

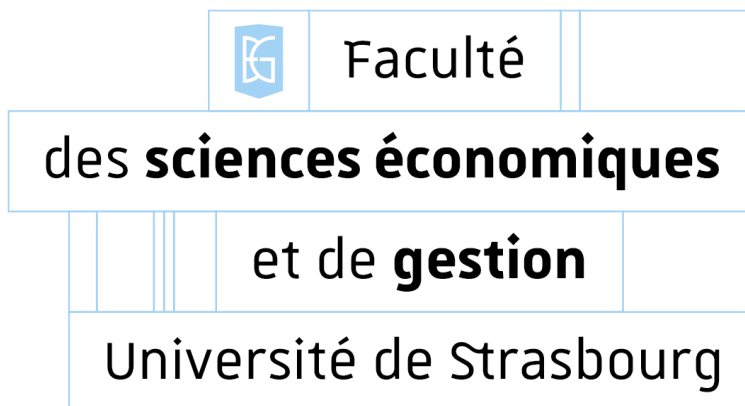
# Une étude de la convergence économique au sein des pays de L'OCDE : une analyse menée avec des données de panel

Nathan Schild

29 février 2024

## Table des matières

<b>Introduction</b>	<b>2</b>
<b>1 Présentation des données et statistiques descriptives</b>	<b>2</b>
1.1 Présentation des données . . . . .	3
1.2 Statistiques descriptives . . . . .	3
<b>2 Résultats des estimations économétriques</b>	<b>4</b>
2.1 Résultats des estimations . . . . .	4
2.2 Validité statistique du modèle . . . . .	5
2.2.1 Tests d'existence des effets spécifiques . . . . .	5
2.2.2 Test d'indépendance d'Hausman . . . . .	6
<b>Conclusion</b>	<b>6</b>



# Introduction

L'objectif de cette étude est d'analyser s'il existe un phénomène de convergence économique pour 23 pays de l'OCDE sur la période 1970-2015. Les résultats suggèrent bel et bien la présence d'un phénomène de convergence au sein de ce groupe de pays. Solow (1956) nous expose un modèle de croissance qui prédit une convergence conditionnelle entre les pays, cette forme de convergence se traduit par le fait que les pays ayant des caractéristiques structurelles identiques vont tendre vers le même niveau de développement économique. Empiriquement, la convergence se traduit par un coefficient négatif et significatif associé au niveau de revenu par habitant initial, cela signifie qu'un revenu par habitant initial élevé se traduira par un taux de croissance moindre et inversement. Il s'ensuit alors tout une littérature souhaitant vérifier cette convergence par diverses méthodes empiriques. Une des plus importantes études est celle de Baumol (1986) qui va étudier cette convergence sur une longue période (1870-1979) pour certains pays. Les résultats suggèrent effectivement cette convergence entre les pays et indiquent également qu'il existe des groupes de convergence, ce qui pourrait conforter cette hypothèse de convergence conditionnelle. Par la suite, Barro (1991) va également étudier cette question de la convergence économique pour 98 pays sur la période 1960-1985. Les résultats de cette étude suggèrent, en effet, une convergence entre les 98 pays présents dans l'étude. Enfin, Barro (2016) va mener une étude empirique encore plus large en travaillant sur des données de panel concernant 89 pays sur la période 1960-2010. Encore une fois, les résultats suggèrent une convergence entre les pays.

En se basant sur cette littérature, cette étude propose donc une étude empirique de la convergence économique au sein de 23 pays appartenant à l'OCDE sur la période 1970-2015 en utilisant des données de panel. Nous utiliserons principalement les méthodes et variables utilisées dans l'article de Barro (2016) et de Barro (1991) afin de vérifier cette hypothèse. La première section évoque les caractéristiques des données utilisées dans cette étude (sources, méthodes de calcul) ainsi que des statistiques descriptives afin d'en dégager les tendances principales. La deuxième section présente le modèle économétrique à estimer, les résultats des estimations OLS, Within et GLS puis une discussion des résultats dans laquelle nous considérerons des tests statistiques (F-test et LM test) afin de justifier l'existence des effets fixes et temporels puis un test d'indépendance (Hausman test) permettant de choisir entre l'estimation Within et GLS.

Le modèle considéré est alors le suivant :

$$g_{i,t,t-1} = \alpha \log(y_{i,t-1}) + X'_{i,t-1} \beta + u_{i,t}$$

où  $g_{i,t,t-1}$  est le taux de croissance annuel du revenu par habitant de l'économie  $i$  entre les dates  $t$  et  $t-1$ ,  $y_{i,t-1}$  est le revenu par tête à la date  $t-1$ ,  $X$  un vecteur de variables de contrôle que l'on considère à la date  $t-1$  également. Pour qu'il y ait une convergence entre les pays, il faut que  $\alpha$  soit négatif et significatif.

