

# L'efficacité des EHPAD en France

Brigitte DORMONT, Cécile MARTIN

## Résumé

L'efficacité-coût des EHPAD en France est étudiée en appliquant plusieurs méthodes d'analyse de frontière : estimation sur données en coupe par maximisation de vraisemblance et modèle de données de panel à erreurs composées. La robustesse des résultats est testée par comparaison avec une méthode de régression quantile. Cette étude nous fournit une estimation de l'impact de la qualité de la prise en charge, de la forme institutionnelle et de la taille des établissements sur leurs coûts.

## 1 Introduction

Les dépenses publiques consacrées à la perte d'autonomie des personnes âgées en France s'élèvent à 1.7% du PIB en 2011. L'importance de ces dépenses et l'anticipation de leur hausse invitent actuellement à une réflexion sur un changement global de la régulation de l'offre de soins de long-terme en Etablissement d'Hébergement pour Personnes Agées Dépendantes (EHPAD), i.e. unités de soins de longue durée, maison de retraite ou logement-foyer médicalisés pouvant accueillir des personnes âgées dépendantes.

Une réforme de la régulation tarifaire est en projet. Inspirée de la réforme du financement du secteur hospitalier, elle prévoit le passage d'un système de paiement rétrospectif à une tarification à l'activité fonction du degré de dépendance des résidents accueillis. L'idée sous-jacente d'une telle réforme est qu'il existerait une hétérogénéité des coûts non expliquée entre les établissements qu'il conviendrait de résorber. On observe en effet une grande variabilité des coûts par résident selon le statut de l'institution (public, privé à but lucratif, privé à but non lucratif) et selon sa localisation géographique. Certaines associations d'EHPAD s'opposent néanmoins à ce projet en soutenant qu'il pourrait engendrer un nivellement vers le bas de la qualité du service et une sélection des résidents si leurs caractéristiques ne sont pas prises correctement en compte dans le calcul du tarif remboursé. Afin d'en évaluer les conséquences, on estime par des méthodes d'analyse paramétriques de frontière la part de l'hétérogénéité des coûts expliquée par les caractéristiques de la patientèle (âge, degré de dépendance, pathologies spécifiques), par la qualité de la prise en charge et celle résultant d'inefficacités techniques et allocatives des établissements. Les scores d'efficacité obtenus sont étudiés afin d'examiner si le statut institutionnel influe sur l'efficacité.

Une autre évolution récente de la régulation du secteur réside dans l'encouragement donné à la constitution d'établissements de grande taille. Des autorisations d'ouverture plus strictes et des incitations au regroupement d'EHPAD<sup>1</sup> concrétisent cette quête de la taille critique. La mise en commun de personnel, un plus grand pouvoir de négociation vis à vis des institutions financières ainsi que le partage de tâches administratives permettraient aux établissements de profiter d'économies d'échelle substantielles. Nous vérifions cette hypothèse en évaluant par divers modèles d'estimation la taille critique des institutions. Nous testons la robustesse du résultat pour toute catégorie de résidents accueillis, selon le degré de dépendance et le mode d'hébergement choisi (hébergement permanent, accueil de jour ou hébergement temporaire).

## 2 Contexte des soins de long terme en institution en France

### 2.1 L'objet d'étude

Différents types d'établissements peuvent accueillir des personnes âgées en France :

- les EHPAD qui délivrent des soins à des résidents en perte d'autonomie ;
- les Etablissement d'hébergement pour personnes âgées (EHPA) non EHPAD qui ne sont pas médicalisés et

---

1. La loi *HPST* du 21 juillet 2009 introduit ainsi l'obligation d'appels à projet systématiques pour toute autorisation de création ou de transformation des établissements médico-sociaux, avec pour seule exception à cette obligation le cas d'opérations de regroupements entre établissements préexistants. Ceci est perçu par l'UNIOSS et par d'autres acteurs du secteur comme une évidente "incitation au regroupement".[32].

accueillent des personnes âgées non dépendantes.

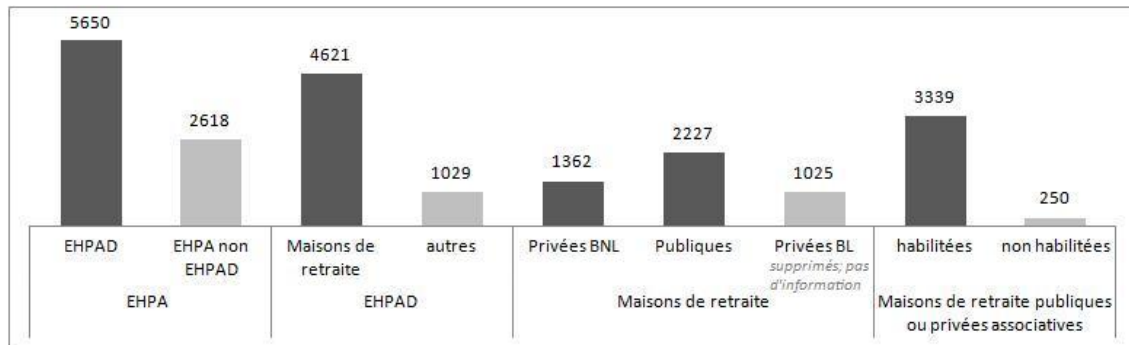
Nous nous intéressons ici uniquement aux EHPAD. Cette catégorie peut recouvrir différentes formes d'institutions : maisons de retraite, logement-foyers, unités de soins de longue durée (USLD), mais aussi certains services de centres hospitaliers, des centres d'accueil de jour, d'hébergement temporaire et des établissements expérimentaux. Afin de conserver un groupe homogène d'établissements que nous pourrions comparer entre eux, nous ne conservons dans cette étude que les maisons de retraite.

