

Méthodes micro-économétriques - Contrôle Continu

Charbel AHOUANDOKOUN, Nazifou AFOLABI, Youssef DIR

15 février, 2023

Exercice 1

1- Importance de l'analyse:

En microéconomie, l'analyse des déterminants des dépenses en soins dentaires peut aider à comprendre les comportements de consommation et les décisions des ménages en matière de soins dentaires. Cela peut avoir des intérêts divers et variés tels que:

- Déterminer les segments de marché et les profils des consommateurs pour les soins dentaires, ce qui peut être utile pour les professionnels de la santé pour cibler leur marketing et leur offre de services.
- Évaluer les effets des politiques publiques sur les dépenses en soins dentaires, telles que les réglementations, les incitations fiscales et les subventions pour les soins de santé.
- Étudier les inégalités de santé et les disparités en matière d'accès aux soins dentaires, y compris les différences en matière d'âge, de revenu et de niveau d'éducation.
- L'étude des régimes d'assurance santé peut montrer comment ils influencent les décisions de consommation des ménages en matière de soins dentaires en termes de coût et de couverture.

En conclusion, l'analyse des déterminants des dépenses en soins dentaires en microéconomie peut fournir une compréhension plus approfondie des comportements des consommateurs et des décisions des ménages en matière de soins dentaires. Cela peut être utile pour les professionnels de la santé pour cibler leur marketing et leur offre de services, ainsi que pour les décideurs politiques pour élaborer des politiques publiques efficaces en matière de soins de santé qui encouragent un accès abordable et une qualité élevée des soins dentaires pour les consommateurs.

2- Tableaux de statistiques descriptives.

Commentaire:

- SR(Sans réponse): A refusé de répondre
- NV: Non vérifié.

Table 1: Statistiques descriptives variables quantitatives

	Min	1er Quartile	Médiane	Moyenne	3è Quartile	Max
age	19	33	48	57.51161	62.0	996
educyr	0	12	14	22.04850	16.0	99
inctot	-309948	10380	24960	36072.08944	48851.5	731653
dvexptot	0	0	0	295.61671	188.0	81000

Table 2: Fréquence des variables qualitatives par année

Années	Sex		hinotcov		workev					himcare		Total
	Homme	Femme	Non	Oui	SR	NV	Ne sait pas	Non	Oui	Non	Oui	
2015	11789	13493	2630	22435	103	38	76	21521	3761	19998	5284	25282
2016	11599	13373	2745	21991	128	44	64	21562	3410	19459	5513	24972
2017	10803	12492	2538	20543	150	19	45	20706	2589	17518	5777	23295
2018	10549	12060	2299	20083	149	1	77	20517	2092	16305	6304	22609
2019	10105	11499	2158	19235	140	1	70	19704	1900	15294	6310	21604

La colonne total du deuxième tableau montre que le nombre d'individus a diminué au fil des années. Il y a sûrement là un problème d'attrition. **Notre panel n'est donc pas cylindré.**

3- Modèle poolé:

Les coefficients interprétés se retrouvent dans la *Table6* de l'annexe. Tous les coefficients sont significatifs contrairement aux MCO simple.

- 1,9% de la variabilité des dépenses en soins dentaires est expliquée par le modèle.
- Grandir d'un an diminue les dépenses en soins dentaires d'un individu de 0.184.
- Etre de sexe féminin augmente en moyenne les dépenses annuelles dentaires de 76,39 par rapport aux hommes.
- Une année d'étude supplémentaire diminue les dépenses de 0,43 pour un individu.
- Le revenu total personnel a un effet marginal annuel positif même si très faible(0,003) sur les dépenses annuelles en soin dentaire.
- Les individus qui ont déjà travaillé ont en moyenne une dépense annuelle en soin dentaire plus élevé que les autres enquêtés; ceux qui ne savent pas ont les moins élevées.
- Avoir la couverture medicare augmente les dépenses annuelles en soins dentaire de 178,188 en moyenne par rapport aux individus sans cette couverture.
- Les individus n'ayant aucune assurance santé ont une dépense annuelle en soins dentaire moins élevé par rapport aux individus ayant au moins une assurance.

