

---

**Les investissements en équipements médicaux réduisent-ils la durée  
moyenne des séjours à l'hôpital ?**  
Une étude sur données de panel

**Économétrie des données de panel**

Enseignant : Mr. EL OUARDIGHI

Réalisé par  
Nawel ARIF

## 1. Introduction

Dans un monde où l'accès aux soins de santé est un pilier essentiel du bien-être et du progrès social, l'efficacité des infrastructures hospitalières reste une préoccupation centrale.

Les investissements en équipements médicaux, tels que les unités IRM et les scanners CT, soulèvent une question cruciale : **permettent-ils réellement de réduire la durée des séjours à l'hôpital, allégeant ainsi le fardeau des systèmes de santé tout en améliorant la qualité des soins pour les patients ?**

Nous pouvons dès à présent poser les hypothèses pour comparer les résultats de la littérature en lien avec ce sujet et nos résultats empiriques personnels. Ces hypothèses sont les suivantes.

- Une augmentation des équipements médicaux tels que les unités IRM ou les scanners CT est associée à une réduction de la durée moyenne des séjours.
- Une augmentation des lits d'hôpitaux peut soit augmenter, soit réduire la durée moyenne des séjours en fonction de l'offre et de la gestion hospitalière.
- Les effets spécifiques temporels et régionaux (pays) jouent un rôle important dans l'interprétation des résultats.

Ces hypothèses découlent directement des principes de gestion hospitalière et d'économie de la santé. D'un côté, une meilleure disponibilité des équipements médicaux devrait réduire le temps d'attente et accélérer les diagnostics, entraînant ainsi une diminution de la durée des hospitalisations.

D'un autre côté, l'expansion des infrastructures hospitalières, notamment l'augmentation des lits disponibles, pourrait produire des effets contrastés : un effet de réduction des files d'attente favorisant des séjours plus courts, ou au contraire, une hausse de la demande et une utilisation prolongée des ressources.

De manière générale, ces hypothèses sont partiellement vérifiées dans la littérature en économie de la santé, bien que les résultats puissent varier en fonction des spécificités locales et des contraintes budgétaires. Il est donc essentiel d'examiner empiriquement l'impact des équipements médicaux sur la durée moyenne des séjours hospitaliers en tenant compte des différences structurelles entre établissements et des effets temporels.

Pour répondre à cette problématique, nous mobiliserons **un modèle de panel à effets fixes bidirectionnels** afin de contrôler les spécificités propres aux hôpitaux et aux périodes étudiées.

$$Hospital\_Stay_{it} = \beta_0 + \beta_1 MRI\_Units_{it} + \beta_2 CT\_Scanners_{it} + \beta_3 Hospital\_Beds_{it} + \mu_i + \epsilon_t + \nu_{it}$$

Où :

- $\mu_i$  représente un effet invariant dans le temps, capturant les caractéristiques spécifiques à chaque hôpital.
- $\epsilon_t$  correspond à un effet invariant entre les établissements mais variant dans le temps, reflétant les tendances globales du secteur hospitalier.
- $\nu_{it}$  est un terme d'erreur aléatoire.

Ce modèle permet de contrôler les effets spécifiques et individuels qui pourraient biaiser notre estimation. En effet, certaines caractéristiques hospitalières, comme la gestion des ressources ou les politiques locales, peuvent influencer la durée des séjours de manière constante sur toute la période étudiée. En tenant compte des effets fixes, nous évitons ainsi que ces biais affectent nos résultats.

Dans un premier temps, nous allons explorer notre base de données et déterminer quelles sont les informations pertinentes que nous pourrions en tirer. Ensuite, nous détaillerons les estimations et leurs résultats, pour enfin conclure et identifier les éventuelles améliorations à apporter à notre analyse.

## 2. Le cadre statistique

Les données utilisées proviennent de la plateforme Kaggle

(<https://www.kaggle.com/datasets/babyoda/healthcare-investments-and-length-of-hospital-stay>). Cette base de données, intitulée « Healthcare Investments and Length of Hospital Stay », regroupe des observations annuelles sur quatre variables pour 32 pays à travers le monde, couvrant la période de 1990 à 2018.