La moitié environ de ces maisons de retraite sont publiques, et un tiers d'entre elles appartiennent au secteur privé à but non lucratif (caisses de retraite, mutuelles, fondations). En raison d'une non-disponibilité de données, nous ne prenons pas en compte dans cette analyse les EHPAD privés à but lucratif.

Parmi les EHPAD la plupart sont régulés, i.e. ont leurs tarifs fixés par le Conseil Général et l'Autorité Régionale de Santé (ARS). Ils sont en contrepartie habilités à l'aide sociale, ce qui signifie qu'ils peuvent accueillir des résidents bénéficiaires de l'aide sociale départementale. Certains EHPAD ne sont à l'inverse pas habilités à recevoir cette clientèle et peuvent fixer librement une partie de leurs prix. Cela concerne principalement mais non exclusivement des établissements privés à but lucratif. Nous restreignons là encore l'objet de notre analyse en étudiant uniquement les maisons de retraite pour lesquelles les tarifs sont imposés par les autorités de régulation.

In fine, l'objet de notre étude concerne 40.4% des établissements pour personnes âgées. Nous présentons à la Figure 1 le nombre d'établissements concernés par chacune des restrictions effectuées.

FIGURE 1 – Restriction du champ d'étude



Source : EHPA 2007

## 2.2 Le projet de réforme de la tarification

Le système de tarification actuel des EHPAD repose sur un mécanisme de type *prix de journée*. La tarification est ternaire, 3 tarifs correspondant aux 3 missions des EHPAD (hébergement, dépendance et soins) sont fixés. Les charges d'exploitation et d'investissement sont alors réparties entre ces 3 catégories selon des clés de répartition légalement imposées (art. R.314 CASF).

Ces tarifs ne sont pas tous à la charge du résident. L'assurance maladie verse à l'établissement le montant lié au tarif soin, les résidents n'ont à payer que le tarif hébergement et le tarif dépendance. Ils reçoivent en outre l'Allocation Personnalisée pour l'Autonomie (APA) qui couvre une partie du prix dépendance, et peuvent recevoir sous condition de ressources l'aide sociale départementale<sup>2</sup> afin de réduire leur reste à charge lié à l'hébergement.

Les Conseils Généraux et les ARS fixent 3 tarifs dépendance et 3 tarifs soins par établissement, correspondant à 3 degrés différents de dépendance des résidents<sup>3</sup>. Lorsque l'établissement est habilité à l'aide sociale, le tarif hébergement est également fixé par le Conseil Général

Les inconvénients d'un tel système de régulation tarifaire sont l'absence d'incitation à la réduction des coûts et une éventuelle inéquité historique des budgets alloués entre établissements. Lors de la négociation quinquennale<sup>4</sup> au temps  $t=1$  entre un établissement  $e$  et le Conseil Général de son département de résidence, les tarifs

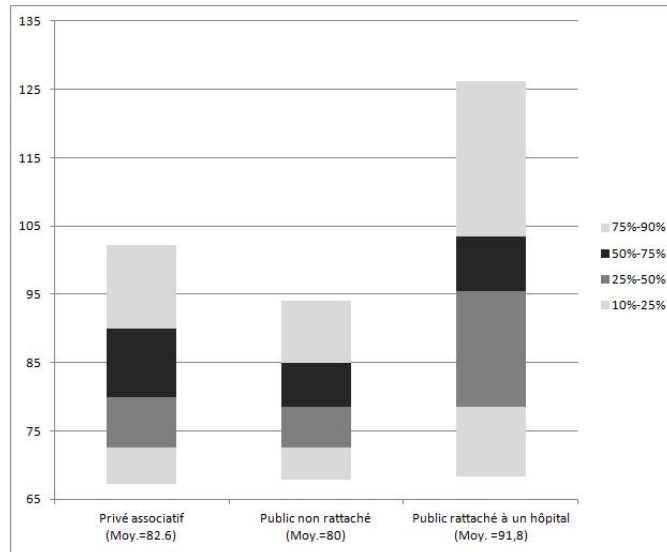
2. Uniquement si l'établissement est habilité à recevoir des bénéficiaires de l'aide sociale.

3. La notion de *dépendance* et sa mesure sont définies par une grille nationale *AGGIR*, permettant de classer les individus dans 6 groupes iso-ressources (GIR) selon les activités qu'ils parviennent ou non à faire seuls, des individus les plus dépendants (GIR 1) aux moins dépendants (GIR 6).

4. Les EHPAD ont, depuis 1999, une obligation de conventionnement avec le Conseil Général et les services déconcentrés de l'Etat. Ces *conventions pluriannuelles tripartites* doivent être signées tous les 5 ans, et permettent de fixer le tarif dépendance,

hébergement, dépendance et soins (i.e. le budget  $B_{t=1}^e$ ) sont en pratique déterminés par les tarifs antérieurs de cet EHPAD (son budget  $B_{t=0}^e$ ), et par un taux  $\tau^e$  que l'on peut caractériser comme étant le poids relatif de l'établissement par rapport aux autorités de tutelle dans cette négociation. Le budget de l'établissement  $e$  serait ainsi partiellement reconduit tous les 5 ans :  $B_{t=1}^e = B_{t=0}^e * (\tau^e + 1)$ . Ceci peut probablement expliquer une partie de la forte hétérogénéité du coût moyen par résident entre les établissements (Figure 2), et notamment selon leur forme institutionnelle. En raison de différences en termes de coûts et de qualité qui semblent exister entre les établissements publics rattachés à une structure hospitalière et ceux non rattachés selon un rapport récent de l'IGAS[4], nous distinguons dans cette étude ces 2 types de structures. Le coût moyen est ainsi plus variable et globalement plus élevé pour les établissements publics rattachés à un hôpital. Ce coût journalier est quasiment identique entre les 2 premiers quartiles des établissements privés associatifs et ceux des maisons de retraite publiques non rattachées, il est en revanche plus élevé sur les 2 derniers quartiles pour les maisons de retraite associatives.

