**2 0.621853 T x R**2 19.899
 Mean of Dependent Variable 113.37125000
 Std Error of Dependent Variable 109.40119200
 Standard Error of Estimate 104.67480110
 Sum of Squared Residuals 295833.97762
 Regression F(4,27) 1.7157
 Significance Level of F 0.17561861
 Log Likelihood -191.51512
 Durbin-Watson Statistic 1.799196

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-218.7084268	688.1537324	-0.31782	0.75306718
2. AGE	-3.7815685	2.8266437	-1.33783	0.19211185
3. PROP	106.2299565	46.2985307	2.29446	0.02976787
4. REVENU	386.2743134	664.7616700	0.58107	0.56601149
5. REVENU2	-87.1465100	157.3231081	-0.55393	0.58418348

Linear Regression - Estimation by Least Squares

Dependent Variable DEPENSES

Usable Observations 32 Degrees of Freedom 27
 Centered R**2 0.043333 R Bar **2 -0.098396
 Uncentered R**2 0.563557 T x R**2 18.034
 Mean of Dependent Variable 433.40781250
 Std Error of Dependent Variable 403.32857468
 Standard Error of Estimate 422.70596107
 Sum of Squared Residuals 4824368.8971
 Regression F(4,27) 0.3057
 Significance Level of F 0.87156333
 Log Likelihood -236.18131
 Durbin-Watson Statistic 2.301373

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-150.3870646	790.7011101	-0.19019	0.85057943
2. AGE	-1.8942153	13.4781417	-0.14054	0.88927676
3. PROP	-72.2118716	190.2806215	-0.37950	0.70728368
4. REVENU	218.3497030	266.9536326	0.81793	0.42055650
5. REVENU2	-13.7474016	20.5661237	-0.66845	0.50952198

Fstat 16.30769

F(27,27)= 16.30769 with Significance Level 0.00000000

estimation avec la méthode des MCG

premier cas : la variance des aléas est fonction de la variable revenu*

Linear Regression - Estimation by Weighted Least Squares

Dependent Variable DEPENSES

Usable Observations 72 Degrees of Freedom 67
 Centered R**2 0.141615 R Bar **2 0.090368
 Uncentered R**2 0.541800 T x R**2 39.010
 Mean of Dependent Variable 134.69147422
 Std Error of Dependent Variable 145.13568504
 Standard Error of Estimate 138.42258449
 Sum of Squared Residuals 1283774.3972
 Log Likelihood -495.58650
 Durbin-Watson Statistic 2.209249

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-181.8706365	165.5190768	-1.09879	0.27579268
2. AGE	-2.9350103	4.6033307	-0.63758	0.52591743
3. PROP	50.4936420	69.8791405	0.72259	0.47244945
4. REVENU	202.1694048	76.7815214	2.63305	0.01049686
5. REVENU2	-12.1136376	8.2731405	-1.46421	0.14781228

second cas : la variance des aléas est fonction de la variable revenu2*

Linear Regression - Estimation by Weighted Least Squares

Dependent Variable DEPENSES

Usable Observations 72 Degrees of Freedom 67
 Centered R**2 0.059369 R Bar **2 0.003212
 Uncentered R**2 0.545293 T x R**2 39.261
 Mean of Dependent Variable 72.075544968
 Std Error of Dependent Variable 70.211227514
 Standard Error of Estimate 70.098384667
 Sum of Squared Residuals 329223.49671
 Regression F(4,67) 1.0572

Significance Level of F 0.38463889
 Log Likelihood -487.62844
 Durbin-Watson Statistic 2.168244

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-114.1088694	139.6874965	-0.81689	0.41688836
2. AGE	-2.6941852	3.8073063	-0.70764	0.48162598
3. PROP	60.4487737	58.5508875	1.03241	0.30559044
4. REVENU	158.4269775	76.3911543	2.07389	0.04193608
5. REVENU2	-7.2492899	9.7243373	-0.74548	0.45858952

estimation avec la méthode des mco, mais avec calcul de la matrice de variance covariance*

Linear Regression - Estimation by Least Squares

Robust Standard Error Calculations

Dependent Variable DEPENSES

Usable Observations 72 Degrees of Freedom 67
 Centered R**2 0.243578 R Bar **2 0.198418
 Uncentered R**2 0.552669 T x R**2 39.792
 Mean of Dependent Variable 262.53208333
 Std Error of Dependent Variable 318.04683131
 Standard Error of Estimate 284.75079529
 Sum of Squared Residuals 5432562.0332
 Log Likelihood -506.48876
 Durbin-Watson Statistic 2.212934

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	-237.1465136	212.9905298	-1.11341	0.26553092
2. AGE	-3.0818140	3.3016612	-0.93341	0.35060669
3. PROP	27.9409084	92.1877767	0.30309	0.76182363
4. REVENU	234.3470270	88.8663517	2.63707	0.00836249
5. REVENU2	-14.9968442	6.9445635	-2.15951	0.03081074