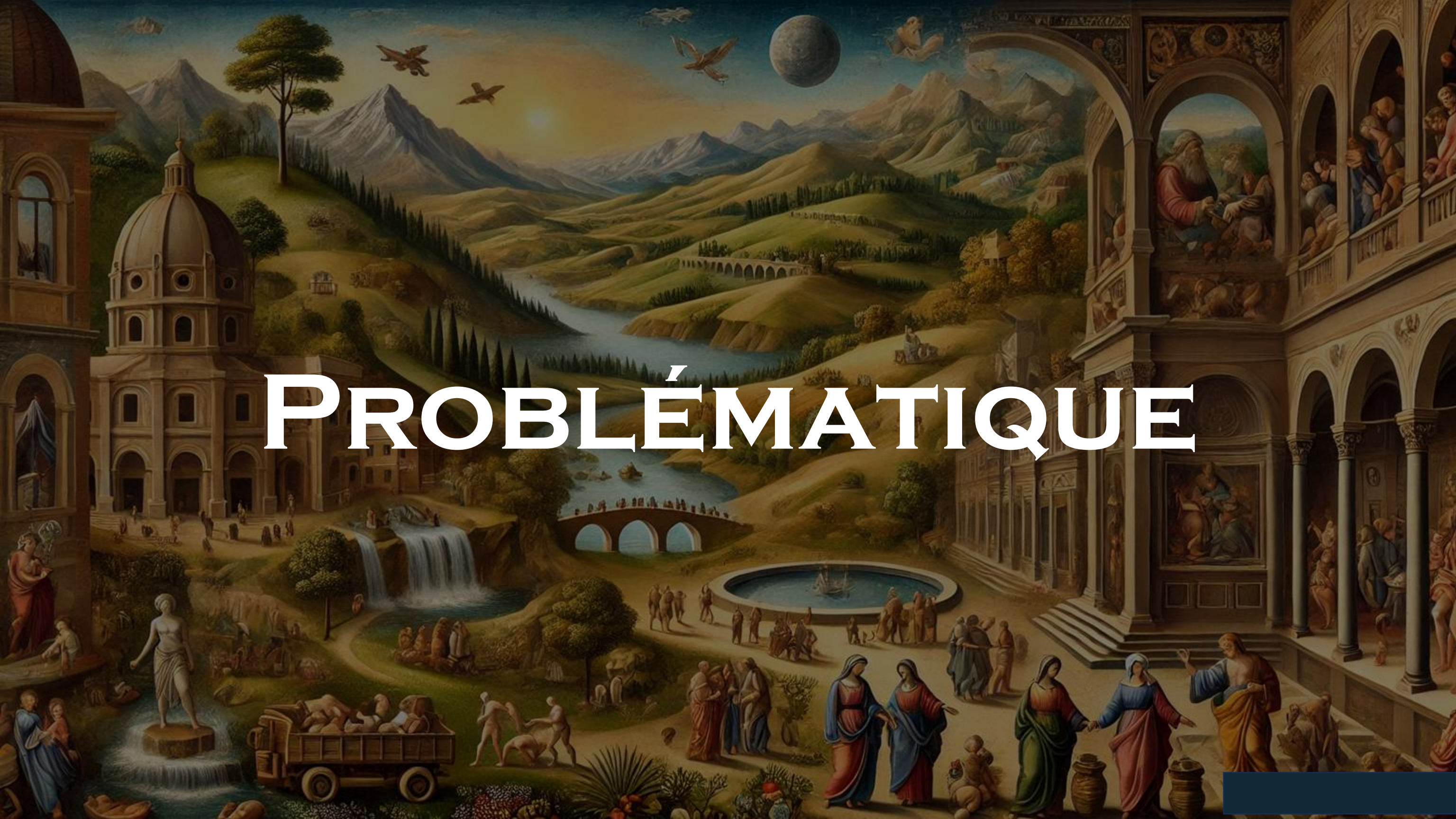


Estimation de l'espérance de vie des pays en développement - Données de Panel



PROBLÉMATIQUE



Introduction

1. Présentation des données

1.1 Description des données brutes (nombre d'observations, périodes, groupes).

1.2 Traitement des données

2. Tests de spécification du modèle

2.1. Test de l'autocorrélation (Hypothèse FE.6)

2.2. Test de l'effet fixe vs aléatoire (Hypothèse FE.3)

3. Estimation finale du modèle

3.1 Estimation du modèle à effets fixes avec erreurs robustes (corrigées pour le clustering) :

3.2 Résultats principaux

4. Discussion des résultats

4.1 Analyse critique des coefficients obtenus pour healthexpenditure et sanitation.

4.2 Comparaison avec les résultats de la littérature existante.

4.3 Limites du modèle

Conclusion

Méthodologie - Données

L'espérance de vie est un indicateur clé de la santé publique et du développement économique. Comprendre les déterminants principaux qui influencent cet indicateur est essentiel pour orienter les politiques publiques.

Identifier les facteurs ayant un impact significatif sur l'espérance de vie à travers une analyse de données de panel.

Évaluer les effets des dépenses de santé, de l'assainissement, du chômage et des émissions de CO2 sur l'espérance de vie.

Analyse des données de panel pour capturer les variations inter et intra-groupes.

Traitement des données pour supprimer les variables avec trop de valeurs manquantes

Hypothèses testées pour justifier le choix du modèle (autocorrélation, variabilité temporelle, effets fixes ou aléatoires).

1.1 Description des données brutes

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
countryname	0				
countrycode	0				
region	0				
incomegroup	0				