FIGURE 2 – Coût journalier moyen par résident en maison de retraite (en Euros)



Source : EHPA 2007 (Calcul sur 3339 maisons de retraite EHPAD)

Afin de réduire ces hétérogénéités de coûts et d'inciter à une plus grande efficacité, le projet de réforme tarifaire<sup>5</sup> propose de passer d'une tarification rétrospective à une tarification à la ressource, qui serait fonction du degré de dépendance des résidents<sup>6</sup> et de leurs pathologies<sup>7</sup>. Les risques néanmoins connus de la mise en place d'un tel système - réduction de la qualité de la prise en charge, sélection à l'entrée des résidents - freinent l'application de cette réforme. Il existe en outre ici un risque supplémentaire à l'application de cette réforme, puisqu'il n'est prévu une tarification à la ressource que pour les tarifs dépendance et soins et non pour le tarif hébergement qui resterait un prix de journée. On peut donc craindre un transfert des charges des sections dépendance et soins sur la section hébergement. Ce phénomène est d'autant plus à redouter qu'il est prévu dans cette même réforme une double tarification pour l'hébergement des établissements habilités à l'aide sociale : un tarif qui resterait fixé par le Conseil Général pour les bénéficiaires de l'aide sociale et un tarif libre pour les non-bénéficiaires. Ces derniers pourraient ainsi se voir contraints de payer une partie des charges dépendance et soins de l'ensemble des résidents, si les tarifs administrés sont inférieurs à ce qu'ils devraient être compte-tenu des besoins spécifiques de l'EHPAD. Ceci pourrait poser des difficultés financières d'accès de certaines populations à l'institutionnalisation puisque le tarif hébergement constitue la plus grande

soin et, lorsque l'établissement est habilité à l'aide sociale, le tarif hébergement.

5. Projet de décret relatif à la tarification des établissements hébergeant des personnes âgées dépendantes en application du nouvel article L.314-2 du CASF, 23 juin 2009.

6. Calcul d'un GIR moyen pondéré (GMP) par établissement : un nombre de *points GIR* est attribué administrativement à chaque catégorie GIR ; le GMP correspond à la somme des points GIR correspondant à l'ensemble des résidents de l'établissement, rapportée au nombre de résidents.

7. Outil PATHOS : calcul d'un Pathos Moyen Pondéré (PMP) par établissement, afin d'intégrer d'éventuels surcoûts liés à certaines pathologies non prises en compte dans le calcul du GMP.

part du reste à charge des résidents non bénéficiaires de l'aide sociale.

L'objectif de notre étude est de fournir une analyse des conséquences éventuelles d'une telle réforme tarifaire, en étudiant grâce à diverses méthodes de frontière les différences de coûts entre les EHPAD et leurs causes. Nous résumons à la Table 1 le mécanisme de tarification actuel et celui proposé par le projet de réforme, pour chacune des sections tarifaires.

TABLE 1 – Récapitulatif : système de financement actuel des EHPAD et projet de réforme

	<b>Financement actuel</b>	<b>Réforme proposée</b>	<b>Aide éventuelle</b>
<b>Tarif hébergement</b>	Prix de journée <i>si l'établissement est habilité</i>  Tarif libre <i>si l'établissement n'est pas habilité</i>	Tarif administré <i>pour les bénéficiaires de l'aide sociale</i>  Tarif libre <i>pour les non bénéficiaires</i>	Aide sociale départementale <i>-sous condition de ressources</i>  <i>-si l'établissement est habilité</i>
<b>Tarifs dépendance</b> <b>GIR 12, GIR 34,</b> <b>GIR 56</b>	Prix de journée	Forfait dépendance départemental <i>fonction du GMP</i>	APA <i>pour les résidents de GIR 1 à 4</i>
<b>Tarifs soins</b> <b>GIR 12, GIR 34,</b> <b>GIR 56</b>	Prix de journée	Forfait soins national <i>fonction des GMP et PMP</i>	Dotation soins globale <i>(Assurance Maladie)</i>

### 3 Revue de littérature

Les premières analyses du secteur des soins de long terme pour personnes âgées mettent en évidence, dès le début des années 80, les facteurs de coûts en institution, en ayant recours à des estimations classiques de fonctions de coûts en moindres carrés ordinaires (MacKay 1988[21], Dor 1989[10]). Ce n'est en revanche que relativement récemment qu'ont été appliquées des méthodes d'estimation spécifiques de l'efficacité à ce secteur. Filipini (1999)[13] et Farsi et al.(2005)[12] ont ainsi utilisé plusieurs méthodes paramétriques d'étude de l'efficacité sur un échantillon de maisons de retraite suisses : analyse stochastique de frontière, approche déterministe de frontière, ainsi que diverses méthodes de données de panel adaptées à l'étude de l'efficacité. Certains auteurs ont également récemment eu recours à des méthodes de régression quantile, à l'instar de Knox et al.(2007)[16], afin de tester la robustesse de leurs résultats d'estimation.