4- Modèle à erreurs composées:

Sur la base de la *Table6*, on constate que les coefficients et les écarts-types ont été ajustés mais les premiers cités conservent les mêmes signes que ceux estimés dans le modèle précédent: les relations n'ont pas changés même si leurs forces ont plus ou moins diminuées. Cela est dû au fait que le modèle poolé contrairement au modèle random ne pondère pas les observations par l'information portée par ces dernières. Il avait donc tendance à sur-estimer(sous-estimer) les effets des régresseurs sur les dépenses en soins dentaires, ce qui avait conduit à des coefficients et écarts-types pas forcément optimaux.

5- Comparaison:

Les modèles MCO et Poolé ont les mêmes coefficients mais des écarts-types différents. Cela est du fait de la structure non sphériques des erreurs du modèle, probablement due à une variable manquante ou aux valeurs plutôt aberrantes de la variable educyr, qui font que les écarts-types estimés par les MCO ne sont pas corrects.

En plus des différences relevées à la question précédente, il faut ajouter que des 3 modèles, le modèle à erreurs composées à la F-statistique la plus élevée.

6- Modèle à effets fixes individuels:

Le coefficient associé à la variable *educyr* n'a pas été estimé car ce dernier est fixe dans le temps. Seules les variables *sex*, *inctot*, *workev8* et *himcare2* ont un effet sur les dépenses en soins dentaires. Aussi, certains liens entre les régresseurs et les dépenses en soins dentaires ont changé de sens:

- Etre de sexe féminin diminue les dépenses en soins dentaires de 39.013 par rapports à la dépense moyenne.

7- Choix du modèle entre fixed et random:

Choisir le meilleur modèle revient à vérifier la présence d'une hétérogénéité individuelle corrélée aux variables explicatives de notre modèle. Nous utiliserons ici le test Hausmann. Avec les hypothèses suivantes:

- H_0 : les deux modèles sont convergents, l'estimateur du random effect est plus efficace.
- H_1 : le modèle à effets fixe est meilleur que celui à effets aléatoires.

La p-value du test ($7.4272037 \times 10^{-15}$) étant inférieur à 0.05 on rejette H_0 , le modèle à effets fixes est à privilégier.

8- Choix du meilleur de nos modèles:

De toutes les comparaisons précédentes et de la question 7, on choisira le modèle à effets fixes.

9- Explication des facteurs des dépenses en soins dentaires:

- Etre de sexe féminin diminue les dépenses en soins dentaires de 39.013 par rapport à la dépense moyenne en soins de l'individu.
- Une augmentation d'une unité du revenu personnel d'un individu par rapport à la moyenne de son revenu, entraîne une augmentation de 0.001 de ses dépenses en soins dentaires par rapport à sa moyenne de dépenses en soins dentaires sur la durée de l'enquête.
- Avoir la couverture medicare, augmente les dépenses en soins dentaires de 229.546 par rapport à la dépense moyenne en soins de l'individu.