year	1,083	2010	5.479756	2001	2019
lifeexpect~k	988	78.31561	3.351955	66.70927	84.35634
prevelance~t	817	3.33011	2.397558	2.5	24.5
co2	950	239832.4	759874.2	40	5753490
healthexpe~e	988	7.426201	2.811096	1.525117	24.23068
educatione~e	742	4.914143	1.513369	1.02195	12.0046
unemployment	912	7.007714	4.145033	.1	27.47
corruption	0				
sanitation	873	80.746	16.896	21.78866	100
id_country	1,083	83.31579	47.78837	4	167
gr_1	1,083	1	0	1	1
gr_2	1,083	0	0	0	0
gr_3	1,083	0	0	0	0
gr_4	1,083	0	0	0	0

```
.
```

1.2 Traitement des données

```
. duplicates report id_country year
```

Duplicates in terms of `id_country year`

copies	observations	surplus
1	1083	0

On assure que chaque pays et chaque année sont uniques

```
misstable summarize
```

Variable	Obs=.	Obs>.	Obs<.	Obs<.		
				Unique values	Min	Max
lifeexpect~k	95		988	>500	66.70927	84.35634
prevelance~t	266		817	81	2.5	24.5
co2	133		950	>500	40	5753490
healthexpe~e	95		988	>500	1.525117	24.23068
educatione~e	341		742	>500	1.02195	12.0046
unemployment	171		912	>500	.1	27.47
corruption	1,083		0	0	.	.
sanitation	210		873	>500	21.78866	100

On vérifie si certaines variables critiques ont des valeurs manquantes :

1.2 Traitement des données

- La variable corruption doit être retirée car elle n'a aucune donnée (100% manquantes).
- Variable educationexpenditure, avec 31% de données manquantes, on décide de retirer la variable si elle n'est pas essentielle à l'analyse.
- Les autres variables ont des données manquantes, mais pas au point de rendre l'analyse impossible.
- La variable co2 a été transformée en log pour le modèle afin de convenir à l'échelle de la variable dépendante

Le fait qu'il n'y ait pas de doublons garantit que le modèle de panel peut être appliqué sans problème structurel.

