

# **Impact des facteurs socio-économiques sur l'émission de CO2**

ABUBAKER Mohamed  
M1 ARB

FRAYSSE Christopher  
M1 ARB

## Résumé

L'étude a permis de mettre en évidence une importante relation entre l'émission de CO2 par habitant et le revenu par habitant ainsi que l'intensité énergétique d'un pays dans le cas d'un modèle en coupe transversale. L'ajout de variables indicatrices permettant de quantifier les effets régionaux a été pertinent, cependant, l'analyse aurait pu être plus poussée en se penchant sur des sous-groupes plus précis mais le manque d'observation empêchait une trop grande quantité de variables explicatives. Dans un dernier temps l'intégration de la courbe environnementale de Kuznets au modèle est fonctionnel, cependant il est nécessaire de prendre avec des pincettes se résultat, la théorie derrière cette courbe étant encore largement débattu dans la littérature.

Pour la totalité des modèles, il est nécessaire de s'interroger sur la pertinence des résultats et leur cause structurelle, pour se faire une analyse sur des séries temporelles permettra de mettre en évidence les effets structurelles des politiques environnementales et les effets structurelles économiques mondiaux. Ainsi, cela permettrait de différencier à la fois les effets à court termes des tendances à long termes mais aussi des impacts réels d'une croissance économique, d'une variation de revenu par habitant et des politiques environnementales sur la quantité des CO2 émises par habitants.

## 1 Introduction

Les débats lors de la dernière COP manifeste un paradoxe socio-économique profond : les nations en phase de transition économique sont confrontées à une bifurcation critique entre la nécessité de progression de leur PIB et les exigences internationales visant la diminution du volume de dioxyde de carbone émis. C'est bloquer dans cette occlusion qu'elles doivent regarder nostalgiquement les marchés développés qui, lors de la ruée vers le nouveau monde de l'expansion économique, ont cinglés les rivages de l'expansion sans être freinées par les turbulences des contraintes environnementales. Ces premiers explorateurs qui, après avoir bénéficié de ce nouveau monde libre et lucratif, exigent des marchés émergents de naviguer vers une route inexplorée en adoptant un modèle de croissance soucieux de l'environnement. Ce nouveau cap représente pour ces économies un voyage mouvementé sur des mers inconnus qu'ils se doivent de cartographier avec des réserves financières modestes tout en étant pris en tenaille par des vents contraires : ceux de la conscience écologique et ceux d'une demande interne en hausse constante. Cette dualité entretient le débat ardu sur l'égalité environnementale et la justice climatique, interrogeant la répartition équitable entre expansion économique, obligation écologique tout en veillant à ne pas oublier les responsabilités historiques.

Ce travail se propose capitaine d'une navigation analytique à travers les brumes qui associent les dynamiques socio-économiques aux émissions de CO2, orientant son cap vers une étude transversale de l'année 2019 et s'appuyant sur les données de 173 pays. Cette étude se concentrera sur l'examen des liens entre le revenu par habitant, les spécificités géographiques, l'intensité énergétique relative au PIB, ainsi que d'autres variables. Son but sera de découvrir et cartographier les relations entre ces variables et leur influence sur les émissions de CO2, en identifiant des tendances distinctes, tout en s'abstenant de prétendre offrir une solution aux dilemmes socio-économiques actuels.

L'année de départ sera l'année 2019 afin de minimiser les perturbations temporelles liés à l'utilisation d'une coupe transversales qui auraient pu trouver leur cause dans l'impact de la pandémie du COVID-19 dans les années suivantes. Il convient de souligner que ce travail vise à éclairer les dynamiques de 2019 et n'entend pas extrapoler ses conclusions à l'année 2023.

## 1.1 Définition des données

Les données exploitées proviennent d'un ensemble de base de donnée reconnues telles que la Banque Mondiale [9], de l'Energy Information Administration [5] ou encore de spécialiste [1]. Après la récolte, un processus de nettoyage et de normalisation a été appliqué à ces données ayant pour vocation de limiter les erreurs intrinsèques aux données originales.

Notons malgré tout des contraintes dans la couverture géographique. Des obstacles telles que les disparités entre les bases de données ou de situation géopolitique ont été une entrave à la collecte de données cohérente pour l'ensemble des pays. Ainsi, certains pays notables tel que Taiwan<sup>1</sup> ou Israël<sup>2</sup> mais aussi tous les petits états sont malheureusement absents de l'analyse. De ce fait, l'étude portera sur un total de 173 observations.

### 1.1.1 Variable expliquée

La variable dépendante sera :

- *CO2globalC* : Il s'agit d'une estimation la quantité moyenne de CO2 émise par individu dans un pays, exprimée en tonnes. Cette mesure reflète les émissions générées directement à l'intérieur des frontières nationales et ne tient pas compte des émissions attribuables au commerce international de biens, ni celles issues des transports maritimes et aériens internationaux.

Le choix de l'émissions de CO2 par personne a été guidé par des soucis d'équités et de comparabilité, permettant d'évaluer de manière plus juste les pays aux populations variables. Cela permet la comparaison entre deux pays différents tout en mettant en relief le poids individuelles de chaque individus dans les émissions totales de CO2.

### 1.1.2 Variable explicative

Les variables de contrôle seront :

- *Ener\_Intens* : La quantité d'énergie consommée par habitant, mesurée en quadrillions d'unités thermiques britanniques (BTU). La consommation d'énergie par tête peut révéler l'efficacité avec laquelle un pays utilise l'énergie pour répondre aux besoins de sa population.
- *Ener\_GDP* : Le ratio entre la consommation totale d'énergie et le produit intérieur brut (PIB), ajusté par la parité de pouvoir d'achat (PPA), ce qui permet de comparer le niveau d'intensité énergétique entre pays en neutralisant les effets des taux de change. Une intensité énergétique élevée par unité de PIB peut indiquer une efficacité énergétique plus faible et une économie plus dépendante de l'énergie.
- *Energy\_Prod* : La production totale d'énergie d'un pays, englobant les combustibles fossiles (pétrole brut, gaz naturel) et les énergies renouvelables. Cette variable peut refléter la capacité d'un pays à générer de l'énergie.
- *Density* : La densité de population d'un pays, obtenue en divisant le nombre total d'habitants par la superficie du pays en km<sup>2</sup>. Une densité de population élevée peut indiquer une pression plus grande sur les ressources et l'environnement et peut affecter les émissions de CO2.
- *Income* : Le PIB par habitant ajusté en PPA, ce qui permet une comparaison équitable du niveau de vie entre les pays. Un revenu plus élevé par habitant peut être associé à une consommation d'énergie plus importante, donc potentiellement à des émissions de CO2 plus élevées.
- *Life\_Exp* : L'espérance de vie moyenne des habitants d'un pays. Une espérance de vie plus longue peut indiquer un meilleur niveau de santé et de bien-être, qui peut être corrélé à des niveaux de vie plus élevés et donc à une consommation d'énergie plus importante.

---

1. Non reconnu par l'entière des organismes qui ont servi à l'établissement de la base de donnée

2. Données contradictoires entre certaines bases de données

- *Avg\_School* : Le nombre moyen d'années d'éducation reçues par les habitants d'un pays. Un niveau d'éducation plus élevé peut être lié à une plus grande conscience environnementale et à des comportements plus économes en énergie.

### 1.1.3 Variable indicatrice

La variable indicatrice sera :

- *Zone géographique*<sup>3</sup> : La zone géographique où se situe le pays. La méthodologie de la Banque mondiale [6] est utilisée pour examiner les émissions de CO2 par zone géographique plutôt que par niveau de revenu, afin de mettre en évidence les facteurs uniques tels que les politiques environnementales, le développement économique et les conditions climatiques qui sont influencés par ces émissions, par d'autres variables.

## 1.2 Statistiques de variable

### 1.2.1 Statistiques descriptives

Variable	Moyenne	Médiane	E.T.	Min	Max
CO2GlobalC	4.43	2.10	5.49	0.03	36.03
Ener_Intens	87.20	49.70	116.00	1.09	724.00
Ener_GDP	3.87	3.35	2.36	0.80	17.70
Ener_Prod	3.50	0.17	13.50	0.00	124.00
Density	196.00	82.30	641.00	2.08	$7.97 \times 10^3$
Income	1.95	1.30	1.99	766.00	$1.07 \times 10^5$
Life_Exp	72.50	73.60	7.43	52.90	84.40
Avg_School	8.84	9.03	3.18	2.12	14.10

TABLE 1 – Statistiques descriptives, utilisant les observations 1 - 173

L'analyse préliminaire des statistiques descriptives des variables indique plusieurs tendances significatives et points d'attention pour le modèle. Dans un premier temps, une médiane inférieure à la moyenne pour certaines variables la potentielle distribution asymétrique positive suggérant une inclinaison vers la droite. Ceci couplé à un écart-type substantiel et une forte disparité entre les valeurs minimales et maximales interrogent sur la présence de valeurs aberrantes.

L'analyse graphique<sup>4</sup> confirme les observations en montrant des relations approximativement linéaires entre les émissions de CO2 et des variables telles qu'*Income*<sup>5</sup> et *Ener\_Intens*<sup>6</sup>, tout en soulignant la nécessité de traiter des valeurs aberrantes marquées. La dispersion pour certaines variables tels que *Density* indique à la fois une variabilité notable et interroge sur une hétérogénéité significative dans les observations.

Ainsi, les statistiques descriptives et les représentations graphiques suggèrent l'exploration de la nécessité de transformations des variables, peut-être non-linéaires, afin de mieux comprendre les distributions et d'améliorer potentiellement l'ajustement du modèle aux données en réduisant l'impact

3. A.1

4. A.2

5. Income et CO2GlobalC

6. Ener\_Intens et CO2GlobalC

des valeurs extrêmes. Il n'est cependant pas possible de conclure sans avoir vu la distribution des résidus du modèle linéaire.