### 2.1 Statistiques Descriptives

Tableau 1 : Statistiques Descriptives

VARIABLES	DEFINITIONS	N	MEAN	STD_DEV	MIN	MEDIAN	MAX
HOSPITAL_STAY	Durée moyenne des séjours hospitaliers (en jours).	518	7,140	2,567	3,400	6,650	32,700
MRI_UNITS	Nombre d'unités d'IRM (Imagerie par Résonance Magnétique) disponibles par million d'habitants	518	10,566	8,686	0,100	8,765	55,210
CT_SCANNERS	Nombre de scanners CT (Tomodensitométrie) disponibles par million d'habitants.	518	19,647	14,353	1,480	15,375	111,490
HOSPITAL_BEDS	Nombre de lits d'hôpital pour 1 000 habitants.	518	2,159	0,802	0,095	2,279	4,029

L'analyse des statistiques descriptives du tableau 1 permet d'identifier les tendances générales et les disparités entre pays. On observe notamment que la durée moyenne des séjours hospitaliers est de 7,14 jours, avec une médiane inférieure (6,65 jours) et un écart-type de 2,57 jours.

Cela indique que la majorité des séjours se concentrent autour de cette moyenne, bien que certaines hospitalisations puissent être nettement plus longues. La valeur minimale est de 3,4 jours, tandis que certaines observations atteignent 32,7 jours, ce qui peut refléter des différences structurelles dans l'organisation des soins, comme au Japon.

Cette distribution asymétrique est également visible sur la figure 1, qui met en évidence un pic autour de la médiane et une traîne plus étendue vers les séjours longs.

Le nombre d'IRM est en moyenne de 10,57 unités par million d'habitants, mais avec un écart-type élevé de 8,69, témoignant d'une forte hétérogénéité entre les systèmes de santé. Certains pays disposent de moins d'une IRM par million d'habitants, alors que d'autres en comptent jusqu'à 55 unités.

De même, les scanners CT sont plus nombreux, avec une moyenne de 19,65 unités par million d'habitants, mais une dispersion encore plus grande (écart-type de 14,35). Ces variations significatives peuvent s'expliquer par les disparités de financement public, les choix d'investissement dans le secteur de la santé et l'accès à des options de diagnostic.

Le nombre de lits d'hôpital présente aussi une variabilité significative, avec une moyenne de 2,16 (en logarithme), soit environ 8,6 lits pour 1000 habitants.

Toutefois, la variation reste notable avec un minimum de 0,095 et un maximum de 4,02, ce qui reflète des stratégies différentes dans la gestion des infrastructures hospitalières. Certains pays privilégient une capacité hospitalière importante pour absorber la demande, tandis que d'autres misent sur des soins ambulatoires pour limiter les hospitalisations prolongées.

D'après la figure 2, une corrélation négative apparaît entre *Hospital Stay* et les équipements médicaux : le nombre d'IRM (-0.06) et de scanners CT (-0.25) semble indiquer qu'une meilleure dotation en matériel pourrait être associée à des séjours plus courts. Cependant, cette tendance n'est pas systématique, comme le montre le cas du Japon, où les hospitalisations restent longues malgré une infrastructure bien développée, en

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



Le modèle (1) est une régression multiple estimée par les moindres carrés ordinaires (*Pooled OLS*). Cette approche permet d'obtenir une estimation globale des coefficients sans tenir compte des hétérogénéités individuelles ou temporelles, ce qui peut entraîner un biais d'omission si ces effets sont significatifs. Par exemple, une augmentation de 1 % du nombre d'IRM est associée à une hausse de 0,183 % de la durée d'hospitalisation, tandis qu'une augmentation de 1 % du nombre de scanners CT entraîne une augmentation de 0,082 % de cette durée. En revanche, une augmentation de 1 % du nombre de lits hospitaliers ( $\log\_Hospital\_Beds$ ) est associée à une réduction de 3,235 % de la durée d'hospitalisation.

Le modèle (2) correspond à la spécification retenue et affinée au fur et à mesure de l'étude. Cette sélection repose sur des critères de pertinence statistique et économique, assurant une meilleure adéquation aux caractéristiques des données analysées. Contrairement au modèle (1), les résultats montrent qu'une augmentation de 1 % du nombre de scanners CT est associée à une baisse de 0,045 % de la durée des séjours hospitaliers. Par ailleurs, une augmentation de 1 % du nombre de lits hospitaliers réduit la durée d'hospitalisation de 0,930 %, tandis que l'effet du nombre d'IRM n'est pas significatif.