2. Tests de spécification du modèle

1

Tester la variabilité des variables dans le temps (Hypothèse FE.4)

`xtsum lifeexpectancyworldbank healthexpenditure unemployment sanitation co2`

2

Tester l'autocorrélation (Hypothèse FE.6)

`xtserial lifeexpectancyworldbank healthexpenditure unemployment
sanitation co2`

Tests des hypotheses et creation du modèle

1) Tester la variabilité des variables dans le temps (Hypothèse FE.4)

. xtsum lifeexpectancyworldbank healthexpenditure unemployment sanitation co2						
Variable		Mean	Std. Dev.	Min	Max	Observations
lifeex~k	overall	78.31561	3.351955	66.70927	84.35634	N = 988
	between		3.13633	70.06307	82.89155	n = 52
	within		1.256356	73.6565	82.0443	T = 19
health~e	overall	7.426201	2.811096	1.525117	24.23068	N = 988
	between		2.670482	1.962694	15.60953	n = 52
	within		.949128	1.859993	19.51751	T = 19
unempl~t	overall	7.007714	4.145033	.1	27.47	N = 912
	between		3.345694	.4698947	15.98105	n = 48
	within		2.491746	-1.044391	18.66561	T = 19
sanita~n	overall	80.746	16.896	21.78866	100	N = 873
	between		15.66743	32.51863	100	n = 46
	within		6.693634	27.01731	105.2286	T-bar = 18.9783
co2	overall	239832.4	759874.2	40	5753490	N = 950
	between		765199.4	56.31579	5306051	n = 50
	within		54631.61	-248498	687272.1	T = 19

Tests des hypotheses et creation du modèle

2) Tester l'autocorrélation (Hypothèse FE.6)

```
. xtserial lifeexpectancyworldbank healthexpenditure unemployment sanitation co2

Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation
    F( 1,      40) =      81.365
      Prob > F =      0.0000
```

Le test de Wooldridge pour l'autocorrélation dans les données de panel donne les résultats suivants :

$F(1, 40) = 81.365$

$\text{Prob} > F = 0.0000$

- L'hypothèse nulle (H_0) de ce test est : Il n'y a pas d'autocorrélation d'ordre 1 dans les erreurs du modèle. Comme la p-value est très faible (0.0000), on rejette l'hypothèse nulle.

Tests des hypotheses et creation du modèle

3

Comparer effets fixes vs aléatoires (Hypothèse FE.3)

```
xtreg lifeexpectancyworldbank healthexpenditure unemployment sanitation co2, fe  
estimates store fixed
```

```
xtreg lifeexpectancyworldbank healthexpenditure unemployment sanitation co2, re  
estimates store random  
hausman fixed random
```

4

Effectuer la régression à effets fixes

```
xtreg lifeexpectancyworldbank healthexpenditure unemployment sanitation co2_log, fe  
vce(cluster id_country)
```

```
xtregar lifeexpectancyworldbank healthexpenditure unemployment sanitation co2_log, fe
```


Comparer effets fixes vs aléatoires (Hypothèse FE.3)

```
. hausman fixed random
```

	—— Coefficients ——			
	(b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
healthexpe~e	.6700706	.6606824	.0093883	.0112919
unemployment	.0017628	.0004474	.0013154	.
sanitation	.1235506	.1198649	.0036856	.0009875

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 12.31
Prob>chi2 = 0.0064
(V_b-V_B is not positive definite)

```
.
```

Le modèle à effets fixes (FE) est plus adapté dans ce cas, car il tient compte de la corrélation entre les effets non observés et les variables explicatives.

```
. xtregar lifeexpectancyworldbank healthexpenditure unemployment sanitation co2_log, fe
```

```
FE (within) regression with AR(1) disturbances   Number of obs   =       737
```

```
Group variable: id_country                     Number of groups =       41
```

```
R-sq:                                           Obs per group:
```

```
    within = 0.0705                             min =       17
```

```
    between = 0.0079                             avg =      18.0
```

```
    overall = 0.0013                             max =       18
```

```
F(4, 692) =      13.13
```

```
corr(u_i, Xb) = -0.4862                       Prob > F      =      0.0000
```