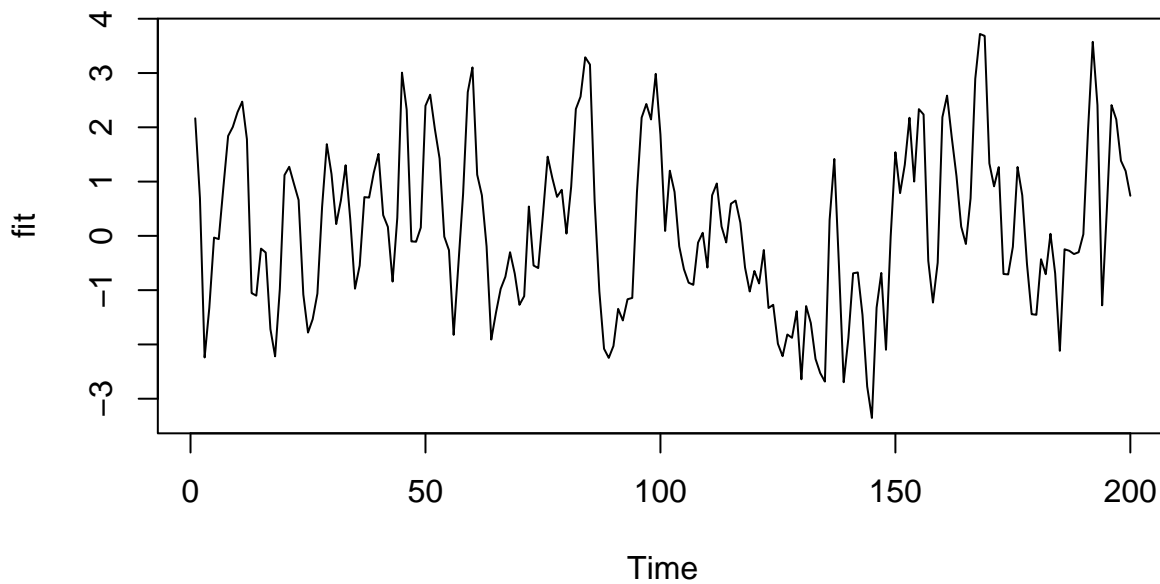
10- Valeurs aberrantes:

Oui, pour certains individus le nombre d'années d'études est supérieur à l'âge ou encore certains présentent un revenu total personnel négatif ce qui n'est pas possible car il s'agit du revenu et non du patrimoine ou la trésorerie. Les régressions faites par en excluant valeurs sont présentes en annexe. On constate une augmentation en générale des écarts-types(diminution néanmoins sur certains) et une modifications de la force du lien entre les dépenses en soins dentaires et les variables de notre modèle. En particulier pour le modèle à effets fixes, les coefficients des variables *Sex* et *workev8* ne sont pas significatifs contrairement au modèle avec la base de départ tandis que celui de *workev9* a subi l'effet inverse.