Le modèle (3) adopte une structure de régression à effets aléatoires, intégrant les effets individuels et temporels sous l'hypothèse qu'ils sont non corrélés avec les régresseurs. Toutefois, dans le cadre de notre étude, cette hypothèse a été jugée trop restrictive, rendant ce modèle moins approprié. Selon cette spécification, une augmentation de 1 % du nombre de lits hospitaliers est associée à une réduction de 1,263 % de la durée d'hospitalisation, tandis que l'effet des scanners CT et des IRM n'est pas significatif.

Enfin, nous justifierons notre choix de modèle et expliquerons pourquoi les modèles (1) et (3) présentent des limites dans notre contexte spécifique. De plus, l'application d'une transformation logarithmique à la variable 'lits d'hôpital' permet d'atténuer l'influence des valeurs extrêmes, d'améliorer la linéarité de la relation estimée et de faciliter l'interprétation des coefficients en termes d'élasticité.

### 2.3 Validité statistique

### 2.3.1 Pooled OLS

Le modèle de régression en pooled OLS correspond à une régression multiple estimée par la méthode des moindres carrés ordinaires (MCO) appliquée à des données de panel. Cette approche fait abstraction des effets spécifiques individuels et temporels, en imposant l'hypothèse d'homogénéité des coefficients à travers les unités et les périodes. Or, dans une étude sur les structures hospitalières où l'on observe d'importantes disparités entre pays et au fil du temps, cette hypothèse est souvent irréaliste. De surcroît, l'absence de prise en compte des effets fixes propres à chaque hôpital peut induire un biais dans les estimations, remettant en cause la validité statistique des résultats. Pour ces raisons, cette méthode n'a pas été retenue dans notre analyse.

$$Hospital\_Stay_{it} = \beta_0 + \beta_1 MRI\_Units_{it} + \beta_2 CT\_Scanners_{it} + \beta_3 Hospital\_Beds_{it} + \epsilon_{it}$$

### 2.3.2 Effets fixes (WITHIN)

Le modèle à effets fixes capture les effets propres à chaque localisation et période, éliminant les biais liés aux caractéristiques inobservées constantes. Les tests de Lagrange et F-tests (Tableaux 1 à 6 en annexe) confirment la significativité des effets individuels et temporels, justifiant statistiquement leur utilisation.

L'analyse révèle que l'effet des scanners CT devient négatif et significatif, suggérant qu'une augmentation de leur nombre réduit la durée des séjours hospitaliers. De plus, le nombre de lits d'hôpital conserve un effet négatif et significatif, confirmant que l'augmentation de l'offre hospitalière diminue la durée des hospitalisations.

$$Hospital\ Stay_{it} = \beta_0 + \beta_1 MRI\ Units_{it} + \beta_2 CT\ Scanners_{it} + \beta_3 Hospital\ Beds_{it} + \mu_i + \epsilon_t + \nu_{it}$$

### 2. 3. 3 Effets aléatoires (FGLS)

Le modèle à effets aléatoires suppose l'absence de corrélation entre les effets individuels, temporels et les variables explicatives, plusieurs éléments remettent en cause la pertinence de ce modèle.

Dans le **Tableau 2, CT Scanners** devient non significatif sous les effets aléatoires, contrairement aux effets fixes, remettant en cause cette approche car certaines relations attendues semblent disparaître.

Le **test d'Hausman** (Tableau 7),  $\chi^2 = 422.65, p < 2.2e - 16$  compare les effets fixes et aléatoires.

La rejection de l'hypothèse nulle indique une corrélation entre les effets spécifiques et les régresseurs, rendant le modèle à effets aléatoires biaisé et incohérent.

Le **modèle à effets fixes** est préféré, car il contrôle mieux l'hétérogénéité et conserve des coefficients significatifs. Le **Pooled OLS** est rejeté, ne tenant pas compte de ces effets. Ce choix assure des résultats plus robustes et cohérents.

$$Hospital\_Stay_{it} = \beta_0 + \beta_1 MRI\_Units_{it} + \beta_2 CT\_Scanners_{it} + \beta_3 Hospital\_Beds_{it} + \mu_i + \epsilon_t + vit$$

### 3. Conclusion

Cette étude a analysé l'impact des investissements en équipements médicaux sur la durée des séjours hospitaliers à l'aide d'un modèle à effets fixes bidirectionnels. Les résultats montrent que les scanners CT et l'augmentation du nombre de lits réduisent significativement la durée des hospitalisations, tandis que les unités IRM n'ont pas d'effet clair. Ces conclusions soulignent le rôle crucial des infrastructures hospitalières dans l'optimisation des soins.