L'intérêt de telles études est de pouvoir identifier les principaux facteurs de coûts des établissements du secteur afin d'ajuster au mieux les mécanismes de tarification qui le réglementent. L'impact de la taille de l'établissement et la présence de rendements d'échelle peuvent ainsi être testés. Dor (1989)[10] note la présence de déséconomies d'échelles, et également de déséconomies de gamme entre des patients bénéficiaires du Medicare et des patients privés, sur un échantillon national de maisons de retraite aux Etats Unis en 1986. Les seuls travaux empiriques portant à notre connaissance sur les coûts des EHPAD en France sont ceux de Dervaux et al.(2006)[9]. A partir d'un échantillon régional de maisons de retraite et par une méthode d'analyse non paramétrique d'enveloppement de données, ils évaluent une taille critique d'environ 60 lits ; nous vérifions dans cette étude si ce résultat est généralisable à l'ensemble du territoire national, et s'il est robuste au type d'hébergement proposé et au degré de dépendance des résidents accueillis.

Certaines études proposent une comparaison des effets de la catégorie institutionnelle des établissements sur leurs coûts et leur efficacité. Knox et al.(2007)[16] observent ainsi une moindre efficacité des institutions privées à but non lucratif relativement à celles privées à but lucratif sur un échantillon de maisons de retraite au Texas en 1999 et 2002. Peu d'études empiriques ont néanmoins comparé les établissements publics et privés à but non lucratifs. C'est à cette problématique que nous nous intéressons. Les analyses théoriques pouvant justifier a priori d'une plus grande efficacité d'un statut sur l'autre sont plus nuancées lorsqu'il s'agit de comparer entre elles des structures à but non lucratif. Newhouse (1970)[24] montre que lorsque le degré consurrentiel est suffisamment faible, l'allocation des ressources dans des établissements de santé privés à but non lucratif peut ne pas être optimale et se caractériser par une qualité trop élevée : "*there is a bias against producing lower quality products*". A l'inverse, les théories de la bureaucratie soutiennent l'idée d'une plus grande inefficacité dans les institutions publiques puisque les droits de propriété ne peuvent s'exercer pleinement (Niskanen, 1971[26]). Filippini (1999)[13] remarque ainsi une plus grande efficacité des

maisons de retraites privées associatives relativement aux institutions publiques en Suisse. Nous regardons si ce même constat peut être effectué en France, en comparant des résultats d'estimations excluant et intégrant des variables de qualité, afin d'étudier si l'observation éventuelle d'une moindre efficacité d'une catégorie d'établissement serait caractérisée par une plus grande qualité des soins fournis.

La recherche de variables de qualité et le traitement de leur caractère endogène est un apport original de notre étude. Peu d'études sur ce secteur ont approfondi ce problème. De manière générale, l'absence de variables de qualité ou l'intégration de variables de qualité non pertinentes dans une analyse d'efficacité peut biaiser les résultats puisqu'un établissement jugé alors inefficace peut tout simplement être un établissement délivrant plus de qualité. Ceci est d'autant plus vrai pour le secteur des soins de long terme pour personnes âgées au sein duquel la qualité se caractérise principalement par une disponibilité accrue du personnel. Or les charges de personnel représentent environ 80% des coûts en EHPAD. La recherche de l'efficacité dans ce secteur est donc une toute autre problématique que celle de l'efficacité en hôpital. On peut en effet supposer que tout gain en "efficacité" va ici se traduire par une réduction de personnel ou une modification de la structure des qualifications, et donc en une dégradation symétrique de la qualité. Or les ratios d'encadrement semblent être déjà faibles en France - 5.7 pour 10 résidents (Rapport national sur la dépendance, 2011 [28]) - relativement à d'autres pays européens comme l'Allemagne où le taux d'encadrement en institution est de 12 (Sénat, 2004 [20]). Des ouvrages récents retraçant l'expérience d'anciens acteurs du secteur décrivent les conséquences négatives de ce manque de personnel sur la qualité de la prise en charge : pose de couches à des personnes non incontinentes, mise sous tranquillisants ou sous somnifères, sous-alimentation, absence de personnel soignant la nuit et le week end ; *"les effectifs en maison de retraite sont bien au-dessous de ceux de l'hôpital. Certes, il y a moins de soins à faire, mais on devrait être en capacité d'assurer une certaine qualité relationnelle. (...) Pour pouvoir travailler aujourd'hui en maison de retraite sans risquer la déprime, il faut rabaisser le patient au rang d'objet."* (William Réjault, 2009)[29].

Si les gains en efficacité ne peuvent se faire que par un arbitrage coûts-qualité, l'intérêt d'une réforme de la tarification en France serait dans cette situation très limité. La question que l'on se pose est donc celle d'un possible gain en efficacité sur les 20% de charges restantes après soustraction de la masse salariale, ou éventuellement sur certaines charges de personnel qui ne seraient pas liées à la qualité.