### 1.2.2 Matrice des coefficients de corrélation

Variables	CO2GlobalC	Ener_Intens	Ener_GDP	Ener_Prod	Density	Income	Life_Exp	Avg_School
CO2GlobalC	1,0000							
Ener_Intens	0,8829	1,0000						
Ener_GDP	0,5885	0,5903	1,0000					
Ener_Prod	0,2929	0,2413	0,2389	1,0000				
Density	0,0657	0,3969	0,0930	-0,0388	1,0000			
Income	0,7241	0,8309	0,2526	0,1689	0,3072	1,0000		
Life_Exp	0,5270	0,5670	0,3223	0,1491	0,1732	0,7164	1,0000	
Avg_School	0,5144	0,5284	0,3580	0,1420	0,0686	0,6779	0,7688	1,0000

TABLE 2 – Matrices des coefficients de corrélation

La matrice de corrélation a pour but de mettre en évidence associations significatives entre les variables. Ainsi, il est possible de voir une forte corrélation positive entre les émissions de CO2 par personne et l'intensité énergétique (0,8829) ainsi que le revenu par habitant (0,7241).

Ce résultat est tout somme assez logique, les économies les plus riches et qui consomment le plus d'énergies sont celles qui tendent à avoir les plus grosses émissions de CO2. La forte corrélation entre l'intensité énergétique et le revenu par habitant (0,8309) peut sous-entendre qu'un pays à haut revenu a une industrie généralement plus énergivore, il sera nécessaire de comprendre les causes structurelles de ce phénomène pour pouvoir poser un lien plus concluant.

La matrice de corrélation montre aussi des coefficients très proche de zéros supposant le peu de relation linéaire entre l'émissions de CO2 par personne et la densité de population (0,0657)

## 2 Modèles et méthodologie économétrique

### 2.1 Modèle 1

En suivant la littérature empirique en économie de l'énergie, il est plausible de formuler la relation long-terme entre les émissions de CO2, le revenu sous une combinaison de fonctions logarithmique confirmant les observations lors de l'analyse des statistiques descriptives. Le premier modèle n'inclura pas les variables indicatrices pour permettre d'observer l'importance des effets régionaux :

$$\ln(\text{CO2globalC}) = \text{const} + \beta_1 \text{Density} + \beta_2 \ln(\text{Income}) + \beta_3 \text{Ener\_Intens} + \beta_4 \ln(\text{Ener\_GDP}) + \beta_5 \text{Life\_Exp} + \beta_6 \text{Avg\_School} + \epsilon \quad (1)$$

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	t de Student	p. critique
const	-9,54	0,389	-24,52	0,0000
l_Ener_GDP	0,996	0,0583	17,08	0,0000
Ener_Intens	-0,000729	0,000376	-1,939	0,0542
Energy_Prod	0,00254	0,00195	1,304	0,1942
Density	-0,000147	4,48e-005	-3,289	0,0012
l_Income	1,06	0,0577	18,33	0,0000
Life_Exp	-0,00784	0,00659	-1,190	0,2356
Avg_School	-0,0173	0,0152	-1,137	0,2572
Somme carrés résidus	17,8047	Somme des carrées expliqués		333,168
Somme Carrés Total	350,972	$F(7, 165)$		441,0772
$R^2$	0,949270	$R^2$ ajusté		0,947118

TABLE 3 – Modèle 1 : MCO, sans indicatrices, 173 observations

Avec un  $R^2 = 0,949270$  le modèle explique 94.92% de la variance de la variable expliquée. Le modèle peut être considéré comme potentiellement efficace pour expliquer la variabilité de la variable dépendante ( $\ln \text{CO2GlobalC}$ ) en montrant une importante corrélation entre les régresseurs et la variable dépendante.

Néanmoins, il ne permet pas de statuer sur la nature causal des relations ni sur sa capacité à généraliser à des données qui diffère de l'échantillon testé. Le  $R^2$  n'informe pas non plus sur la validité de la régression il est alors nécessaire de s'intéresser aux test de diagnostic pour déterminer si la régression est valide.

Le modèle présente une statistique  $Fc = 441,0772 > F(6, 165)$  ce qui permet de déterminer que l'hypothèse nulle  $H_0 : \forall \beta_i = 0$  (Tous les coefficients de régression sont égaux à zéro) est rejeté au risque de première espèce  $\alpha = 0,05^7$ , cela implique qu'il y a eu moins un coefficient de régressions significativement différents de 0 et que le modèle est donc significatif.

#### 2.1.1 Information sur la validité de la régression

Après avoir effectué la régression il est nécessaire de déterminer si les propriétés statistiques sont respectés [8] pour s'assurer de la validité des résultats. Il est alors important de vérifier si l'estimateur par moindres carrés est le "BLUE", cela passe par un traitement de la linéarité des paramètres dans

7. Constant pour tout le reste de l'étude

le modèle, la stricte exogénéité des régresseurs<sup>8</sup>, l'absence de multicollinéarité parfaite, la présence d'une matrice de variance-covariance sphérique pour les erreurs. Dans une moindre mesure il est possible de s'intéresser à la normalité des erreurs bien que non nécessaire à l'établissement du meilleur estimateur linéaire non biaisé.

### Linéarité du modèle

La théorème de Gauss-Markov implique que la première hypothèse à l'estimation d'un estimateur "BLUE" est la linéarité du modèle. Cette hypothèse est validé par la forme de l'équation 1, qui bien que les variables soient transformés pour capturer des effets supposés non linéaire, reste linéaire dans ses paramètres.

### Absence de multicollinéarité

Pour tester la multicollinéarité de le modèle il est nécessaire d'analyser la valeur VIF des variables :

Variable	Valeur
ln(Ener_GDP)	1,727
Ener_Intens	3,060
Energy_Prod	1,102
Density	1,310
ln(Income)	6,869
Life_Exp	3,819
Avg_School	3,732

TABLE 4 – Facteurs d'inflation de variance

Toutes les valeurs supérieures à 10 sont considérées comme étant indicatrice d'une potentielle colinéarité. Ainsi, le modèle ne présente aucune valeur supérieur ou proche de ce seuil, de façon général les valeurs semblent assez éloigné du seuil de 10 permettant de conclure en faveur de l'hypothèse que le modèle ne présente aucune multicollinéarité.

### La matrice de variance-covariance des erreurs est sphérique

Cette hypothèse sous-entend deux sous-hypothèse ; La première est celle d'homoscédasticité des erreurs. La seconde serait celle d'absence d'auto-corrélation des erreurs. Dans cette partie, ne sera traité qu'uniquement l'homoscédasticité en posant l'hypothèse forte que le modèle ne comprend pas d'auto-corrélation des erreurs.<sup>9</sup> Pour tester cette hypothèse, un test de *White* sera effectué. Ce dernier résulte en une statistique de test  $nR^2 = 52,084215 < \chi^2(35)$  rejetant alors l'hypothèse  $H_0$  pour la présence d'homoscédasticité. Le modèle contient de l'hétéroscédasticité.

Puisque le test de *White*[7] a révélé l'existence d'hétéroscédasticité dans les résidus, il n'est pas nécessaire de procéder à un test de *Breusch-Pagan*[2] ; l'hypothèse  $H_0$  est donc rejetée, le modèle présente de l'hétéroscédasticité. Pour remédier à ce problème, il est nécessaire d'examiner de plus près les résultats du test de *White*, qui suggèrent que les relation incluant la variable *Avg\_School* puissent être la cause de cette hétéroscédasticité observée.

8. Supposé vraie pour la totalité de l'étude

9. Aucun test à la disposition des auteurs ne permet de traiter l'auto-corrélation spatiale des erreurs

### Normalité des erreurs

La normalité des erreurs est une hypothèse importante pour les petits échantillons de valeurs. Le modèle présenté possédant un total de 173 observations alors le théorème de la limite centrale pousse à croire que même si les termes d'erreurs ne sont pas parfaitement distribués autour de la moyenne alors les estimateurs des coefficients obtenus par les moindres carrés ordinaires auront néanmoins une distribution qui se rapproche de la distribution normale.

### Significativité des paramètres

Dans cette partie, il sera testé si l'entière des variables sont utiles pour l'estimation de la variable expliquée. Pour se faire, il sera important de se concentrer sur les valeurs critiques (p. critique) présentes dans le tableau 1 qui sont supérieures à  $p = 0,05$ [3].

Ainsi, il est possible de voir que *Energy\_prod*, *Life\_Exp* et *Avg\_School* ne sont pas statistiquement significatifs leur p. critique étant respectivement égal à 0,1942, 0,2356 et 0,2572.

Quant à *Ener\_Intens* légèrement supérieur à 0,05, avec une valeur p de 0,0542, il sera choisis de la maintenir dans l'analyse ultérieure pour déterminer si sa signification statistique s'améliore après l'élimination des trois autres variables.

#### 2.1.2 Conclusion économétrique

Ce modèle, bien que satisfaisant de part son fort pouvoir explicatif les problèmes identifiés d'hétéroscédasticité et de variable non-pertinente nécessitent une ré-évaluation et un ajustement. Le prochain modèle supprimera les variables non significatives (*Energy\_Prod*, *Life\_Exp*, et *Avg\_School*) ce qui devrait aussi régler le problème d'hétéroscédasticité par la suppression de la variable *Avg\_School*. Dans un objectif double, il inclura les variables indicatrices pour capturer les effets régionaux, afin de déterminer l'impact spécifique des conditions régionales<sup>10</sup> sur les émissions de CO2.