lifeexpectancyw~k	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
healthexpenditure	-.0432319	.027533	-1.57	0.117	-.0972901	.0108263
unemployment	.0313621	.0075465	4.16	0.000	.0165454	.0461788
sanitation	.0380117	.0158144	2.40	0.016	.0069618	.0690617
co2_log	-.7977638	.1694973	-4.71	0.000	-1.130554	-.4649732
_cons	88.70605	.1247607	711.01	0.000	88.4611	88.95101
rho_ar	.95089887					
sigma_u	3.2530696					
sigma_e	.23787058					
rho_fov	.99468163	(fraction of variance because of u_i)				

```
F test that all u_i=0: F(40, 692) = 6.51                      Prob > F = 0.0000
```

xtregar : Corrige pour une autocorrélation AR(1) dans les modèles à effets fixes.

```
. xtreg lifeexpectancyworldbank healthexpenditure unemployment sanitation co2_log, fe vce(cluster id_country)
> untry)
```

```
Fixed-effects (within) regression
Group variable: id_country

Number of obs      =      778
Number of groups   =      41

R-sq:
    within  = 0.6101
    between = 0.0386
    overall  = 0.0654

Obs per group:
    min = 18
    avg  = 19.0
    max  = 19

F(4, 40)      = 66.70
Prob > F      = 0.0000

corr(u_i, Xb) = -0.7309
```

(Std. Err. adjusted for 41 clusters in id_country)

lifeexpectancyw~k	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
healthexpenditure	.6448778	.1317048	4.90	0.000	.3786924	.9110632
unemployment	-.0261857	.0276207	-0.95	0.349	-.0820093	.0296379
sanitation	.1278966	.0121108	10.56	0.000	.1034198	.1523735
co2_log	-2.345323	.5040592	-4.65	0.000	-3.364065	-1.326582
_cons	89.85377	5.485468	16.38	0.000	78.76722	100.9403
sigma_u	4.2584719					
sigma_e	.82379854					
rho	.96392729	(fraction of variance due to u_i)				

vce(cluster id_country) : Ajuste les erreurs standards pour tenir compte de l'autocorrélation intra-groupe.

Analyse des modèles

Modèle 1 : xtregar

Ce modèle est utile si l'on suspecte une forte autocorrélation AR(1) dans les résidus.

Cependant, l'autocorrélation résiduelle n'est pas explicitement signalée comme corrigée ici.

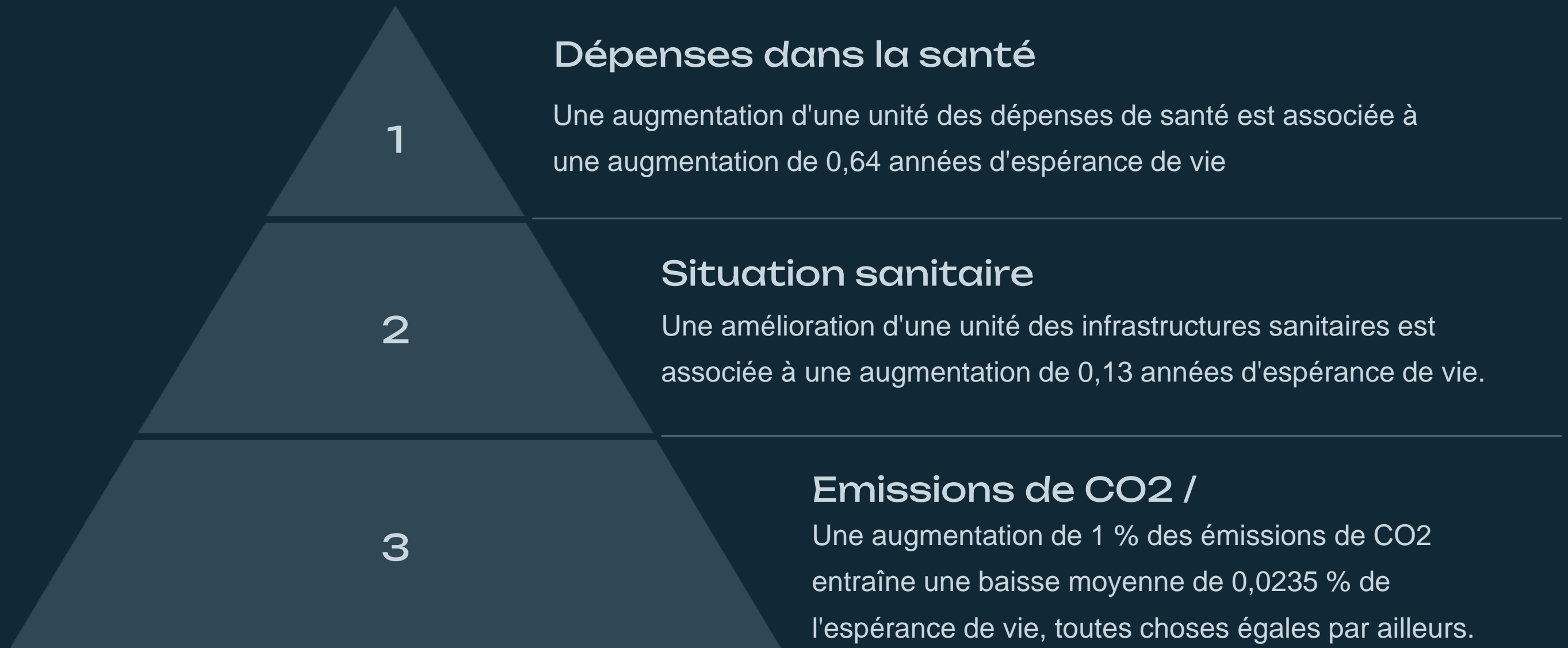
- Faible R^2 , ce qui indique que les variables explicatives expliquent peu de la variance intra-groupe.
- Moins robuste aux biais potentiels qu'un modèle avec erreurs robustes (cluster).