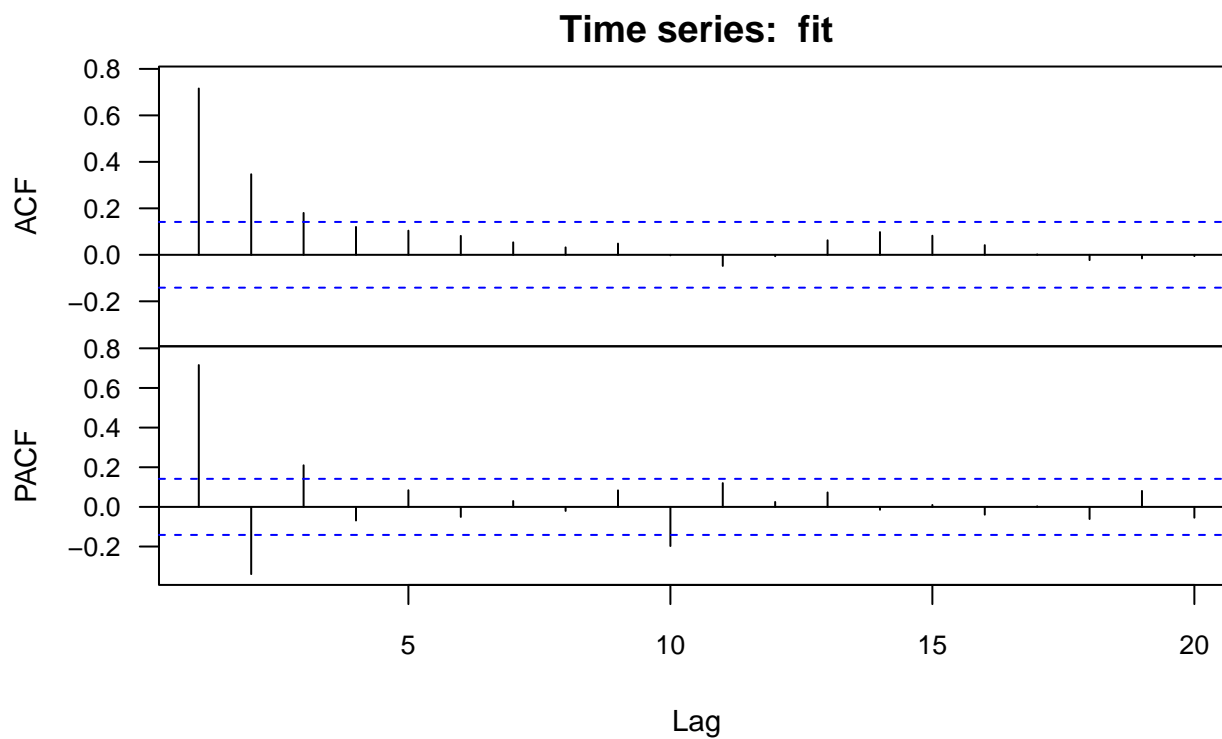
Exercice 2

1- Simulation:

L'équation est celle d'un processus ARMA(1,1) ou ARIMA(1,0,1).



2- Fonction d'autocorrélation et fonction d'autocorrélation partielle:



Commentaire:

- *ACF*: on remarque une oscillation même si elle n'est pas régulière, cela indique la présence d'une composante **moyenne mobile**; Le dernier pic significatif est au lag 3.
- *PACF*: il y a un pic significatif au niveau du lag 10, il y a donc une autocorrelation partielle au lag 10.
- La *PACF* et l'*ACF* nous font penser à une $MA(3)$ et une $AR(10)$, ce qui n'est pas raccord avec l'équation. **Il faudra donc procéder à des ajustements**

3- Ajustements:

a- Un ARMA(1,1)

Table 3: ARMA(1,1)

<i>Dependent variable:</i>	
	fit
ar1	0.469*** (0.080)
ma1	0.608*** (0.083)
intercept	0.178 (0.203)
Observations	200
Log Likelihood	-275.272
σ^2	0.913
Akaike Inf. Crit.	558.544

Note: * p<0.1; ** p<0.05; *** p<0.01

b- Un AR(1)

Table 4: AR(1)

<i>Dependent variable:</i>	
	fit
ar1	0.719*** (0.049)
intercept	0.197 (0.261)
Observations	200
Log Likelihood	-293.698
σ^2	1.100
Akaike Inf. Crit.	593.396

Note: * p<0.1; ** p<0.05; *** p<0.01

c- Un MA(1)

Table 5: MA(1)

<i>Dependent variable:</i>	
	fit
ma1	0.882*** (0.039)
intercept	0.166 (0.135)
Observations	200
Log Likelihood	-287.852
σ^2	1.034
Akaike Inf. Crit.	581.704

Note: * p<0.1; ** p<0.05; *** p<0.01

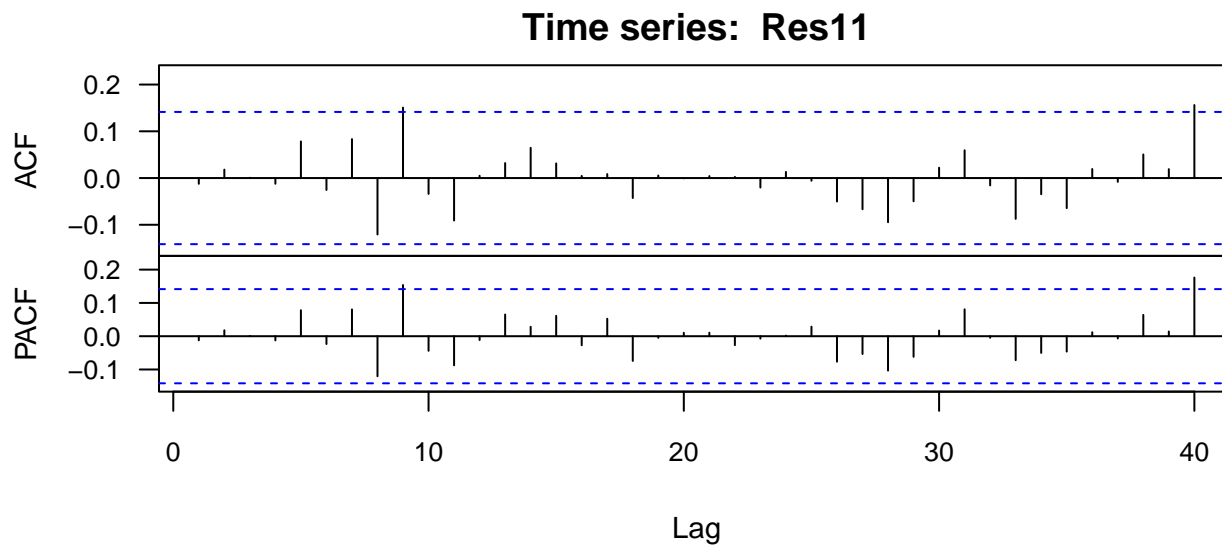
4- Nature des résidus pour chaque modèle:

Ici, on fait un box test (test de blancheur) sur les résidus (BB).

- H_0 : les résidus sont des Bruit Blanc
- H_1 : les résidus ne sont pas des Bruit Blanc

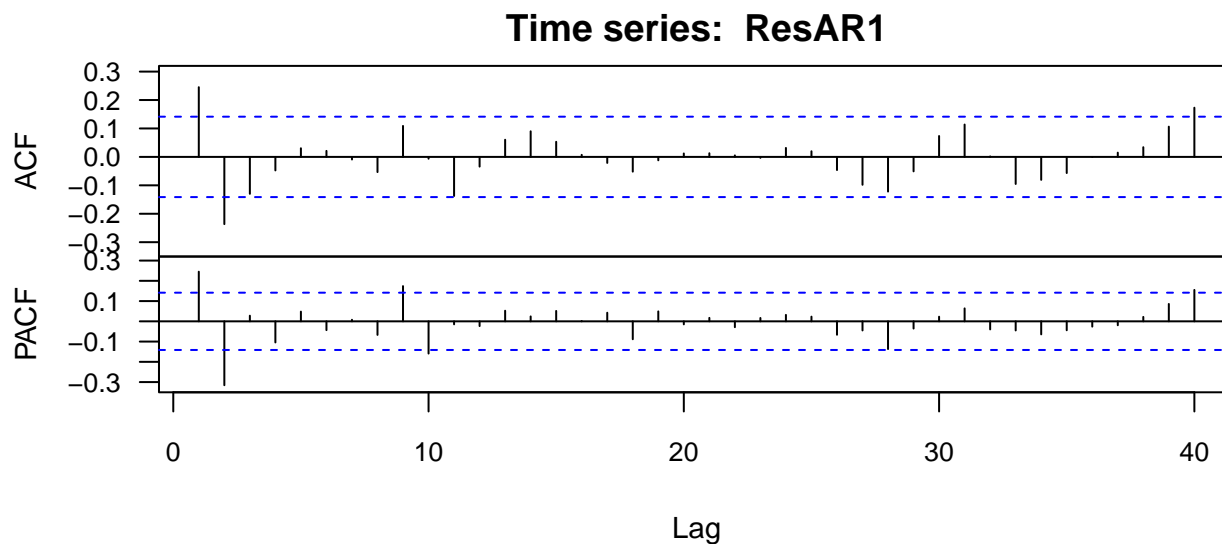
Règle de décision : si p-value < 0.05 On rejette H_0

a- residus ARMA (1,1):