### 2.2 Modèle 2

Voici le modèle 1 corrigé auquel seront ajouté les variables indicatrices des zones géographiques pertinentes<sup>11</sup> :

$$\begin{aligned} \ln(\text{CO2globalC}) = & \text{const} + \beta_1 \text{Density} + \beta_2 \ln(\text{Income}) + \beta_3 \text{Ener\_Intens} \\ & + \beta_4 \ln(\text{Ener\_GDP}) + \beta_5 \text{SAS} + \beta_6 \text{ECS} \\ & + \beta_7 \text{AFE} + \beta_8 \text{LCN} + \epsilon \end{aligned} \quad (2)$$

---

10. B.2

11. C.1



Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	t de Student	p. critique
const	-9,88	0,325	-30,40	0,0000
l_Ener_GDP	0,973	0,0549	17,72	0,0000
Ener_Intens	-0,000940	0,000376	-2,500	0,0134
Density	-0,000171	4,26e-005	-4,012	0,0001
l_Income	1,04	0,0375	27,69	0,0000
ECS	-0,257	0,0705	-3,653	0,0003
AFE	-0,160	0,0820	-1,952	0,0527
LCN	-0,228	0,0739	-3,077	0,0024
Somme carrés résidus	16,5951	Somme des carrées expliqués		334,377
Somme Carrés Total	350,972	$F(7, 165)$		299,8074
$R^2$	0,952717	$R^2$ ajusté		0,950711

TABLE 5 – Modèle 2 : MCO, avec indicatrices, 173 observations

Avec un  $R^2 = 0,952717$ , le modèle a gagné en efficacité en incluant les variables indicatrices même après le rejet de certaines variables. De la même façon que lors du modèle 1, il n'est pas possible de statuer sur les relations causales qui lient les variables expliquées et explicatives, il ne permet que de statuer sur la nature corrélée entre régresseurs et variable dépendante. Pour reprendre la même analyse que lors du premier modèle, ce modèle présente une statistique  $F_c = 299,8074 > F(7, 165)$  ainsi, l'hypothèse  $H_0 : \rho^2 = 0$  est rejeté au risque de première espèce  $\alpha = 0,05$  le modèle est significatif (bien ajusté).

### 2.2.1 Information sur la validité de la régression

Il est nécessaire de tester sa validité statistique de ce nouveau modèle de régression. Il est toujours important de vérifier si l'estimateur par moindres carrés est le "BLUE". Cela passera par des test sur l'absence de multicolinéarité parfaite et la présence d'une matrice de variance-covariance sphérique pour les erreurs. La linéarité étant conservée de par la structure linéaire en fonction des paramètres et la normalité la conclusion débattu lors du premier modèle sera conservé ici.

#### Absence de multicolinéarité

L'une des conclusions de l'analyse de la première régression a été de dire que la multicolinéarité était absente du premier modèle. Le test effectué, un test de *VIF* a éclairé sur l'absence de valeur supérieur à 10, Désormais, il est nécessaire d'observer si en rajoutant les variables indicatrices, la multicolinéarité apparaît :

Variable	Valeur
ln(Ener_GDP)	1,641
Ener_Intens	3,278
Density	1,276
ln(Income)	3,105
ECS	1,691
AFE	1,282
LCN	1,383

TABLE 6 – Facteurs d'inflation de variance

Ainsi toutes les valeurs sont encore très éloigné du seuil de 10 permettant de conclure en l'absence de multicolinéarité dans le modèle. Les variables sont indépendantes les une des autres même après l'ajout des variables indicatrices pour essayer de quantifier l'effet spatiale.

### Présence d'homoscédasticité

L'autre conclusion importante du premier modèle était d'incriminer la variable *Avg\_School* comme étant la cause de cette hétéroscédasticité<sup>12</sup>. Pour vérifier l'absence d'hétéroscédasticité dans ce nouveau modèle un test de *White* sera effectué et si ce dernier est concluant, un test de *Breusch-Pagan* sera aussi effectué. La valeur la plus faible sera conservée et permettra de statuer sur la présence ou non d'hétéroscédasticité.

Ainsi, le test de *White* donne une statistique de test :  $TR^2 = 31,526135$  qui est inférieur à  $\chi^2(29)$ . L'hypothèse  $H_0$  est acceptée au risque de première espèce et le modèle présente bien de l'homoscédasticité selon le test de *White*.

Il est nécessaire d'utiliser le test de *Breusch-Pagan* pour vérifier les résultats du test de *White* qui donne à son tour une statistique de test :  $LM = 12,465654$  qui elle aussi est inférieure à  $\chi^2(7)$ . Le modèle présente donc bien de l'homoscédasticité selon le test de *Breusch-Pagan*.

### Significativité des paramètres

Dans ce modèle, la presque totalité des variables possède un t. de student ( $\frac{\hat{\beta}}{\sigma_{\hat{\beta}}}$ ) > 1.96 ce qui permet de déterminer la significativité des variables. Une seule variable n'est pas significative ; *AFE*, mais il sera décidé malgré tout de la conserver car pertinente pour la suite de l'analyse.

#### 2.2.2 Conclusion économétrique

À la différence du premier modèle, caractérisé uniquement par un  $R^2$  satisfaisant, les conditions nécessaires à l'établissement d'un BLUE sont remplies par ce nouveau modèle. Les ajustements effectués, tels que l'élimination des variables non significatives et l'introduction de variables indicatrices permettant de quantifier l'effet régionale permet d'en apprendre plus sur la quantité de CO2 émis par personne. Ainsi, il est possible de voir une élasticité presque proportionnelle entre l'intensité énergétique (*Ener\_GDP*) et l'augmentation des émissions de CO2 par personne. Une variation de 1% de la première entraîne une variation de 0.973% de la quantité de CO2 émis par personne. Les relations d'élasticités du modèle ne s'arrête pas là, au-delà de l'élasticité estimée des émissions de CO2 par personne en fonction de l'intensité énergétique (qui est égal à 0.973) le modèle éclaire sur l'élasticité estimée des émissions de CO2 par personne en fonction du revenu par habitant qui est de 1.04. Ainsi, une variation de 1% du revenu par habitant entraîne une variation de 1.04% de la quantité de CO2 émise par personne.

Les résultats obtenus concernant les variables, *Ener\_Intens* et *Density*, éclaire sur le fait qu'une augmentation nette de 1% de ces variables, toutes choses étant égales par ailleurs, devrait en moyenne décroître les émissions de CO2 par personne de respectivement 0.094% et 0.017%.

Quant aux résultats pour les variables indicatrices, il est nécessaire de se concentrer sur les changements relatifs, ainsi, la quantité de CO2 émise par personne en Europe et en Asie Centrale est environ  $(\exp(-0.257) - 1) \times 100\% = 22.66\%$  moins élevée que dans le reste du monde<sup>13</sup>, elle est  $(\exp(-0.228) - 1) \times 100\% = 20.39\%$  moins élevée en Amérique latine et Caraïbes que dans le reste du monde et elle est  $(\exp(-0.16) - 1) \times 100\% = 14.79\%$  moins élevée en Afrique de l'Est et du Sud que dans le reste du monde.

Dans cette perspective d'amélioration continue, le Modèle 3 est envisagé comme une nouvelle étape

12. Annexe 11, valeurs en gras

13. Le reste du monde concerne toutes les régions qui ne sont pas conservés dans le modèle

dans la quête de compréhension. Ce modèle vise spécifiquement à explorer l'hypothèse de la courbe environnementale de Kuznets[4], ajoutant ainsi une dimension cruciale à la compréhension des interactions entre le revenu et les émissions de CO2.

### 3 Modèle et théorie économique

#### 3.1 Modèle EKC

Dans cette partie, le modèle 2 va être transformé. La variable *Ener\_Intens* va se voir supprimer et la variable  $\ln(\text{Income})$  va devenir *centered*\_ $\ln(\text{Income})$ . Le but de cette transformation va être d'ajouter une nouvelle variable noté *centered*\_ $\ln(\text{Income})^2$  permettant de tester l'hypothèse de la courbe environnementale de Kuznets :

$$\begin{aligned} \ln(\text{CO2globalC}) = & \beta_1 \ln(\text{Ener\_GDP}) + \beta_2 \text{Density} \\ & + \beta_3 \text{ECS} + \beta_4 \text{AFE} + \beta_5 \text{LCN} \\ & + \beta_6 \text{centered\_}\ln(\text{Income})^2 + \beta_7 \text{centered\_}\ln(\text{Income}) + \epsilon \end{aligned} \quad (3)$$

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	t de Student	p. critique
l_Ener_GDP	0,834791	0,0273645	30,51	0,0000
Density	-0,000179502	3,91802e-005	-4,581	0,0000
ECS	-0,218249	0,0626637	-3,483	0,0006
AFE	-0,179294	0,0755085	-2,374	0,0187
LCN	-0,270673	0,0633452	-4,273	0,0000
sq_centered_ln_income	-0,105778	0,0215494	-4,909	0,0000
centered_ln_income	1,10181	0,0300865	36,62	0,0000
$F(7, 166)$	637,6507	P. critique (F)	1,9e-116	
$R^2$	0,964143	$R^2$ ajusté	0,954957	

TABLE 7 – Modèle 3 : MCO, hypothèse EKC

Avec un  $R^2 = 0,964143$ , ce nouveau modèle semble mieux spécifié que les deux autres. Pour les mêmes raisons que les autres modèles, conclure sur les relations causales qui lient les variables expliquées et explicatives n'est pas possible. Il ne permet que de statuer sur la nature corrélée entre régresseurs et variable dépendante et notamment sur la relation entre le revenu par habitant et l'émission de CO2 par habitant qui permettra d'affirmer ou refuser l'hypothèse de la courbe environnementale de Kuznets. De plus avec un  $F_c = 637,6507 > F(7, 166)$ , l'hypothèse  $H_0 : \rho^2 = 0$  est rejeté au risque de première espèce  $\alpha = 0,05$  le modèle est significativement différent de 0 ce qui se traduit par le fait qu'il possède au moins un coefficient significativement différent de 0.

### Absence de multicollinéarité

Variable	Valeur
l_Ener_GDP	1,410
Density	1,141
ECS	1,581
AFE	1,283
LCN	1,325
sq_centered_ln_income	1,304
centered_ln_income	1,828

TABLE 8 – Facteurs d'inflation de variance

L'observation des facteurs d'inflation de variance permet de déterminer l'absence de multicollinéarité dans le modèle. Toutes les valeurs étant très éloignées du seuil limite de 10.

### Présence d'homoscédasticité

A l'image du modèle 2, le test de *White* sera utilisé en premier lieu et validé avec le test de *Breusch-Pagan* considéré comme plus efficace pour déterminer l'hétéroscédasticité.

Ainsi, le premier test donne une statistique de test :  $TR^2 = 30,169414$ , avec p. critique =  $P(\chi^2(28) > 30,169414) = 0,355165$  qui permet de conclure en faveur de l'hypothèse  $H_0$  : le modèle ne présente pas d'hétéroscédasticité.