## 4 Méthodes

### 4.1 Modélisation

On suppose que les EHPAD transforment des facteurs de production (diverses catégories de travail  $L_i$ ,  $i$ =aides-soignants, infirmiers, agents de services hospitaliers (ASH),...) en une certaine quantité de journées-résidents en hébergement permanent, temporaire ou en accueil de jour ou de nuit (output  $Y$ ). La qualité  $q$  du service de soins et de nursing fait partie intégrante du processus de production. Nous faisons également l'hypothèse que plusieurs variables catégorielles ( $z = \sum_k z_k$ )<sup>8</sup> caractérisant les spécificités de l'environnement de l'EHPAD (variables liées au bâti, accueil de patients atteints d'Alzheimer ou de maladies apparentées), ainsi que le degré de dépendance moyen des résidents  $G$  peuvent influencer sa production. La technologie de production peut ainsi être caractérisée par la fonction de transformation suivante :  $F(Y, q, L, G, z, u) = 0$ , avec  $u$  : inefficacité technique aléatoire de production. Grâce aux propriétés duales des fonctions de production, on peut alternativement expliciter la technologie de production par la fonction de coût total :  $CT = C(Y, q, w, G, z, v)$ , avec  $w_i$  le salaire moyen par catégorie  $i$  de personnel pour l'établissement, et  $v$  un aléa de coût. Nous préférons utiliser une fonction de coût car cette approche permet de formaliser correctement l'activité multiproduit de l'EHPAD (différents modes d'hébergement et degrés de dépendance des résidents). L'estimation d'une fonction de coût requiert en outre comme variables explicatives, non pas les quantités d'inputs, mais leurs prix, probablement moins endogènes.

Afin d'être compatible avec la théorie microéconomique du producteur, cette fonction de coût doit vérifier diverses propriétés que nous examinons après estimation : croissance, concavité et homogénéité de degré 1 par rapport à la quantité d'output et aux prix d'inputs.

8.  $z_k$  peut être décomposé en 2 catégories de variables :  $z_{k,1}$  est un vecteur de variables supposées affecter directement les coûts mais non la relation entre output et coûts, à l'inverse des variables  $z_{k,2}$  qui peuvent avoir un effet sur cette relation output-coûts.

**Modèle 1.** Deux formes fonctionnelles sont utilisées. Nous choisissons tout d'abord une fonction translog (Christensen et al.1973 [5]) qui consiste à approximer la fonction de coût par un développement limité de Taylor à l'ordre 2. Cette forme fonctionnelle présente quelques avantages par rapport à des fonctions moins flexibles de type *Cobb-Douglas* ou à *élasticité de substitution constante (Constant Elasticity of Substitution-CES)*. Elle n'impose notamment pas de restriction *a priori* sur la nature de la technologie et permet ainsi une estimation des rendements d'échelle. Les coefficients de 1<sup>er</sup> ordre peuvent en outre être interprétés comme des élasticités-coût. Notre premier modèle de coût estimé est le suivant :

$$\begin{aligned}
\ln\left(\frac{CT}{w_1}\right) &= \alpha_0 + \alpha_Y \ln(Y) + \sum_{i,i \neq 1} \alpha_{w_i} \ln\left(\frac{w_i}{w_1}\right) + \frac{1}{2} \alpha_{Y^2} (\ln Y)^2 + \sum_{i,i \neq 1} \frac{1}{2} \alpha_{w_i^2} \left(\ln \frac{w_i}{w_1}\right)^2 \\
&+ \sum_{i,i \neq 1} \alpha_{Y.w_i} (\ln Y) \left(\ln \frac{w_i}{w_1}\right) + \sum_{i,i \neq 1} \sum_{j,j \neq i} \alpha_{w_i.w_j} \left(\ln \frac{w_i}{w_1}\right) \left(\ln \frac{w_j}{w_1}\right) \\
&+ \alpha_G G + \sum_{k^1} \alpha_{z_{k^1}} z_{k^1} + \sum_{k^2} \alpha_{z_{k^2}} z_{k^2} + \sum_{k^2} \alpha_{Y.z_{k^2}} Y.z_{k^2} + \alpha_q q + \epsilon
\end{aligned} \tag{1}$$

La normalisation des coûts et des rémunérations de facteurs par le prix d'un des facteurs de production ( $w_1$ ) garantie l'homogénéité de degré 1 du coût total par rapport aux prix d'inputs. L'utilisation d'une telle fonction limite néanmoins notre analyse puisque la log-linéarisation des variables explicatives ne permet pas d'introduire des outputs dont la quantité pourrait être égale à 0 pour certains établissements.

**Modèle 2.** Pour tenir compte de la multi-dimensionnalité de l'output produit par les EHPAD ( $Y = \sum_j Y_j$ , avec  $j$  représentant soit différents degrés de dépendance des résidents, ou soit divers types d'hébergement proposés) et afin d'étudier d'éventuelles économies de gamme entre ces outputs, nous utilisons également une fonction hybride semblable à celle proposée par Grannemann et al. en 1986[14] :

$$CT = A \prod_i w_i^{\alpha_{w_i}} e^{f(Y_j, z_{k^2})} e^\epsilon$$

avec :

$$\begin{aligned}
\ln A &= \alpha_0 + \alpha_{z_{k^1}} z_{k^1} + \alpha_G G + \alpha_q q \\
f &= \sum_j [\alpha_{Y_j} Y_j + \alpha_{Y_j^2} (Y_j)^2] \\
&+ \sum_{j,j' (j \neq j')} \alpha_{Y_j.Y_{j'}} Y_j.Y_{j'} + \alpha_{Y_j^3} \prod_j Y_j \\
&+ \alpha_G G + \alpha_{z_{k^2}} z_{k^2} + \sum_{k^2} \sum_j \alpha_{Y_j.z_{k^2}} Y_j.z_{k^2}
\end{aligned} \tag{2}$$

## 4.2 Méthodes d'estimation

**Analyse stochastique de frontière.** Nous menons dans un premier temps une analyse stochastique de frontière en coupe transversale, en intégrant un terme d'erreur composé permettant de distinguer ce qui relève de l'inefficacité de l'établissement  $e$  ( $u_e$ ), et ce qui relève du bruit ( $v_e$ ) .

$$\ln CT_e = \ln CT(X_e) + u_e + v_e$$

avec  $u_e \geq 0$ ,  $CT_e$  la variable dépendante associée à l'établissement  $e$ , et  $X_e$  le vecteur des variables explicatives de cet établissement  $e$ .