Modèle 2 : xtreg avec erreurs robustes (cluster)

Ce modèle est plus robuste, car il ajuste les erreurs standard pour tenir compte de l'autocorrélation et de l'hétéroscédasticité au niveau des groupes.

- R^2 plus élevé, indiquant une meilleure capacité explicative.
- L'effet de healthexpenditure devient significatif, ce qui est intuitivement attendu.
- Estimations robustes aux biais liés à l'hétéroscédasticité et à la corrélation intra-groupe, ce qui est souvent le cas dans les données de panel.

4. Discussion des résultats



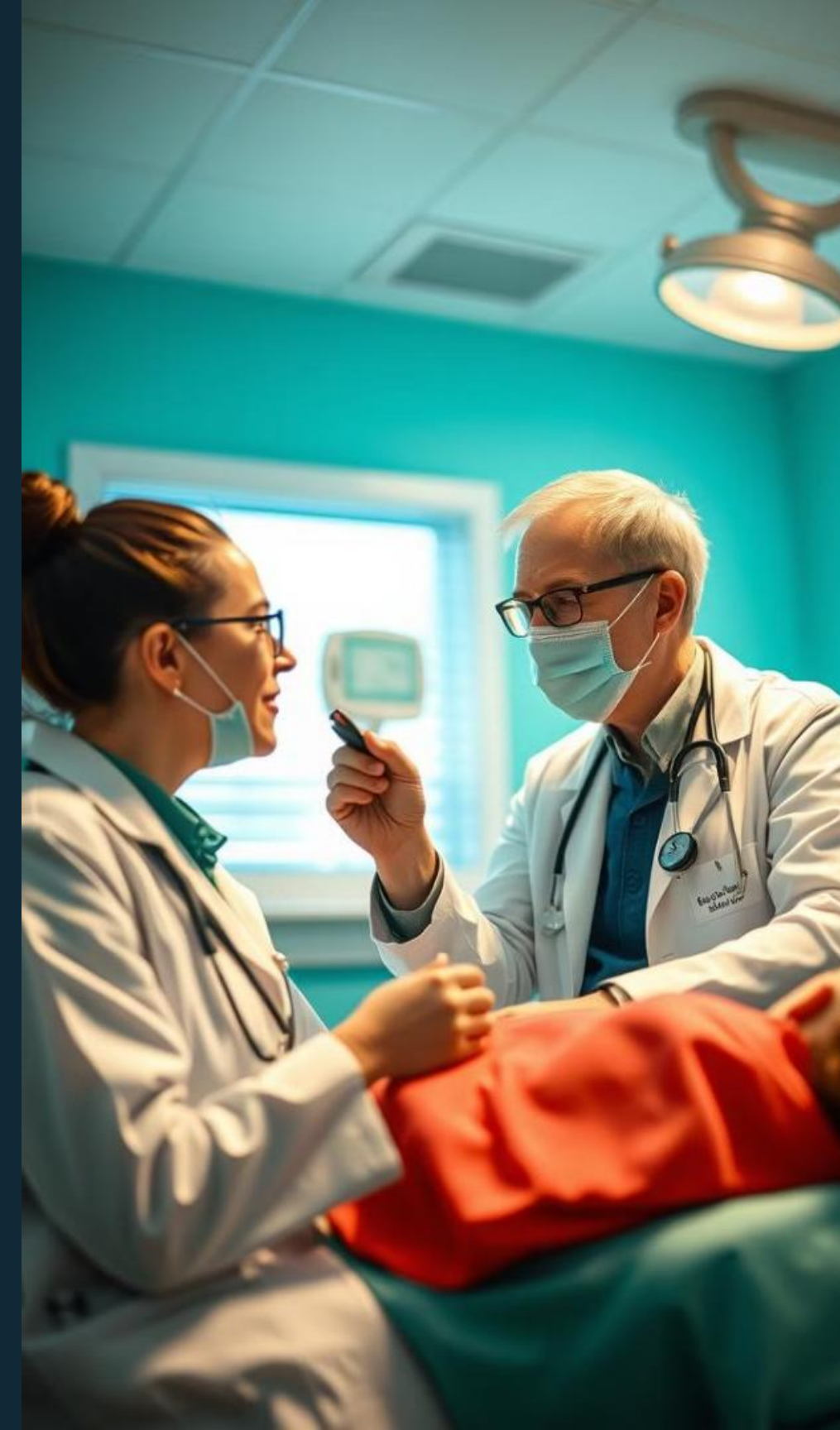
4. Discussion des résultats

Cet effet est attendu, car des investissements accrus dans la santé améliorent l'accès aux soins, la prévention, et les traitements, ce qui se traduit par une augmentation de la longévité.

Bien que le chômage puisse être lié à des conditions de vie moins favorables (stress, pauvreté, etc.), les données ici ne montrent pas d'effet clair.

Pour la situation sanitaire cela est cohérent, car de meilleures infrastructures sanitaires réduisent la propagation des maladies, améliorent les conditions de vie et favorisent une meilleure santé générale.