Le test d'Hausman ( $\chi^2 = 422.65, p < 2.2e - 16$ ) a rejeté l'hypothèse d'absence de corrélation entre les effets spécifiques et les régresseurs, confirmant que le modèle à effets aléatoires est biaisé et incohérent. Les tests de Lagrange et F-tests ont validé l'existence d'effets individuels et temporels, justifiant l'adoption d'un modèle à effets fixes pour une meilleure précision des estimations.

Cependant, certaines limites demeurent. Une possible endogénéité entre les équipements et la durée des séjours pourrait biaiser les résultats, nécessitant des méthodes plus robustes comme les variables instrumentales. De plus, l'absence de certaines variables explicatives (qualité des soins, politiques hospitalières) pourrait limiter l'interprétation des résultats. Enfin, l'hétérogénéité des systèmes de santé entre pays pourrait être mieux prise en compte à travers des modèles interactifs ou des données plus détaillées.

Enfin, cette étude met en évidence que les scanners CT et l'expansion des infrastructures hospitalières contribuent à réduire la durée des hospitalisations. Des recherches futures intégrant davantage de variables et des méthodes économétriques avancées permettraient de renforcer ces conclusions et d'affiner la compréhension des mécanismes sous-jacents.

## Annexe : Tableaux

### 4. F- Test

**Table 1 : F test for individual effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df1	df2	P value	Alternative hypothesis
75.16	31	483	< 2.2e-16	significant effects

**Table 2 : F test for time effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df1	df2	P value	Alternative hypothesis
2.317	28	486	0.0001929 ***	significant effects

**Table 3 : F test for twoways effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df1	df2	P value	Alternative hypothesis
48.58	59	455	< 2.2e-16	significant effects

### 4.1 Lagrange Multiplier Test

**Table 4 : Lagrange Multiplier Test - individual effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df	P value	Alternative hypothesis
1661	1	< 2.2e-16	significant effects

**Table 5 : Lagrange Multiplier Test - time effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df	P value	Alternative hypothesis
0.0066694	1	0.9349	significant effects

**Table 6: Lagrange Multiplier Test - two-ways effects :  $Hospital\_Stay \sim MRI\_Units + CT\_Scanners + log\_Hospital\_Beds$**

Test statistic	df	P value	Alternative hypothesis
1661	2	< 2.2e-16	significant effects

## 4.2 Hausman Test

Table 7: Hospital\_Stay ~ MRI\_Units + CT\_Scanners + log\_Hospital\_Beds

Test statistic	df	P value	Alternative hypothesis
422.65	3	< 2.2e-16	one model is inconsistent

## 5. Annexe : Figures

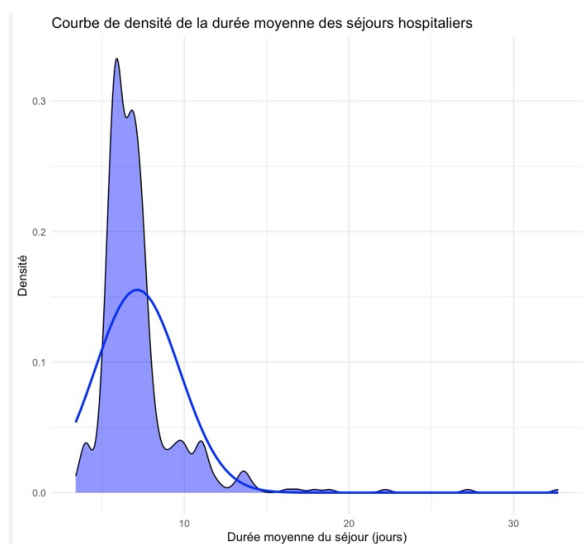


Figure 1 : Courbe de densité de la durée Moyenne des séjours hospitaliers