On a alors :

$$CT_e = CT(X_e) * \exp(v_e) * \exp(u_e)$$

$$\text{et } CE_e = \frac{CT_e}{CT(X_e) \cdot \exp(v_e)} = \exp(u_e) \geq 1$$

avec  $CE_e$  le score d'efficacité-coût de l'établissement  $e$ .

$u_e$  et les scores d'efficacité  $CE_e$  sont estimés en utilisant la formulation suggérée par Jondrow et al.(1982)[15]

Cette méthode d'estimation par maximisation de vraisemblance, proposée en 1977 par Aigner et al.[1] et Meeusen et Van den Broeck[22], requiert des hypothèses fortes sur la forme fonctionnelle du modèle de coût ainsi que sur la distribution du terme d'efficacité. Ainsi la forme de la fonction de vraisemblance, et de ce fait les résultats d'estimation, vont notamment dépendre du choix arbitraire de la distribution du terme  $u_e$  (Newhouse, 1994[25]). Nous optons pour une distribution de la perturbation Normale-Normale tronquée, avec  $v_e \sim iidN(0, \sigma_v^2)$ ,  $u_e \sim iidN^+(\mu, \sigma_u^2)$  et  $\mu = \delta m_e + \eta_e$  ( $m$  : variables environnementales explicatives de l'efficacité et  $\eta$  bruit statistique). Nous utilisons ainsi une *approche en une étape* (Battese et Coelli, 1995[2]) pour laquelle le terme d'inefficacité est remplacé dans la fonction de coût par ses variables explicatives. Nous aurions pu utiliser une autre approche rencontrée dans la littérature (Pitt et Lee 1981[27], Vitaliano et Toren 1994[33]) qui consiste à estimer en première étape la frontière stochastique de coût et les scores d'efficacité en supposant ces derniers indépendamment et identiquement distribués, puis à régresser ces scores sur diverses variables explicatives. Cette 2ème étape vient de fait contredire l'hypothèse d'indépendance des termes d'efficacité. Une approche en 2 étapes peut ainsi mener à des estimations non efficaces (Battese et Coelli, 1995[2]), voire des estimations biaisées des scores d'efficacité (Wang et Schmidt 2002[34]).

Compte tenu des choix arbitraires effectués pour cette analyse stochastique de frontière quant à la distribution du terme d'inefficacité et à la forme de la fonction de coûts, il est important de tester la robustesse des résultats d'estimation et des scores d'efficacité. Nous les confrontons donc à des résultats obtenus par des méthodes d'estimation n'utilisant pas ces hypothèses fortes sur la distribution du terme d'erreur : méthode de régression quantile et estimations en données de panel.

**Régression quantile.** Considérons la régression multiple  $CT_e = \alpha + \beta X_e + \epsilon_e$ . La méthode des moindres carrés ordinaires (MCO) repose sur une minimisation de la différence entre  $CT_e$  observé et  $\hat{CT}_e$ , moyenne de  $CT_e$  conditionnellement à  $X_e$ . La méthode de régression quantile, proposée par Koenker et Basset en 1978[17], reprend cette même approche, mais cette fois-ci avec  $\hat{CT}_e$  correspondant au  $k^{ieme}$  percentile de la fonction de coût conditionnellement à  $X_e$ . Il ne s'agit alors plus d'une minimisation de la somme des écarts ( $CT_e - \beta X_e$ ) au carré, mais d'une minimisation des écarts absolus comme suit :

$$\operatorname{argmin}_{\beta} \left[ \sum_{CT_e \geq \beta X_e} k |CT_e - \beta X_e| + \sum_{CT_e < \beta X_e} (1 - k) |CT_e - \beta X_e| \right]$$

En faisant l'hypothèse que le  $k^{ieme}$  percentile (conditionnel à  $X_e$ ) représente les établissements efficaces, la régression quantile QR(k) permet une estimation de la frontière des coûts. Cette estimation serait même plus proche de la *vraie* frontière puisque, relativement à une méthode de maximisation de vraisemblance, l'estimation en régression quantile est moins sensible à la présence d'observations extrêmes.

Si cette méthode permet de s'affranchir d'une décision arbitraire quant à la distribution du terme d'inefficacité, elle nécessite néanmoins un choix quant au percentile à partir duquel les établissements seront considérés comme efficaces. A l'instar de Liu et al.(2008)[19] qui retiennent le 80<sup>ieme</sup> percentile pour estimer une frontière de production, nous choisissons le 20<sup>ieme</sup> percentile pour estimer la frontière de coût. Les termes résiduels obtenus à partir de cette estimation peuvent alors être appréhendés comme des scores d'inefficacité.