## 1 Présentation des données et statistiques descriptives

Cette section a pour but de décrire la méthodologie utilisée dans cette étude, les sources des données ainsi que de procéder à des statistiques descriptives. Cela nous permettra d'avoir un premier contact avec les données utilisées dans l'étude afin de mieux interpréter les résultats des estimations.

## 1.1 Présentation des données

Premièrement, il convient de noter que dans notre étude, la variable dépendante, le taux de croissance du PIB/habitant est calculé sur une période de 5 ans comme l'étude de [Barro \(2016\)](#). La période considérée ici est celle allant de 1970 à 2015, il y a donc 9 périodes dans notre étude. En effet, le taux de croissance est calculé sur les périodes suivantes : 1970-1975, 1975-1980, ..., 2010-2015. En prenant en compte le fait que nous étudions la convergence pour 23 pays de l'OCDE et que notre échantillon est cylindré (i.e. toutes les observations sont disponibles), nous avons  $N = 23$  et  $T = 9$  donc au total  $N * T = 207$  observations.

Conformément au modèle exposé en introduction, les variables de contrôle sont considérées durant la période  $t - 1$ , plus particulièrement dans le cadre de notre étude, cela correspond aux années 1970, 1975, 1980, ..., 2010.

Les différentes variables prises en compte dans cette étude sont similaires à celles des papiers de [Barro \(1991\)](#) et [Barro \(2016\)](#). En effet, nous considérerons les variables suivantes dans le cadre de cette étude :

- Le PIB/habitant ainsi que le taux de croissance du PIB/habitant proviennent de l'OCDE. Ils sont en prix constants en dollars des Etats-Unis où l'année de référence est 2010. Vous retrouverez les données présentes dans l'étude en cliquant sur ce [lien](#).
- La fécondité et l'espérance vie à la naissance proviennent de la base de données de la Banque mondiale et plus précisément sur la base concernant les indicateurs de développement dans le monde. Les données sont accessibles [ici](#).
- Les données concernant le nombre d'années d'étude proviennent de la base de données de Barro et Lee. Ici, nous prenons en compte le nombre d'années d'étude moyen de la population totale âgée de 15 à 64 ans. Vous pouvez retrouver les données sur l'éducation [ici](#).
- L'inflation provient de la base de données de l'OCDE. Les données sont accessibles [ici](#).
- Le degré d'ouverture (i.e.  $\frac{X+M}{Y}$ ) provient de la version 7.1 des Penn World Tables. Les données sont accessibles [ici](#).

## 1.2 Statistiques descriptives

Avant de passer aux résultats des estimations, il convient de considérer des statistiques descriptives concernant les variables d'intérêt. La Table 1 reporte les statistiques descriptives concernant les deux principales variables de l'étude. Comme énoncé précédemment, il n'y a pas de valeurs manquantes et il y a 207 observations. Il est intéressant de noter que les deux distributions ont une skewness positive, ce qui signifie que l'asymétrie est à droite donc que quelques valeurs remontent la moyenne vers le haut, on a donc bien la moyenne qui est supérieure à la médiane. Par ailleurs, la distribution du taux de croissance du PIB/habitant semble peu dispersée tandis que la distribution du PIB/habitant semble beaucoup plus dispersée. En effet, l'écart-type du taux de croissance du PIB/habitant est de l'ordre de 0.096 tandis que celui du PIB/habitant est largement plus élevé.

TABLE 1 – Statistiques descriptives

	growthrate	gdpcap
nobs	207	207
NAs	0	0
Minimum	-0.160	2,495.284
Maximum	0.566	85,514.870
1. Quartile	0.047	21,233.050
3. Quartile	0.153	36,121.570
Mean	0.108	29,316.940
Median	0.096	28,037.950
Variance	0.009	153,816,104.000
Stdev	0.096	12,402.260
Skewness	1.321	1.124
Kurtosis	3.712	3.364