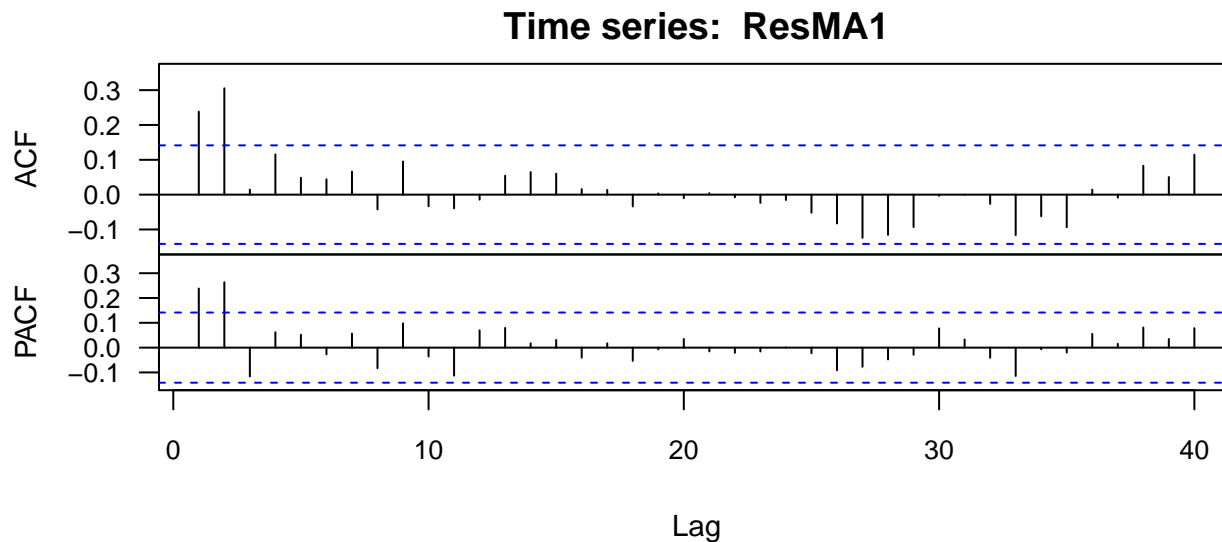
On a un p-value de 0.862, au test de blancheur, supérieur au seuil de 0.05; on ne rejette pas H_0 donc les résidus sont des Bruits Blancs.

b- residus AR(1):



La p-value du test de blancheur (0.001) est inférieur à 0.05, on rejette H_0 , les résidus ne sont donc pas des Bruits Blancs.

c- residus MA(1):



La p-value du test de blancheur (0.001) est inférieur à 0.05, on rejette H_0 , les résidus ne sont donc pas des Bruits Blancs.

Conclusion:

Seul le ARMA(1,1) a des residus qui sont des Bruits Blancs.

5- Autre critère de décision :

L'autre critère pour le choix de modèle est l'AIC. L'AIC (Akaike information criterion) est un critère de sélection de modèle qui est utilisé pour évaluer la qualité d'un modèle statistique par rapport à d'autres modèles. Plus précisément, l'AIC est une mesure du trade-off entre la précision du modèle (capturée par la qualité de la déviation des données observées par rapport aux données prévues) et la complexité du modèle (capturée par le nombre de paramètres du modèle). Elle mesure donc la perte de précision du modèle du fait de sa complexité.

Principe:

On choisit le modèle ayant l'AIC le plus faible.

Choix:

	Récapitulatif des AIC de nos 3 modèles		
	ARMA	AR	MA
AIC	558.5438	593.3965	581.704

Dans notre cas le choix porte sur le modèle ARMA(1,1) car il a l'AIC le plus faible.

Cohérence du choix:

Oui, il y a une cohérence entre le modèle ARMA(1,1) choisi et le modèle de la question 1 car l'équation de la question 1 est celle d'un ARMA(1,1).