**Estimation à effets individuels aléatoires.** Nous effectuons ensuite une analyse sur données de panel afin de s'affranchir de nouveau de l'hypothèse sur la distribution du terme d'erreur (Newhouse, 1994[25]), mais également ici de la non corrélation du terme d'inefficacité avec les régresseurs du modèle (Schmidt et Sickles,

1984[30]). Cette hypothèse de non-corrélation est sans doute trop forte pour une analyse de fonction de coûts, puisque divers facteurs environnementaux peuvent agir à la fois sur les coûts et sur les caractéristiques de la patientèle, la taille des établissements,... Ceci est d'autant plus vrai pour les variables de qualité  $q$  introduites, fortement liées au montant des dépenses autorisées et de ce fait à l'ampleur du pouvoir de négociation de l'établissement vis à vis des autorités de régulation tarifaire. Sous l'hypothèse que ces variables de qualité  $q$  soient corrélées avec l'effet spécifique individuel  $u_e$  et non avec la perturbation  $v_{et}$ , i.e. qu'elles soient corrélées à de l'hétérogénéité non observée constante temporellement, l'utilisation d'un modèle à effets aléatoires avec effets individuels corrélés permet de traiter cette endogénéité. Un autre grand avantage à l'utilisation de données de panel réside dans la prise en compte de l'hétérogénéité non observée dans le calcul des scores d'efficacité, qui inclue certaines dimensions de la qualité difficile à mesurer. Puisque nous ne disposons que de 2 années d'observations, il nous est impossible d'opter pour une estimation en effets fixes, la variance intra-individuelle (Within) relative à la variance inter-individuelle (Between) étant trop faible pour que l'on puisse obtenir des estimations non biaisées. Nous effectuons donc une estimation en effets aléatoires, adaptée à l'analyse de l'efficacité par Schmidt et Sickles, 1984[30] :

$$\ln CT_{et} = \beta_0^* + \beta_X \ln X_{et} + v_{et} + u_e^*$$

avec :

$$\begin{aligned} \beta_0^* &= \beta_0 + E(u_e) \\ u_e^* &= u_e - E(u_e) \end{aligned}$$

Or

$$\min[u_e^*] = -E(u_e), \text{ avec } u_e = 0$$

D'où

$$u_e = u_e^* - \min[u_e^*]$$

Cette estimation de  $\beta$  par moindres carrés généralisés (MCG) engendre des résultats biaisés (biais d'hétérogénéité) si les termes d'inefficacité sont corrélés avec les régresseurs. Pour pallier cette difficulté, nous utilisons un modèle à effets individuels corrélés (spécification de Mundlak, 1978) adapté au modèle de Schmidt et Sickles par Farsi et al.(2005)[12]. On pose alors  $u_e^* = \delta_{X_e} \overline{\ln X_e} + \eta_e$  et  $\overline{\ln X_e} = \frac{1}{T} \sum_t \ln X_{et}$ .

Nous ajoutons pour chacune des méthodes présentées ci-dessus (en coupe et en données de panel) des estimations par méthode de variables instrumentales<sup>9</sup> afin de traiter l'endogénéité des variables de qualité.

A partir de ces estimations, les élasticités d'échelle globales et les économis de gamme sont évaluées. L'utilisation de diverses méthodes d'estimation permet de tester la robustesse des résultats d'estimation, des rendements d'échelles calculés et des scores d'inefficacité obtenus. Nous ne nous attendons pas à obtenir des termes d'inefficacité identiques, la valeur absolue de ce terme étant difficilement interprétable : *"the different efficiency scores should not (...) be interpreted as accurate point estimates of efficiency, but might more usually be interpreted as indicating general trends in efficiency"* (Jacobs, 2001). Nous étudions en revanche dans quelle mesure la variabilité de ces termes d'inefficacité est importante, et si la catégorie institutionnelle et la prise en compte de la qualité ont un effet sur l'efficacité. L'impact du statut sur l'efficacité-coût est étudiée par des tests paramétrique et non paramétrique (Kruskal-Wallis). Nous cherchons ainsi, par comparaison de ces diverses méthodes, à montrer la robustesse de ces tendances.

### 4.3 Variables

Il est fondamental lorsque l'on veut étudier l'efficacité d'acteurs agissant dans un secteur économique particulier, de bien appréhender toutes les spécificités propres à ce secteur afin d'avoir une fonction de

9. Le traitement de l'endogénéité est présenté en annexe.



production (ou une fonction de coûts par propriétés duales) caractérisant le mieux possible leur processus de production. Si la complexité de la technologie de production n'est en effet pas prise en compte dans la formalisation de la fonction de coût, alors cette complexité non modélisée va se reporter sur les termes d'erreurs, et les estimations des scores d'efficacité seront biaisées (Kumbhakar et Lovell, 2000[18]).

Nous avons supposé ci-dessus qu'une maison de retraite produit un output  $Y$ , correspondant au nombre annuel de journées-résidents. Selon le modèle étudié cet output est soit intégré comme tel, soit décomposé selon le degré de dépendance des résidents (en fonction du groupe GIR, *Groupe Iso-Ressources*, des résidents :  $Y_{GIR12}$ ,  $Y_{GIR34}$  et  $Y_{GIR56}$ ) ou selon le mode d'hébergement proposé (journées-résidents en hébergement permanent  $Y_{HP}$ , en hébergement temporaire  $Y_{HT}$ , ou en accueil de jour  $Y_{AJ}$ ).

La production de cet output génère un coût  $CT$ , que l'on assimile ici au coût total annuel de l'établissement. Ce coût est un coût de long-terme, incluant à la fois les coûts variables (coût de fonctionnement, rémunération du personnel) et également les coûts fixes de l'EHPAD (coût de la structure, charges d'investissement).

Les prix correspondant aux inputs travail  $w_i$  correspondent aux salaires moyens bruts annuels du personnel de l'établissement exerçant une fonction  $i$ . Nous retenons dans notre modélisation les 3 principales catégories de personnel présentes en maison de retraite : aide-soignant ou aide médico-psychologique  $w_{AS}$ , agent de service hospitalier  $w_{ASH}$ , et infirmier  $w_{INF}$ .