## 2 Résultats des estimations économétriques

Dans cette section, nous estimerons les modèles OLS, WITHIN et GLS et nous les commenterons afin de constater si l’hypothèse de convergence conditionnelle est vérifiée ou non dans le cadre particulier de notre échantillon. Il convient de noter que dans le cadre des modèles Within et GLS, les résultats présentés dans la Table 2 prennent en compte les effets individuels et temporels dont nous justifierons le choix dans la sous-section 2.2. L’existence des deux effets conduit à ne pas considérer l’estimateur Between dans cette étude. Par la suite, nous mènerons trois tests différents, les deux premiers vont permettre de tester l’existence des effets individuels et temporels. Nous verrons que les deux premiers tests vont confirmer l’existence des deux effets, ce qui va nous pousser à rejeter le modèle des MCO. Enfin, nous mènerons un test d’Hausman afin de savoir si les effets sont fixes ou aléatoires, si les effets sont fixes, nous considérerons le modèle WITHIN sinon le modèle GLS. Il convient de noter que les 3 modèles ne sont pas sujets à de l’hétéroscédasticité car nous avons adopté des écarts-types robustes.

### 2.1 Résultats des estimations

La Table 2 résume les résultats des estimations OLS, WITHIN et GLS. En guise de rappel pour le lecteur, l’hypothèse de convergence conditionnelle sera vérifiée si le coefficient associé au logarithme du PIB/habitant à la date  $t - 1$  est négatif et significatif. Par la suite, nous ne commenterons que le modèle Within, cela se justifie dans un premier temps par le fait que les deux effets existent, ce qui revient à rejeter le modèle des MCO. De plus, d’après les résultats du test d’Hausman, nous pouvons considérer les effets comme fixes, ce qui revient à ne pas commenter l’estimation GLS. En nous fiant à la Table 2, nous pouvons aisément accepter l’hypothèse de convergence conditionnelle au sein de cet échantillon. En effet, dans le modèle Within, le coefficient associé est négatif et significatif au seuil de 1%. Cela voudrait signifier économiquement qu’un PIB/habitant initial plus élevé à la date  $t - 1$  engendre un taux de croissance du PIB/habitant plus faible. Les résultats de cette étude suggèrent donc qu’il existe un phénomène de convergence au sein du groupe de pays étudié.

De plus, toutes les variables sont significatives au seuil de 10% sauf le coefficient associé au

logarithme de la fécondité. Notons que dans le cadre du modèle Within, les signes des coefficients sont plutôt intuitifs sauf celui associé à l'éducation. En effet, un degré d'ouverture plus élevé implique une plus grande part des échanges dans le PIB, ce qui implique un taux de croissance plus élevé ou encore une inflation plus forte implique un taux de croissance plus faible. Par ailleurs, le modèle capte 29.5% des variations du taux de croissance du PIB/habitant. Néanmoins, le coefficient associé à l'inverse de l'espérance de vie à la naissance semble élevé. De plus, le coefficient de détermination semble assez faible, cela peut amener certaines questions au niveau des variables choisies ou des méthodes d'estimation.

TABLE 2 – Résultats des estimations OLS, WITHIN et GLS

	Variable dépendante : Taux de croissance du PIB/habitant		
	OLS	WITHIN	GLS
log PIB/habitant (t-1)	-0.169*** (0.029)	-0.273*** (0.024)	-0.212*** (0.040)
log Fécondité	-0.073** (0.031)	-0.041 (0.041)	-0.048 (0.031)
1/Espérance de vie à la naissance	19.848 (17.328)	-78.146*** (16.713)	-29.371 (21.637)
Education	0.007* (0.004)	-0.014** (0.007)	-0.0002 (0.006)
Inflation	-0.100** (0.047)	-0.142*** (0.041)	-0.114*** (0.039)
Degré d'ouverture	0.061*** (0.014)	0.031* (0.017)	0.060*** (0.020)
Constante	1.510*** (0.455)		2.660*** (0.585)
Observations	207	207	207
$R^2$	0.416	0.295	0.293
$\hat{\sigma}^2_{\text{Régression}}$	0.0056	0.0027	0.0032
$\hat{\sigma}^2_{\mu}$			0.00111
$\hat{\sigma}^2_{\epsilon}$			0.00074

Note 1 :

\* $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$

Note 2 :

Les chiffres entre parenthèses représentent les écarts-types estimés des coefficients

## 2.2 Validité statistique du modèle

### 2.2.1 Tests d'existence des effets spécifiques

L'objectif de cette sous-section est de tester l'existence de l'effet individuel  $\mu_i$ , de l'effet temporel  $\epsilon_t$  ou alors des deux effets. Pour tester cela, cela revient à tester la nullité de la variance des effets, i.e.  $H_0 : \sigma_{\mu}^2 = 0$ ,  $H_0 : \sigma_{\epsilon}^2 = 0$  contre  $H_A : \sigma_{\mu}^2 > 0$  ou  $H_A : \sigma_{\epsilon}^2 > 0$

Comme nous pouvons le constater avec les Tables 3 et 4, le test de Fisher et celui du multiplicateur de Lagrange confirment l'existence des deux effets. En effet, pour les deux tests, il existe des effets individuels, temporels et combinés au seuil de 1%. Ces tests confirment alors la démarche entreprise dans la section précédente selon laquelle les modèles Within et GLS sont estimés avec les deux effets. Par ailleurs, ces conclusions nous permettent de rejeter le modèle OLS qui ne prend en compte aucun effet.