Quant au test de *Breusch-Pagan*, il donne une statistique de test :  $LM = 9,172822$ , avec p. critique =  $P(\chi^2(7) > 9,172822) = 0,240485$  validant aussi l'absence d'hétéroscédasticité dans le modèle.

### Significativité des paramètres

L'entière des variables du modèle 3 est significatif. Ainsi, il est possible de voir que tous les t. de Student associée à chaque variable sont supérieurs au seuil de 1.96.

Ainsi, toutes les variables du modèle 3 sont pertinentes à l'explication de la variable explicative dans le cadre de l'échantillon de donnée.

## 3.2 Conclusion économique

Le modèle met en évidence une élasticité entre la variable  $\ln(\text{Ener\_GDP})$  et la  $\ln(\text{CO2globalC})$ . En effet, une augmentation de 1% de la première conduira à une augmentation de 0.834791% de la seconde mettant en avant la sensibilité des émissions de CO2 par personne face aux variations de l'intensité énergétique du PIB, toutes choses égales par ailleurs.

Dans un second temps, le modèle éclaire sur une élasticité presque proportionnelle une variation de 1% du coefficient  $\text{centered\_ln}(\text{Income})$  entraîne quant à lui une augmentation de 1.10181% des émissions de CO2 par personne. Cependant, une variation de 1% de  $\text{centered\_ln}(\text{Income})^2$  entraîne une baisse de 0.106% des émissions de CO2 par personne. Cette dualité élasticité croissante pour le premier et décroissante pour le second éclaire sur un point de retournement ou, à partir d'un certain revenu atteint, les émissions de CO2 par personne ne baisseraient en fonction du revenu, observation en lien avec la théorie de la courbe environnementale de Kuznets.

Pour finir, il est possible de voir l'importance des effets régionaux dans le modèle : les trois variables indicatrices sont significatifs montrant que ce sont des régions qui ont tendances à moins polluer que la région de référence (reste du monde), pour rentrer dans les détails : la quantité de CO2 émise par personne en Amérique Latine et dans les Caraïbes est réduite d'environ  $(\exp(-0.270673) - 1) \times 100\% =$

23.71% par rapport à notre région de référence, elle est aussi réduite de 16.41% pour l'Afrique de l'Est et du Sud et réduite de 19.61% pour l'Europe et l'Asie centrale quand la comparaison des émissions de CO2 par personne se fait avec le reste du monde. Les raisons de cette émission plus faible ne peut pas être conclu par l'analyse économétrique seulement ; les économies de ces régions étant très vastes, il se peut que le modèle compare des économies au début de la courbe environnementale de Kuznets et des économies qui ont déjà passés le point de retournement.

## Références

- [1] R. M. ANDREW et G. P. PETERS. *The Global Carbon Project's fossil CO2 emissions dataset. Version 2023v36*. [Data set]. 2023. DOI : 10.5281/zenodo.10177738. URL : <https://doi.org/10.5281/zenodo.10177738>.
- [2] Trevor BREUSCH et Adrian PAGAN. "A simple test for heteroscedasticity and random coefficient variation". In : *Econometrica* 47.5 (sept. 1979), p. 1287. DOI : 10.2307/1911963. URL : <https://doi.org/10.2307/1911963>.
- [3] Ronald Aylmer FISHER. *Statistical methods for research workers*. Oliver Boyd, jan. 1970.
- [4] Gene M. GROSSMAN et Alan KRUEGER. "Economic Growth and the Environment". In : *The Quarterly Journal of Economics* 110.2 (1995), p. 353-377. URL : <https://EconPapers.repec.org/RePEc:oup:qjecon:v:110:y:1995:i:2:p:353-377..>
- [5] *International - U.S. Energy Information Administration (EIA)*. URL : <https://www.eia.gov/international/data/world>.
- [6] Edie PURDIE, Neil FANTOM et Tariq KHOKHAR. *The 2016 edition of World Development Indicators is out : Three features you won't want to miss*. Août 2019. URL : <https://blogs.worldbank.org/opendata/2016-edition-world-development-indicators-out-three-features-you-won-t-want-miss>.
- [7] Halbert WHITE. "A Heteroskedasticity-Consistent covariance matrix estimator and a direct test for heteroskedasticity". In : *Econometrica* 48.4 (mai 1980), p. 817. DOI : 10.2307/1912934. URL : <https://doi.org/10.2307/1912934>.
- [8] Jeffrey M. WOOLDRIDGE. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts et London, England : The MIT Press, 2002.
- [9] *World Development Indicators | DataBank*. URL : <https://databank.worldbank.org/source/world-development-indicators/Series/EN.POP.DNST>.

## A Statistiques descriptives

### A.1 Glossaire

#### Zone géographique

- SAS (Asie du Sud) : Afghanistan, Bangladesh, Bhoutan, Inde, Maldives, Népal, Pakistan, Sri Lanka.
- ECS (Europe et Asie centrale) : Albanie, Arménie, Autriche, Azerbaïdjan, Bélarus, Belgique, Bosnie-Herzégovine, Bulgarie, Croatie, Chypre, République tchèque, Danemark, Estonie, Finlande, France, Géorgie, Allemagne, Grèce, Hongrie, Irlande, Italie, Kazakhstan, Kirghizistan, Lettonie, Lituanie, Luxembourg, Macédoine du Nord, Moldavie, Mongolie, Monténégro, Pays-Bas, Norvège, Pologne, Portugal, Roumanie, Russie, Serbie, Slovaquie, Slovénie, Espagne, Suède, Suisse, Tadjikistan, Turquie, Ukraine, Ouzbékistan.
- MEA (Moyen-Orient et Afrique du nord) : Algérie, Bahreïn, Djibouti, Égypte, Iran, Irak, Israël, Jordanie, Koweït, Liban, Libye, Maroc, Oman, Qatar, Arabie Saoudite, Syrie, Tunisie, Émirats arabes unis, Yémen.
- AFE (Afrique de l'Est et du Sud) : Angola, Botswana, Burundi, Comores, Eswatini, Éthiopie, Kenya, Lesotho, Madagascar, Malawi, Mozambique, Namibie, Rwanda, Sao Tomé-et-Principe, Afrique du Sud, Soudan, Tanzanie, Ouganda, Zambie, Zimbabwe.
- LCN (Amérique latine et Caraïbes) : Antigua-et-Barbuda, Argentine, Barbade, Belize, Bolivie, Brésil, Chili, Colombie, Costa Rica, Cuba, Dominique, République dominicaine, Équateur, Salvador, Grenade, Guatemala, Guyana, Haïti, Honduras, Jamaïque, Mexique, Nicaragua, Panama, Paraguay, Pérou, Saint-Christophe-et-Niévès, Sainte-Lucie, Saint-Vincent-et-les-Grenadines, Suriname, Bahamas, Trinité-et-Tobago, Uruguay.
- AFW (Afrique de l'Ouest et central) : Bénin, Burkina Faso, Cap-Vert, Cameroun, République centrafricaine, Tchad, Congo-Brazzaville, Côte d'Ivoire, Guinée équatoriale, Gabon, Gambie, Ghana, Guinée, Guinée-Bissau, Libéria, Mali, Mauritanie, Niger, Nigéria, Sénégal, Sierra Leone, Togo.
- NAC (Amérique du nord) : Canada, États-Unis.