Annexe:

Table 6: Questions 3-4-5

	<i>Dependent variable:</i>		
	dvexptot		
	<i>OLS</i>	<i>panel</i>	
	MCO	Poolé	Random
	(1)	(2)	(3)
age	−0.184*** (0.034)	−0.184*** (0.019)	−0.183*** (0.019)
sex2	76.399*** (6.411)	76.399*** (6.776)	75.830*** (6.741)
educyr	−0.426*** (0.121)	−0.426*** (0.093)	−0.437*** (0.093)
inctot	0.003*** (0.0001)	0.003*** (0.0001)	0.003*** (0.0001)
workev2	46.411*** (10.668)	46.411*** (10.834)	47.094*** (10.794)
workev7	−65.671 (42.781)	−65.671* (36.801)	−64.911* (36.006)
workev8	−63.719 (106.891)	−63.719*** (20.163)	−74.393*** (21.350)
workev9	−130.656** (59.944)	−130.656*** (19.000)	−128.068*** (19.346)
himcare2	178.188*** (7.707)	178.188*** (10.225)	177.745*** (10.171)
hinotcov2	−139.072*** (10.205)	−139.072*** (6.301)	−138.167*** (6.356)
Constant	108.118*** (11.765)	108.118*** (11.440)	109.846*** (11.424)
Observations	117,762	117,762	117,762
R ²	0.019	0.019	0.018
Adjusted R ²	0.019	0.019	0.018
Residual Std. Error	1,077.114 (df = 117751)		
F Statistic (df = 10; 117751)	228.487***	228.487***	2,123.617***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 7: Random effects et fixed effects (Question 6)

	<i>Dependent variable:</i>	
	dvexptot	
	Random effect	fixed effect
	(1)	(2)
age	−0.183*** (0.019)	−0.054 (0.055)
sex2	75.830*** (6.741)	−39.013** (16.802)
educyr	−0.437*** (0.093)	
inctot	0.003*** (0.0001)	0.001*** (0.0003)
workev2	47.094*** (10.794)	−27.099 (29.853)
workev7	−64.911* (36.006)	8.546 (35.270)
workev8	−74.393*** (21.350)	−659.726* (351.424)
workev9	−128.068*** (19.346)	−40.703 (80.396)
himcare2	177.745*** (10.171)	229.546*** (77.536)
hinotcov2	−138.167*** (6.356)	−34.192 (26.790)
Constant	109.846*** (11.424)	
Observations	117,762	117,762
R ²	0.018	0.001
Adjusted R ²	0.018	−1.630
F Statistic	2,123.617***	3.906*** (df = 9; 44736)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: Les 4 modèles avec la base nettoyée

	<i>Dependent variable:</i> dvexptot			
	<i>OLS</i>		<i>panel linear</i>	
	MCO	Poolé	Random effect	Fixed effect
	(1)	(2)	(3)	(4)
age	-0.336*** (0.042)	-0.336*** (0.029)	-0.330*** (0.028)	-0.043 (0.072)
sex2	73.328*** (7.497)	73.328*** (7.968)	73.085*** (7.930)	-21.073 (15.945)
educyr	12.372*** (0.989)	12.372*** (0.923)	11.934*** (0.905)	
inctot	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)	0.001*** (0.0003)
workev2	42.683*** (16.089)	42.683** (18.539)	44.930** (18.212)	-40.034 (50.356)
workev7	-118.533** (57.453)	-118.533*** (40.117)	-113.812*** (40.700)	-1.247 (57.170)
workev8	60.969 (341.348)	60.969 (74.884)	60.118 (76.050)	-100.509 (221.228)
workev9	-183.245** (89.639)	-183.245*** (29.209)	-179.336*** (29.269)	-185.231* (97.553)
himcare2	210.517*** (9.073)	210.517*** (12.090)	209.614*** (12.025)	265.472*** (90.175)
hinotcov2	-128.382*** (12.877)	-128.382*** (8.086)	-128.115*** (8.166)	-27.727 (32.093)
Constant	-45.221** (19.991)	-45.221** (21.292)	-40.495* (20.932)	
Observations	94,503	94,503	94,503	94,503
R ²	0.020	0.020	0.018	0.001
Adjusted R ²	0.020	0.020	0.018	-1.463
Residual Std. Error	1,130.919 (df = 94492)			
F Statistic	189.918*** (df = 10; 94492)	189.918*** (df = 10; 94492)	1,753.594***	3.580*** (df = 9; 38339)

Note:

* p<0.1; ** p<0.05; *** p<0.01