TABLE 3 – F-test pour l'existence des effets

	F-stat	F-théorique	df1	df2	p-value	Décision
Twoways	5.8856***	1.526	30	170	$1.814 \times 10^{-14}$	<b>Effet significatif</b>
Indiv	4.0513***	1.602	22	178	$6.92 \times 10^{-8}$	<b>Effet significatif</b>
Time	5.9094***	1.987	8	192	$9.141 \times 10^{-7}$	<b>Effet significatif</b>

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

TABLE 4 – LM test pour l’existence des effets

	Chi-sq	$\chi_k^2$	df	p-value	Décision
Twoways	64.895***	5.991	2	$8.095 \times 10^{-15}$	<b>Effet significatif</b>
Indiv	14.077***	3.841	1	0.0001755	<b>Effet significatif</b>
Time	50.818***	3.841	1	$1.013 \times 10^{-12}$	<b>Effet significatif</b>

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

### 2.2.2 Test d’indépendance d’Hausman

L’objectif de cette sous-section est de savoir si les effets sont aléatoires ou fixes, ce qui va nous permettre de trancher entre le modèle Within et le modèle GLS. Sous  $H_0$ , l’estimateur est convergent et efficace contre  $H_A$  où l’estimateur est non convergent. Si l’on rejette l’hypothèse nulle, on pourra considérer les effets comme fixes et le modèle Within à la place du modèle GLS. La statistique de test suit une loi de khi-2 avec k degrés de libertés. Lorsque la statistique de test est significativement supérieure à la valeur de référence, on peut conclure à l’existence d’effets fixes.

TABLE 5 – Test d’Hausman

	Chi-sq	$\chi_k^2$	df	p-value
Hausman	32.455***	12.592	6	$1.334 \times 10^{-5}$

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

Les résultats suggèrent de rejeter l’hypothèse nulle, ce qui revient à considérer les effets comme fixes et donc de considérer le modèle Within plutôt que le modèle GLS. En effet, la p-value est inférieure au seuil  $\alpha = 1\%$ . Cela justifie le choix effectué dans la section 2.1 dans laquelle nous n’avons interprété que le modèle Within et non pas le modèle des MCO ou le modèle GLS.

## Conclusion

Ainsi, cette étude nous a permis de confirmer le phénomène de convergence au sein des 23 pays de l’OCDE sur la période 1970-2015. Cela revient à dire que plus le niveau de PIB/habitant initial est élevé, plus le taux de croissance du PIB/habitant sera faible. Il convient de noter que cette étude comporte certaines limites. La première est le fait que le modèle présenté souffre d’un biais d’omission. En effet, si l’on considère la régression effectuée par Barro (2016), le modèle présenté ici contient plus de variables que l’étude ici présente. Nous pouvons noter comme variables, les termes à l’échange, la part des investissements dans le PIB ou encore des indicateurs démocratiques. De plus, le modèle présenté ici dispose d’une faible dimension individuelle, cela revient à dire que nous disposons d’un faible échantillon de pays et que dans des travaux futurs, nous pourrions considérer un plus grand échantillon. Par ailleurs, la dimension temporelle est assez faible également, vu que les données sont réparties sur 5 ans, cela réduit le nombre d’observations temporelles, il pourra être intéressant de considérer dans des travaux futurs une période plus longue.

De futurs travaux pourront considérer l’impact de l’intelligence artificielle sur le phénomène de convergence entre les pays. Prettnner and Bloom (2020) proposent déjà une nouvelle version du modèle de Solow en y incluant un facteur de production représentant des robots industriels.

## Références

- Barro, Robert J. 1991. "Economic growth in a cross section of countries." *The quarterly journal of economics* 106 (2) :407–443.
- . 2016. "Economic growth and convergence, applied to China." *China & World Economy* 24 (5) :5–19.
- Baumol, William J. 1986. "Productivity growth, convergence, and welfare : what the long-run data show." *The american economic review* :1072–1085.
- Prettner, Klaus and David E Bloom. 2020. *Automation and its macroeconomic consequences : Theory, evidence, and social impacts*. Academic Press.
- Solow, Robert M. 1956. "A contribution to the theory of economic growth." *The quarterly journal of economics* 70 (1) :65–94.