### A.2 Graphique de dispersion

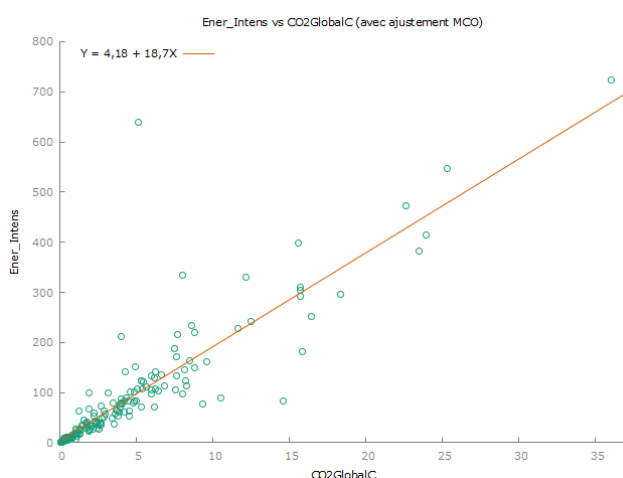


FIGURE 1 – Ener\_Intens et CO2GlobalC



FIGURE 2 – Ener\_GDP et CO2GlobalC

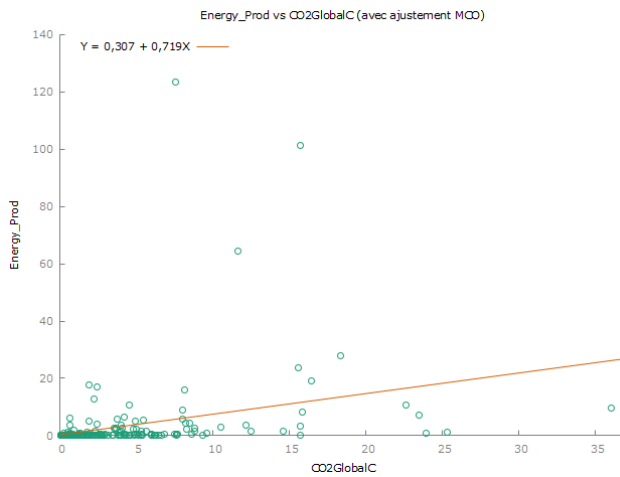


FIGURE 3 – Energy\_Prod et CO2GlobalC

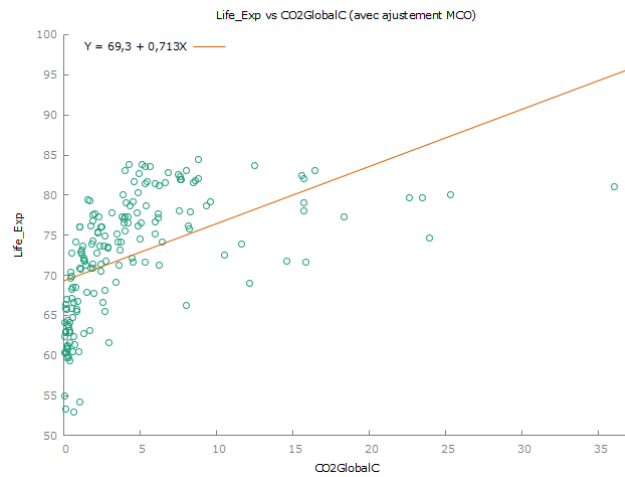


FIGURE 4 – Life\_Exp et CO2GlobalC

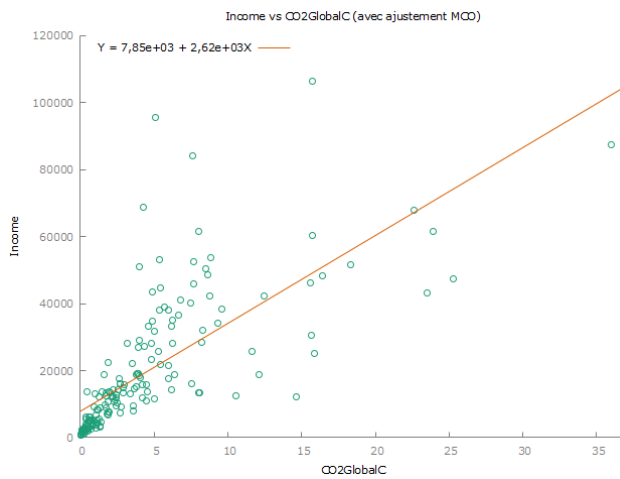


FIGURE 5 – Income et CO2GlobalC

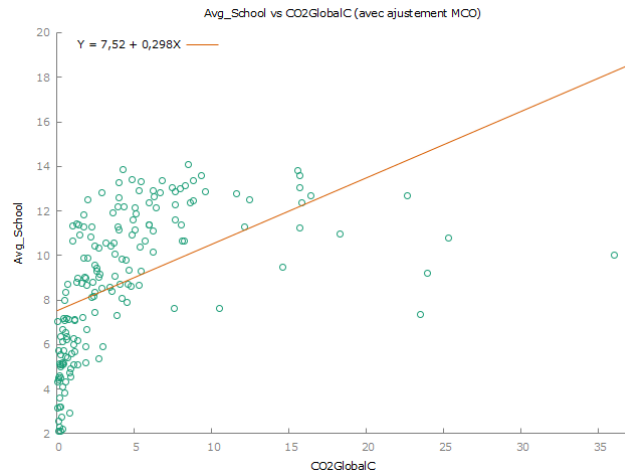


FIGURE 6 – Avg\_School et CO2GlobalC

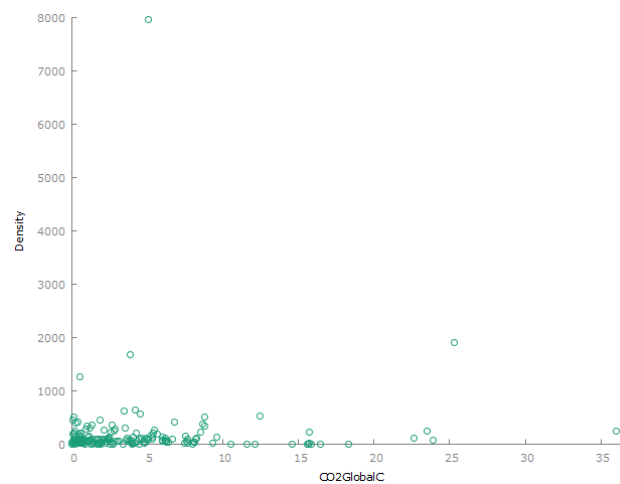


FIGURE 7 – Density et CO2GlobalC



## B Modèle 1

### B.1 Modèle 1

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	<i>t</i> de Student	p. critique
const	−9,53903	0,389068	−24,52	0,0000
l_Ener_GDP	0,996439	0,0583246	17,08	0,0000
Ener_Intens	−0,000729482	0,000376257	−1,939	0,0542
Energy_Prod	0,00254310	0,00195090	1,304	0,1942
Density	−0,000147176	4,47531e−005	−3,289	0,0012
l_Income	1,05810	0,0577391	18,33	0,0000
Life_Exp	−0,00784234	0,00658790	−1,190	0,2356
Avg_School	−0,0172924	0,0152100	−1,137	0,2572
Moyenne var. dép.	0,720946	Éc. type var. dép.	1,428474	
Somme carrés résidus	17,80468	Éc. type régression	0,328492	
$R^2$	0,949270	$R^2$ ajusté	0,947118	
$F(7, 165)$	441,0772	P. critique ( $F$ )	2,6e−103	
Log de vraisemblance	−48,79006	Critère d'Akaike	113,5801	
Critère de Schwarz	138,8064	Hannan–Quinn	123,8143	

TABLE 9 – Modèle 1

#### B.1.1 Multicolinéarité

Variable	Valeur
ln(Ener_GDP)	1,727
Ener_Intens	3,060
Energy_Prod	1,102
Density	1,310
ln(Income)	6,869
Life_Exp	3,819
Avg_School	3,732

TABLE 10 – Facteurs d'inflation de variance, modèle 1

Absence de multicolinéarité.

#### B.1.2 Test pour l'hétéroscédasticité

Test de *White*  $R^2$  non-ajusté = 0,301065

Statistique de test :  $TR^2 = 52,084215$ , avec p. critique =  $P(\chi^2(35) > 52,084215) = 0,031601$

Hypothèse  $H_0$  : présence d'homoscédasticité rejetée. Le modèle présente de l'hétéroscédasticité.

MCO, utilisant les observations 1-173

Variable dépendante :  $\hat{u}^2$ 

Variable	Coefficient	Éc. Type	t de Student	p. critique
const	4,33522	4,99906	0,8672	0,3873
l_Ener_GDP	-1,26352	1,11710	-1,131	0,2600
Ener_Intens	0,0624350	0,0821724	0,7598	0,4487
Energy_Prod	-0,0812883	0,187194	-0,4342	0,6648
Density	-0,000436741	0,00149397	-0,2923	0,7705
l_Income	-1,00353	1,22521	-0,8191	0,4142
Life_Exp	-0,0298694	0,0786951	-0,3796	0,7049
Avg_School	0,651362	0,222426	2,928	<b>0,0040</b>
sq_l_Ener_GDP	0,177698	0,106221	1,673	0,0966
X2_X3	-0,00587584	0,00613104	-0,9584	0,3396
X2_X4	0,00287724	0,0140352	0,2050	0,8379
X2_X5	0,000172714	0,000364760	0,4735	0,6366
X2_X6	0,178668	0,158072	1,130	0,2603
X2_X7	0,000907861	0,0126620	0,07170	0,9429
X2_X8	-0,111588	0,0224049	-4,981	<b>1,88e-06</b>
sq_Ener_Intens	9,97776e-06	6,79932e-06	1,467	0,1445
X3_X4	-0,000107481	0,000124396	-0,8640	0,3891
X3_X5	1,58285e-07	1,30204e-06	0,1216	0,9034
X3_X6	-0,00601794	0,00600387	-1,002	0,3179
X3_X7	7,84441e-05	0,000113378	0,6919	0,4902
X3_X8	0,000606795	0,000199653	3,039	<b>0,0028</b>
sq_Energy_Prod	3,58953e-05	6,67540e-05	0,5377	0,5916
X4_X5	1,81658e-05	4,49554e-05	0,4041	0,6868
X4_X6	0,0160580	0,0231658	0,6932	0,4894
X4_X7	-0,00111828	0,00123256	-0,9073	0,3659
X4_X8	0,00109830	0,00204006	0,5384	0,5912
sq_Density	4,38895e-08	6,96914e-08	0,6298	0,5299
X5_X6	-0,000335807	0,000299991	-1,119	0,2649
X5_X7	4,58653e-05	3,74345e-05	1,225	0,2226
X5_X8	-3,53247e-05	6,96774e-05	-0,5070	0,6130
sq_l_Income	0,0230288	0,0864913	0,2663	0,7904
X6_X7	0,0132894	0,0109155	1,217	0,2255
X6_X8	-0,0761596	0,0270578	-2,815	<b>0,0056</b>
sq_Life_Exp	-0,000681583	0,000793085	-0,8594	0,3916
X7_X8	-0,000830157	0,00293556	-0,2828	0,7778
sq_Avg_School	0,0115151	0,00541472	2,127	<b>0,0352</b>

TABLE 11 – Résultats du Test de White pour l'hétéroscédasticité, modèle 1

**B.1.3 Normalité des erreurs**Nombre de classes = 13, Moyenne =  $-1,11215 \times 10^{-15}$ , Éc. Type = 0,328492Test de l'hypothèse nulle de normalité de la distribution :  $\chi^2(2) = 23,122$  avec  $p$  critique 0,00001.