Le résultat pour les émissions de CO_2 du Pays met en évidence l'impact négatif des émissions de CO_2 sur la santé publique, probablement lié à la pollution de l'air et à d'autres facteurs environnementaux.



Situation sanitaire

Relation entre la pollution de l'air et l'espérance de vie :

Une étude de l'Université de l'Arkansas pour les Sciences Médicales a montré qu'une diminution de $10 \mu\text{g}/\text{m}^3$ des particules fines est associée à une augmentation de l'espérance de vie, particulièrement dans les zones urbaines denses.

Impact des particules fines sur la santé :

Les particules fines (PM 2,5) représentent un risque majeur pour la santé humaine. Une étude de l'Institut de politique énergétique de l'université de Chicago indique que respecter les seuils recommandés par l'OMS pourrait augmenter l'espérance de vie moyenne de 2,3 ans par habitant.





Sources

Estimations de la relation entre la pollution de l'air et l'espérance de vie ou la mortalité

Période : décembre 2012 à mars 2013

Ilias KAVOURAS Université de l'Arkansas pour les Sciences Médicales – Santé Environnementale et Professionnelle – Little Rock – États-Unis

https://www.dna.fr/magazine-sante/2023/09/15/combien-d-annees-d-esperance-de-vie-nous-fait-perdre-la-pollution?utm_source=chatgpt.com