## Code R

```
#####  
# Nom : SCHILD                                #  
# Prénom : Nathan                            #  
# Master APE                                #  
# Code dossier                                #  
#####  
  
#-----  
  
#####  
#---- Installation et chargement des packages nécessaires/base de données ----  
  
# Dans un premier temps, il convient d'installer et de charger les packages nécessaires.  
install.packages("lmtest")  
install.packages("plm")  
install.packages("dplyr")  
install.packages("corrplot")  
install.packages("ggplot2")  
install.packages("stargazer")  
install.packages("ggthemes")  
install.packages("fBasics")  
library(plm)  
library(readxl)  
library(dplyr)  
library(corrplot)  
library(ggplot2)
```

```

library(stargazer)
library(car)
library(lmtest)
library(ggthemes)
library(fBasics)
library(readxl)

# On charge la base de données
data <- read_excel("C:/Users/natha/OneDrive/Bureau/M1/EDP/datap.xlsx")
View(data)

# On transforme la base de données .
pdata<-pdata.frame(data,index=c("country","year"),drop.index=T)

#####
#---- Statistiques descriptives----

# On sélectionne tout d'abord les variables dont on souhaite des statistiques descriptives :

stats<- pdata %>% select(growthrate,gdpcap)

# On visualise les statistiques descriptives à l'aide d'un tableau.
stargazer(basicStats(stats) ,
          type = "text",
          title = "Table: Statistiques descriptives",
          summary=F)

# Pour la présentation, on en extrait le code latex.
stargazer(basicStats(stats),
          type = "latex",
          title = "Table: Statistiques descriptives",
          summary=F)

#####
#---- Estimation des modèles ----

#---- Les MCO ----

# Dans un premier temps, nous omettons le fait que nous travaillons avec des données de panel.
# 3 modèles sont considérés.

```



```

MCO <- lm(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr
          ,data=pdata)
summary(MCO)
coeftest(MCO,vcov=vcovHC(MCO,type="HC1"))

#---- Modèle Within----

WITHIN <- plm(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
              model="within",effect="twoways",data=pdata)
summary(WITHIN)
coeftest(WITHIN,vcov=vcovHC(WITHIN,type="HC1"))

#---- Modèle GLS----

GLS <- plm(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
           model="random",effect="twoways",data=pdata)
summary(GLS)
coeftest(GLS,vcov=vcovHC(GLS,type="HC1"))

# On crée une variable stockant les écarts-type robustes de chaque modèle...
rob_se <- list(sqrt(diag(vcovHC(MCO, type = "HC1"))),
               sqrt(diag(vcovHC(WITHIN, type = "HC1"))),
               sqrt(diag(vcovHC(GLS, type = "HC1"))))

stargazer(list(MCO, WITHIN, GLS),title = "Résultat des estimations OLS, WITHIN et GLS",
           column.labels = c("OLS", "WITHIN", "GLS"), align = T, type = "text",
           no.space = F,se=rob_se)

stargazer(list(MCO, WITHIN, GLS),title = "Résultat des estimations OLS, WITHIN et GLS",
           column.labels = c("OLS", "WITHIN", "GLS"), align = T, type = "latex",
           no.space = F,se=rob_se)

#*****
#---- Tests d'existence et test d'Hausman----

#---- F-test----

F1<- pFtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
            effect="twoways",data=pdata)
F2<- pFtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
            effect="indiv",data=pdata)
F3<- pFtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,

```

```

        effect="time",data=pdata)
F1 ; F2 ; F3 # -> modèle à deux effets.

qf(0.95,30,170)
qf(0.95,22,178)
qf(0.95,8,192)

#---- LM test----

L1 <- plmtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
              effect="twoways",data=pdata,type="bp")
L2 <- plmtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
              effect="indiv",data=pdata,type="bp")
L3 <- plmtest(growthrate~log_gdp+log_fertility+inv_expec+educ+inf+openr,
              effect="time",data=pdata,type="bp")
L1 ; L2 ; L3 # -> modèle à deux effets -> OLS rejeté...

qchisq(0.95,2)
qchisq(0.95,1)
qchisq(0.95,1)

#---- Test d'Hausman----

phtest(WITHIN, GLS) # Effets fixes => Within avec deux effets est retenu au final
qchisq(0.95,6)

```