Intervalle	Centre	Fréquence	Rel.	Cum.
$< -1,0513$	-1,1465	1	0,58%	0,58%
$-1,0513 - -0,86095$	-0,95615	3	1,73%	2,31%
$-0,86095 - -0,67055$	-0,76575	2	1,16%	3,47%
$-0,67055 - -0,48015$	-0,57535	6	3,47%	6,94%
$-0,48015 - -0,28976$	-0,38496	11	6,36%	13,29%
$-0,28976 - -0,099358$	-0,19456	35	20,23%	33,53%
$-0,099358 - 0,091040$	-0,0041590	47	27,17%	60,69%
$0,091040 - 0,28144$	0,18624	44	25,43%	86,13%
$0,28144 - 0,47184$	0,37664	16	9,25%	95,38%
$0,47184 - 0,66223$	0,56704	4	2,31%	97,69%
$0,66223 - 0,85263$	0,75743	3	1,73%	99,42%
$0,85263 - 1,0430$	0,94783	0	0,00%	99,42%
$\geq 1,0430$	1,1382	1	0,58%	100,00%

TABLE 12 – Distribution des fréquences pour les résidus, modèle 1

**Test de l'hypothèse nulle de normalité de la distribution :**

$\chi^2(2) = 25,687$  avec p. critique  $< 0,00001$

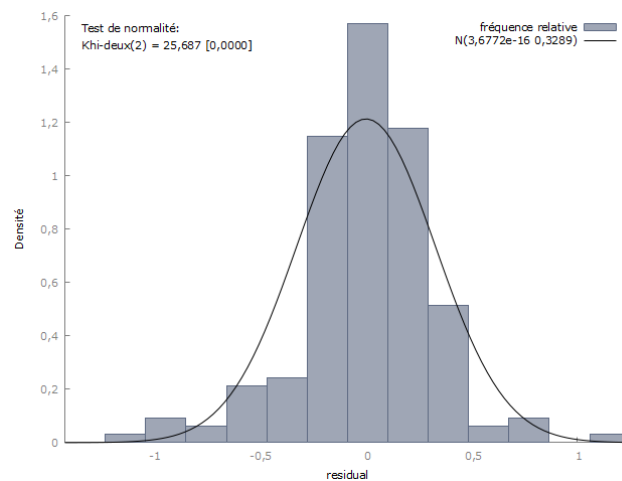


FIGURE 8 – Graphique distributions résidus, modèle 1

**B.2 Modèle 1 corrigé avec les hypothèses données**

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	<i>t</i> de Student	p. critique
const	−9,41776	0,279520	−33,69	0,0000
l_Ener_GDP	0,971549	0,0553873	17,54	0,0000
Ener_Intens	−0,000489834	0,000363254	−1,348	0,1793
Density	−0,000160421	4,41539e−005	−3,633	0,0004
l_Income	0,969990	0,0318020	30,50	0,0000
Moyenne var. dép.	0,720946	Éc. type var. dép.	1,428474	
Somme carrés résidus	18,36316	Éc. type régression	0,330612	
$R^2$	0,947679	$R^2$ ajusté	0,946433	
$F(4, 168)$	760,7399	P. critique ( $F$ )	1,9e−106	
Log de vraisemblance	−51,46161	Critère d'Akaike	112,9232	
Critère de Schwarz	128,6897	Hannan–Quinn	119,3196	

TABLE 13 – Modèle 1.2 (1 corrigé)

**B.2.1 Multicolinéarité**

Variable	Valeur
ln(Ener_GDP)	1,538
Ener_Intens	2,816
Density	1,259
ln(Income)	2,057

TABLE 14 – Facteurs d'inflation de variance, modèle 1.2

**B.2.2 Test pour l'hétéroscédasticité**

Test de *White*  $R^2$  non-ajusté = 0,080038 Statistique de test :  $TR^2 = 13,846591$ , avec p. critique =  $P(\chi^2(14) > 13,846591) = 0,461202$

Hypothèse  $H_0$  acceptée, le modèle présente de l'homoscédasticité

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	5,54112	4,39485	1,261	0,2092
l_Ener_GDP	-0,994580	1,06718	-0,9320	0,3528
Ener_Intens	-0,0279807	0,0718564	-0,3894	0,6975
Density	-0,000184279	0,00107076	-0,1721	0,8636
l_Income	-1,26252	1,10676	-1,141	0,2557
sq_l_Ener_GDP	0,156852	0,0934492	1,678	0,0952
X2_X3	0,00148703	0,00542033	0,2743	0,7842
X2_X4	0,000129939	0,000314315	0,4134	0,6799
X2_X5	0,102482	0,135502	0,7563	0,4506
sq_Ener_Intens	-1,24261e-07	6,23407e-06	-0,01993	0,9841
X3_X4	-2,43880e-07	9,04746e-07	-0,2696	0,7879
X3_X5	0,00198868	0,00542931	0,3663	0,7146
sq_Density	1,82484e-08	5,26584e-08	0,3465	0,7294
X4_X5	-4,70695e-06	0,000132575	-0,03550	0,9717
sq_l_Income	0,0739229	0,0706380	1,047	0,2969

TABLE 15 – Résultats du test de White pour l'hétéroscédasticité, modèle 1.2

Test de *Breusch-Pagan*

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	2,39187	1,72403	1,387	0,1672
l_Ener_GDP	0,538420	0,341619	1,576	0,1169
Ener_Intens	0,000495556	0,00224048	0,2212	0,8252
Density	-0,000112468	0,000272334	-0,4130	0,6801
l_Income	-0,220442	0,196149	-1,124	0,2627

TABLE 16 – Résultats du test de Breusch-Pagan pour l'hétéroscédasticité, modèle 1.2

Somme des carrés expliquée = 15,5662 Statistique de test :  $LM = 7,783109$ , avec p. critique =  $P(\chi^2(4) > 7,783109) = 0,099854$

Hypothèse  $H_0$  : présence d'homoscédasticité acceptée

**B.2.3 Normalité des erreurs**

nombre de classes = 13, moyenne =  $8,31063 \times 10^{-16}$ , éc. type = 0,330612

Intervalle	Centre	Fréquence	Rel.	Cum.
$< -1,0513$	-1,1465	1	0,58%	0,58%
$-1,0513 - -0,86095$	-0,95615	3	1,73%	2,31%
$-0,86095 - -0,67055$	-0,76575	2	1,16%	3,47%
$-0,67055 - -0,48015$	-0,57535	6	3,47%	6,94%
$-0,48015 - -0,28976$	-0,38496	11	6,36%	13,29%
$-0,28976 - -0,099358$	-0,19456	35	20,23%	33,53%
$-0,099358 - 0,091040$	-0,0041590	47	27,17%	60,69%
$0,091040 - 0,28144$	0,18624	44	25,43%	86,13%
$0,28144 - 0,47184$	0,37664	16	9,25%	95,38%
$0,47184 - 0,66223$	0,56704	4	2,31%	97,69%
$0,66223 - 0,85263$	0,75743	3	1,73%	99,42%
$0,85263 - 1,0430$	0,94783	0	0,00%	99,42%
$\geq 1,0430$	1,1382	1	0,58%	100,00%

TABLE 17 – Distribution des fréquences pour les résidus, modèle 1.2

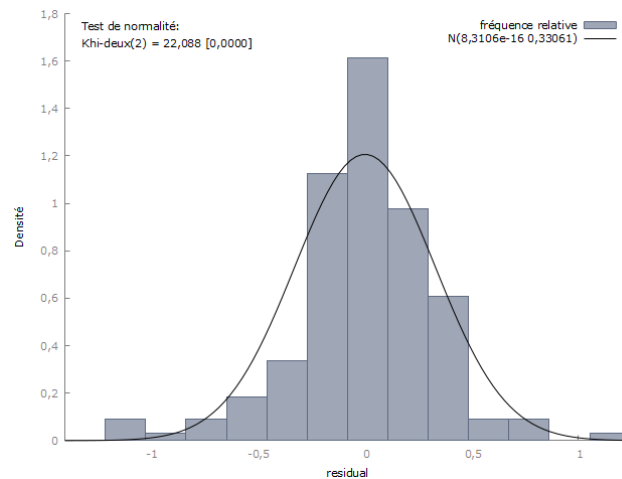


FIGURE 9 – Graphique distributions résidus, modèle 1.2

Test de l'hypothèse nulle de normalité de la distribution :  $\chi^2(2) = 23,122$  avec  $p$  critique 0,00001. Hypothèse de normalité refusé. Les résidus ne sont pas distribués en suivant une loi normale.

## C Modèle 2

### C.1 Modèle préambule

Modèle préambule au modèle 2.

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	<i>t</i> de Student	p. critique
const	−9,72138	0,380941	−25,52	0,0000
l_Ener_GDP	0,945617	0,0596852	15,84	0,0000
Ener_Intens	−0,000924269	0,000398403	−2,320	0,0216
Density	−0,000168143	4,49360e−005	−3,742	0,0003
l_Income	1,02455	0,0405462	25,27	0,0000
SAS	−0,0550719	0,134274	−0,4101	0,6822
ECS	−0,245130	0,0879902	−2,786	0,0060
MEA	0,0956957	0,104406	0,9166	0,3607
AFE	−0,190899	0,107501	−1,776	0,0777
LCN	−0,228813	0,0908101	−2,520	0,0127
AFW	−0,0723493	0,110249	−0,6562	0,5126
NAC	−0,124521	0,248135	−0,5018	0,6165
Moyenne var. dép.	0,720946	Éc. type var. dép.	1,428474	
Somme carrés résidus	16,36221	Éc. type régression	0,317807	
$R^2$	0,953380	$R^2$ ajusté	0,950503	
$F(11, 161)$	331,2930	P. critique ( $F$ )	2,4ee−102	
Log de vraisemblance	−41,48193	Critère d'Akaike	104,9639	
Critère de Schwarz	139,6501	Hannan–Quinn	119,0358	

TABLE 18 – Modèle préambule au modèle 2, avec toutes les variables indicatrices

Variable *NAC*, *AFW*, *MEA*, *SAS* non significatives, leur ensemble sera la référence "reste du monde". *AFE* gagne en significativité en retirant les autres variables.

### C.1.1 Multicolinéarité

Variable	Valeur
l_Ener_GDP	1,924
Ener_Intens	3,649
Density	1,405
l_Income	3,602
SAS	1,356
ECS	2,612
MEA	1,732
AFE	2,187
LCN	2,068
AFW	2,300
NAC	1,200

TABLE 19 – Facteurs d'inflation de variance, modèle préambule

Pas de multicolinéarité.

### C.1.2 Test pour l'hétéroscédasticité

Test de *White*

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	5,73434	5,82760	0,9840	0,3270
l_Ener_GDP	-1,64943	1,47873	-1,115	0,2668
Ener_Intens	-0,0339868	0,0840107	-0,4046	0,6865
Density	-0,000903197	0,00143117	-0,6311	0,5291
l_Income	-1,29071	1,36843	-0,9432	0,3474
SAS	1,38593	1,87013	0,7411	0,4600
ECS	1,62259	2,20447	0,7360	0,4631
MEA	2,36753	1,50982	1,568	0,1194
AFE	0,415983	1,53861	0,2704	0,7873
LCN	-0,358418	1,49528	-0,2397	0,8110
AFW	-1,00991	1,67984	-0,6012	0,5488
NAC	-0,182386	1,00192	-0,1820	0,8558
sq_l_Ener_GDP	0,152415	0,124746	1,222	0,2241
X2_X3	0,00180644	0,00631654	0,2860	0,7754
X2_X4	-0,000302980	0,000402782	-0,7522	0,4533
X2_X5	0,194134	0,172026	1,129	0,2612
X2_X6	-0,0418747	0,309530	-0,1353	0,8926
X2_X7	-0,379374	0,282548	-1,343	0,1818
X2_X8	-0,384207	0,164053	-2,342	0,0208
X2_X9	-0,00893344	0,237963	-0,03754	0,9701
X2_X10	-0,0812250	0,199576	-0,4070	0,6847
X2_X11	0,0480837	0,254441	0,1890	0,8504
X2_X12	0,0742716	0,538996	0,1378	0,8906
sq_Ener_Intens	-1,82924e-06	7,34756e-06	-0,2490	0,8038
X3_X4	6,43688e-07	1,38497e-06	0,4648	0,6429
X3_X5	0,00240328	0,00635453	0,3782	0,7059
X3_X6	0,0102133	0,00689676	1,481	0,1411
X3_X7	0,00219045	0,00166991	1,312	0,1920
X3_X8	0,00211997	0,00123535	1,716	0,0886
X3_X9	-0,000648276	0,00551514	-0,1175	0,9066
X3_X10	-0,000914946	0,00324337	-0,2821	0,7783
X3_X11	-0,00716326	0,00879089	-0,8149	0,4167
sq_Density	2,68112e-08	9,21214e-08	0,2910	0,7715
X4_X5	7,93609e-05	0,000170090	0,4666	0,6416
X4_X6	0,000214258	0,000332870	0,6437	0,5210
X4_X7	-0,000309426	0,000471068	-0,6569	0,5125
X4_X8	0,000452459	0,000397419	1,138	0,2571
X4_X9	-6,10885e-05	0,000418658	-0,1459	0,8842
X4_X10	0,000286716	0,000401019	0,7150	0,4760
X4_X11	0,000330437	0,000730539	0,4523	0,6518
sq_l_Income	0,0737638	0,0834568	0,8839	0,3785
X5_X6	-0,167545	0,216473	-0,7740	0,4404
X5_X7	-0,138490	0,204744	-0,6764	0,5000
X5_X8	-0,224262	0,161495	-1,389	0,1674
X5_X9	-0,0374774	0,172991	-0,2166	0,8288
X5_X10	0,0455868	0,163286	0,2792	0,7806
X5_X11	0,125094	0,195750	0,6391	0,5239

TABLE 20 – Résultats du test de White pour l'hétéroscédasticité, modèle préambule



$R^2$  non-ajusté = 0,364588.

Statistique de test :  $TR^2 = 63,073717$ , avec p. critique =  $P(\chi^2(46) > 63,073717) = 0,047903$

Hypothèse  $H_0$  : présence d'homoscédasticité refusée. Le modèle présente de l'hétéroscédasticité.

### C.1.3 Normalité des erreurs

Intervalle	Centre	Fréquence	Rel.	Cum.
$< -1,0242$	-1,1133	1	0,58%	0,58%
$-1,0242 - -0,84585$	-0,93502	3	1,73%	2,31%
$-0,84585 - -0,66752$	-0,75669	0	0,00%	2,31%
$-0,66752 - -0,48919$	-0,57835	8	4,62%	6,94%
$-0,48919 - -0,31086$	-0,40002	9	5,20%	12,14%
$-0,31086 - -0,13253$	-0,22169	22	12,72%	24,86%
$-0,13253 - 0,045801$	-0,043364	53	30,64%	55,49%
$0,045801 - 0,22413$	0,13497	43	24,86%	80,35%
$0,22413 - 0,40246$	0,31330	26	15,03%	95,38%
$0,40246 - 0,58079$	0,49163	4	2,31%	97,69%
$0,58079 - 0,75912$	0,66996	1	0,58%	98,27%
$0,75912 - 0,93745$	0,84829	2	1,16%	99,42%
$\geq 0,93745$	1,0266	1	0,58%	100,00%

TABLE 21 – Distribution des fréquences pour les résidus, modèle préambule

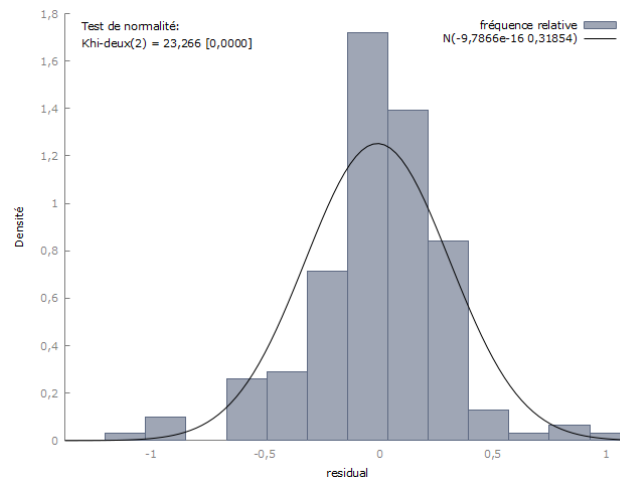


FIGURE 10 – Graphique distributions résidus, modèle préambule

Test de l'hypothèse nulle de normalité de la distribution :  $\chi^2(2) = 23,266$  avec p critique 0,00001.

Hypothèse de normalité refusée. Les résidus ne sont pas distribués en suivant une loi normale.

## C.2 Modèle 2

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	t de Student	p. critique
const	−9,87838	0,324902	−30,40	0,0000
l_Ener_GDP	0,972553	0,0548810	17,72	0,0000
Ener_Intens	−0,000939902	0,000375945	−2,500	0,0134
Density	−0,000171070	4,26358e−005	−4,012	0,0001
l_Income	1,03770	0,0374815	27,69	0,0000
ECS	−0,257481	0,0704806	−3,653	0,0003
AFE	−0,159957	0,0819597	−1,952	0,0527
LCN	−0,227537	0,0739383	−3,077	0,0024
Moyenne var. dép.	0,720946	Éc. type var. dép.	1,428474	
Somme carrés résidus	16,59514	Éc. type régression	0,317138	
$R^2$	0,952717	$R^2$ ajusté	0,950711	
$F(7, 165)$	474,9433	P. critique ( $F$ )	8,0e−106	
Log de vraisemblance	−42,70464	Critère d'Akaike	101,4093	
Critère de Schwarz	126,6356	Hannan–Quinn	111,6435	

TABLE 22 – Modèle 2

### C.2.1 Multicolinéarité

Variable	Valeur
ln(Ener_GDP)	1,641
Ener_Intens	3,278
Density	1,276
ln(Income)	3,105
ECS	1,691
AFE	1,282
LCN	1,383

TABLE 23 – Facteurs d'inflation de variance, modèle 2

Absence de multicolinéarité

### C.2.2 Test pour l'hétéroscédasticité

Test de *White*

$R^2$  non-ajusté = 0,182232 Statistique de test :  $TR^2 = 31,526135$ , avec p. critique =  $P(\chi^2(29) > 31,526135) = 0,341053$

Hypothèse  $H_0$  : présence d'homoscédasticité acceptée

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	0,803781	4,78150	0,1681	0,8667
l_Ener_GDP	-0,992652	1,19856	-0,8282	0,4089
Ener_Intens	0,0177891	0,0830359	0,2142	0,8307
Density	-0,000628238	0,00119161	-0,5272	0,5989
l_Income	-0,0874022	1,20328	-0,07264	0,9422
ECS	1,30092	2,28470	0,5694	0,5700
AFE	0,731380	0,803938	0,9097	0,3645
LCN	-0,265954	1,32829	-0,2002	0,8416
sq_l_Ener_GDP	0,155071	0,103006	1,505	0,1344
X2_X3	-0,00248261	0,00620107	-0,4004	0,6895
X2_X4	-0,000110696	0,000373941	-0,2960	0,7676
X2_X5	0,0940692	0,153015	0,6148	0,5397
X2_X6	-0,389898	0,277608	-1,404	0,1623
X2_X7	0,00327500	0,140021	0,02339	0,9814
X2_X8	-0,132668	0,176555	-0,7514	0,4536
sq_Ener_Intens	4,72348e-06	7,19015e-06	0,6569	0,5123
X3_X4	4,09086e-07	1,03452e-06	0,3954	0,6931
X3_X5	-0,00147348	0,00629443	-0,2341	0,8152
X3_X6	0,00203549	0,00168215	1,210	0,2283
X3_X7	-0,000340699	0,00413622	-0,08237	0,9345
X3_X8	-0,000388063	0,00302412	-0,1283	0,8981
sq_Density	-3,46083e-08	6,14918e-08	-0,5628	0,5744
X4_X5	7,53549e-05	0,000148968	0,5058	0,6137
X4_X6	-0,000721172	0,000419897	-1,717	0,0881
X4_X7	-0,000299244	0,000316696	-0,9449	0,3463
X4_X8	-0,000187289	0,000277646	-0,6746	0,5010
sq_l_Income	-0,000152102	0,0767914	-0,001981	0,9984
X5_X6	-0,0952108	0,211502	-0,4502	0,6533
X5_X7	-0,0741582	0,0980333	-0,7565	0,4506
X5_X8	0,0475088	0,144342	0,3291	0,7425

TABLE 24 – Résultats du test de White pour l'hétéroscédasticité, modèle 2

Test de *Breusch-Pagan*

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	0,719508	2,14105	0,3361	0,7373
l_Ener_GDP	0,668567	0,361656	1,849	0,0663
Ener_Intens	-0,000568049	0,00247742	-0,2293	0,8189
Density	-0,000149006	0,000280963	-0,5303	0,5966
l_Income	-0,0518188	0,246997	-0,2098	0,8341
ECS	-0,148919	0,464456	-0,3206	0,7489
AFE	0,772673	0,540100	1,431	0,1544
LCN	-0,0812856	0,487241	-0,1668	0,8677

TABLE 25 – Résultats du test de Breusch-Pagan pour l'hétéroscédasticité, modèle 2

Somme des carrés expliquée = 24,9313 Statistique de test : LM = 12,465654, avec p. critique =  $P(\chi^2(7) > 12,465654) = 0,086248$

Hypothèse  $H_0$  acceptée. Le modèle présente de l'homoscédasticité

### C.2.3 Normalité des erreurs

Intervalle	Centre	Fréquence	Rel.	Cum.
$< -1,1004$	-1,1921	1	0,58%	0,58%
$-1,1004 - -0,91691$	-1,0086	2	1,16%	1,73%
$-0,91691 - -0,73344$	-0,82518	1	0,58%	2,31%
$-0,73344 - -0,54998$	-0,64171	3	1,73%	4,05%
$-0,54998 - -0,36651$	-0,45824	12	6,94%	10,98%
$-0,36651 - -0,18304$	-0,27477	13	7,51%	18,50%
$-0,18304 - 0,00042915$	-0,091305	50	28,90%	47,40%
$0,00042915 - 0,18390$	0,092163	50	28,90%	76,30%
$0,18390 - 0,36737$	0,27563	29	16,76%	93,06%
$0,36737 - 0,55083$	0,45910	7	4,05%	97,11%
$0,55083 - 0,73430$	0,64257	2	1,16%	98,27%
$0,73430 - 0,91777$	0,82604	2	1,16%	99,42%
$\geq 0,91777$	1,0095	1	0,58%	100,00%

TABLE 26 – Distribution des fréquences pour les résidus, modèle 2

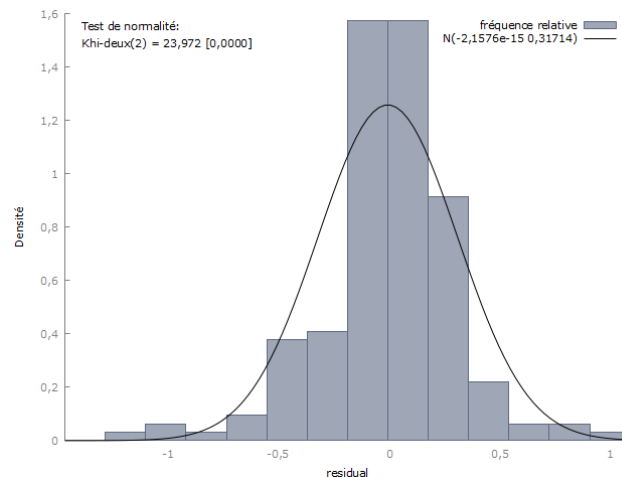


FIGURE 11 – Graphique distributions résidus, modèle 2

Test de l'hypothèse nulle de normalité de la distribution :  $\chi^2(2) = 23,972$  avec p critique 0,00001. Hypothèse de normalité refusée. Les résidus ne sont pas distribués en suivant une loi normale.

Variable dépendante : l_CO2GlobalC				
	Coefficient	Erreur std.	<i>t</i> de Student	p. critique
const	−0,0782468	0,0818283	−0,9562	0,3404
l_Ener_GDP	0,874257	0,0495241	17,65	0,0000
Density	−0,000177475	3,92476e−005	−4,522	0,0000
ECS	−0,197459	0,0663434	−2,976	0,0034
AFE	−0,154735	0,0797752	−1,940	0,0541
LCN	−0,241237	0,0704433	−3,425	0,0008
sq_centered_ln_income	−0,0955466	0,0240646	−3,970	0,0001
centered_ln_income	1,09192	0,0318214	34,31	0,0000
Moyenne var. dép.	0,720946	Éc. type var. dép.	1,428474	
Somme carrés résidus	15,72173	Éc. type régression	0,308680	
$R^2$	0,955205	$R^2$ ajusté	0,953305	
$F(7, 165)$	502,6379	P. critique ( $F$ )	9,3e−108	
Log de vraisemblance	−38,02794	Critère d'Akaike	92,05587	
Critère de Schwarz	117,2822	Hannan–Quinn	102,2900	

TABLE 27 – Modèle EKC

## D Modèle EKC

### D.1 Modèle

#### D.1.1 Multicolinéarité

Variable	Valeur
l_Ener_GDP	1,410
Density	1,141
ECS	1,581
AFE	1,283
LCN	1,325
sq_centered_ln_income	1,304
centered_ln_income	1,828

TABLE 28 – Facteurs d'inflation de variance, modèle EKC

Absence de multicolinéarité

#### D.1.1.1 Test pour l'hétéroscédasticité

Test de *White*  $R^2$  non-ajusté = 0,177978 Statistique de test :  $TR^2 = 30,790163$ , avec p. critique =  $P(\chi^2(28) > 30,790163) = 0,326477$

Hypothèse  $H_0$  : présence d'homoscédasticité acceptée

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	0,131288	0,125843	1,043	0,2986
l_Ener_GDP	-0,0943988	0,146761	-0,6432	0,5211
Density	5,59353e-05	0,000283740	0,1971	0,8440
ECS	0,242754	0,209278	1,160	0,2480
AFE	0,0637863	0,138124	0,4618	0,6449
LCN	0,0359148	0,185968	0,1931	0,8471
sq_centered_ln_income	-0,139130	0,0820708	-1,695	0,0922
centered_ln_income	-0,112591	0,103968	-1,083	0,2806
sq_l_Ener_GDP	0,0870885	0,0479503	1,816	0,0714
X2_X3	-0,000129803	0,000205768	-0,6308	0,5292
X2_X4	-0,223705	0,119557	-1,871	0,0634
X2_X5	-0,0725610	0,0971261	-0,7471	0,4562
X2_X6	-0,0648907	0,133686	-0,4854	0,6281
X2_X7	0,0140439	0,0368908	0,3807	0,7040
X2_X8	0,0523631	0,0563425	0,9294	0,3543
sq_Density	1,59093e-08	5,74903e-08	0,2767	0,7824
X3_X4	-0,000568618	0,000440749	-1,290	0,1991
X3_X5	-0,000124168	0,000312928	-0,3968	0,6921
X3_X6	-0,000129956	0,000248526	-0,5229	0,6018
X3_X7	-5,17875e-05	9,93290e-05	-0,5214	0,6029
X3_X8	0,000140446	0,000172660	0,8134	0,4173
X4_X7	0,153433	0,0925447	1,658	0,0995
X4_X8	0,0377579	0,0961358	0,3928	0,6951
X5_X7	0,0974430	0,0866666	1,124	0,2627
X5_X8	0,0835287	0,150974	0,5533	0,5809
X6_X7	0,100395	0,119914	0,8372	0,4039
X6_X8	0,0793814	0,0885419	0,8965	0,3715
sq_sq_centered_ln_income	0,0128320	0,0125726	1,021	0,3091
X7_X8	-0,0104408	0,0241034	-0,4332	0,6655

TABLE 29 – Résultats du test de White pour l'hétéroscédasticité, modèle EKC

Test de *Breusch-Pagan*

Variable	Coefficient	Éc. Type	T de Student	P. Critique
const	0,519048	0,555785	0,9339	0,3517
l_Ener_GDP	0,465704	0,336372	1,384	0,1681
Density	-0,000224871	0,000266573	-0,8436	0,4001
ECS	-0,257909	0,450610	-0,5724	0,5679
AFE	0,649352	0,541840	1,198	0,2325
LCN	-0,150863	0,478457	-0,3153	0,7529
sq_centered_ln_income	-0,0175124	0,163449	-0,1071	0,9148
centered_ln_income	0,0124725	0,216134	0,05771	0,9541

TABLE 30 – Résultats du test de Breusch-Pagan pour l'hétéroscédasticité, modèle EKC

Somme des carrés expliquée = 18,293 Statistique de test : LM = 9,146480, avec p. critique =  $P(\chi^2(7) > 9,146480) = 0,242310$

Hypothèse  $H_0$  acceptée. Le modèle présente de l'homoscédasticité

### D.1.2 Normalité de l'erreur

Intervalle	Centre	Fréquence	Rel.	Cum.
$< -1,0957$	-1,1873	1	0,58%	0,58%
$-1,0957 - -0,91270$	-1,0042	1	0,58%	1,16%
$-0,91270 - -0,72966$	-0,82118	2	1,16%	2,31%
$-0,72966 - -0,54662$	-0,63814	3	1,73%	4,05%
$-0,54662 - -0,36358$	-0,45510	9	5,20%	9,25%
$-0,36358 - -0,18054$	-0,27206	22	12,72%	21,97%
$-0,18054 - 0,0025023$	-0,089018	42	24,28%	46,24%
$0,0025023 - 0,18554$	0,094022	48	27,75%	73,99%
$0,18554 - 0,36858$	0,27706	35	20,23%	94,22%
$0,36858 - 0,55162$	0,46010	5	2,89%	97,11%
$0,55162 - 0,73466$	0,64314	2	1,16%	98,27%
$0,73466 - 0,91770$	0,82618	2	1,16%	99,42%
$\geq 0,91770$	1,0092	1	0,58%	100,00%

TABLE 31 – Distribution des fréquences pour les résidus, modèle EKC

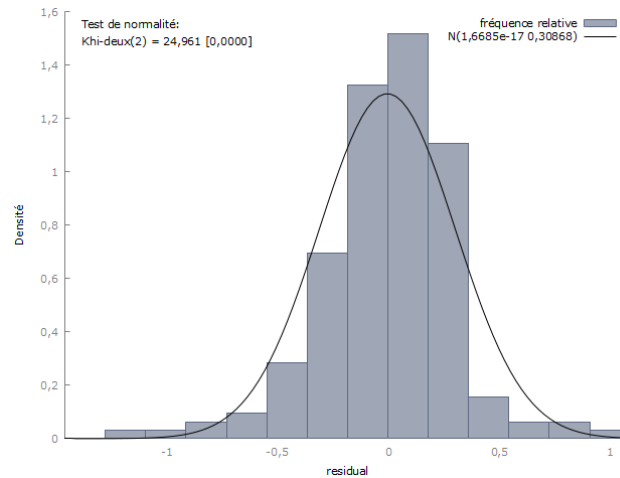


FIGURE 12 – Graphique distributions résidus, modèle EKC

Test de l'hypothèse nulle de normalité de la distribution :  $\chi^2(2) = 24,961$  avec p critique 0,00000. Hypothèse de normalité refusée. Les résidus ne sont pas distribués en suivant une loi normale.