Puisque nous estimons une fonction de coût de long-terme, nous n'avons pas dans notre modélisation de variable de capital. Il nous est en revanche nécessaire d'intégrer une variable de coût du capital, mais il est difficile de disposer d'une telle variable qui soit pertinente. Nous préférons donc intégrer diverses variables liées au coût du bâti, telles que :

-*BATI* : le nombre d'années depuis la construction du bâtiment ou, le cas échéant, de sa dernière rénovation ;  
 -*PROP* : une variable catégorielle apportant une information sur la propriété de la structure : le gestionnaire est-il propriétaire, locataire d'un bâtiment public, d'un bâtiment privé associatif, à but lucratif ou locataire d'un bâtiment HLM ?

-*URB* : le degré d'urbanisation de la commune (ville) de résidence de l'établissement : commune inférieure à 20000 habitants, ville comprenant entre 20000 et 200000 habitants, supérieure à 200000 habitants, ou ville localisée en région parisienne.

Un ensemble de variables représentatives des caractéristiques des résidents accueillis sont également incluses dans le modèle. Il s'agit de la proportion de résidents de chaque groupe GIR dans l'établissement :  $propGIR1$  à  $propGIR6$ . Nous ajoutons également une variable binaire indiquant si l'établissement accueille ou non des patients atteints de la maladie d'Alzheimer ou maladie apparentée *ALZ*. Puisque cette variable peut avoir un effet différent sur les coûts selon l'effectif total de l'établissement (amortissement des aménagements spécifiques pour les patients atteints de démences séniles), nous intégrons une variable d'interaction entre l'output et l'accueil de patients Alzheimer :  $Y.ALZ$ .

Nous prenons de plus en compte diverses caractéristiques de l'environnement de l'établissement qui peuvent avoir un impact sur ses coûts : une variable binaire *OS* et sa variable d'interaction avec l'output  $Y.OS$  informant du choix de l'établissement quant à son option tarifaire liée aux soins<sup>10</sup>, et une autre *PUI* indiquant si l'établissement dispose ou non d'une pharmacie à usage intérieur (PUI). Concernant l'option tarifaire choisie, on s'attend à un effet positif du choix de l'option tarif global sur les coûts, puisque des frais supplémentaires sont alors à la charge de l'établissement. La disposition d'une PUI a également un impact positif attendu sur les coûts, puisqu'elle génère un surcroît de coûts fixes (coût de la structure, rémunération d'un pharmacien salarié). Sa mise en place peut certes se traduire par une meilleure gestion du médicament et une diminution des dépenses qui leurs sont associées, mais celles-ci sont payées par les personnes âgées et n'ont donc pas d'effet sur les charges de l'établissement.

**Variables de qualité.** L'introduction de variables de qualité  $q$  est délicate. Une non prise en compte de la qualité est susceptible de biaiser l'estimation des scores d'efficacité (Mutter et al., 2008[23]), mais la qualité de la prise en charge en établissement médico-social est difficile à observer et à mesurer. Faute d'une meilleure mesure, le taux d'encadrement semble être un proxy pertinent de la qualité en EHPAD. Dans une étude réalisée en 2009, la Fédération Hospitalière de France et le Syndicat National de la Gériatrie pointent les divergences en termes de qualité des soins qui peuvent être engendrées par différents ratios

10. Les EHPAD ont un droit d'option tarifaire (art R.314-167 CASF) concernant les dépenses liées à la section soin : ils peuvent choisir l'option *tarif partiel* ou *tarif global*. Le panier de soins remboursés est alors plus large pour les établissements ayant choisi l'option *tarif global*, puisqu'il inclut les honoraires et prescriptions de médecins généralistes, les examens de biologie et de radiologie. Lorsque l'établissement choisit l'option *tarif partiel*, ces dépenses sont à la charge de la personne âgée, qui est alors remboursée par sa propre caisse d'Assurance Maladie.

personnel soignant/résidents[7]. Une enquête récente réalisée auprès de 1802 résidents en EHPA (Résidents EHPA, DREES) en 2007 montre qu'un des plus grand facteurs d'insatisfaction est le manque de disponibilité du personnel[11]. De manière à prendre en compte la diversité des besoins en personnel soignant selon les EHPAD, nous calculons un encadrement théorique optimal ( $N^*$ ) en fonction des catégories GIR des résidents. Nous utilisons pour cela les normes présentes dans le secteur de la prise en charge à domicile, qui permettent de calculer des équivalents temps plein (ETP) journaliers de personnel nécessaires à chaque personne âgée, afin de déterminer les montants des plans d'aide pour l'APA. Ce calcul d'ETP s'effectue en fonction du GIR attribué à l'individu<sup>11</sup> : une personne en GIR 1 a ainsi besoin quotidiennement de 1 ETP, en GIR 2 de 0.84, en GIR 3 de 0.66, en GIR 4 de 0.42, en GIR 5 de 0.25 et enfin en GIR 6 de 0.07. Nous rapportons ensuite l'encadrement effectif ( $N$ ) (incluant tout le personnel non administratif<sup>12</sup>) à cet encadrement théorique calculé :  $N/N^*$ . Outre le taux d'encadrement, la structure de qualification du personnel a probablement, toutes choses égales par ailleurs, un effet sur la qualité de la prise en charge en établissement. Nous définissons ainsi une deuxième variable de qualité *NonQualif*, calculée comme le rapport de personnel non soignant (ASH et personnel socio-éducatif) sur le personnel soignant (infirmier, aide-soignant et AMP).

Nous vérifions préalablement que ces deux variables sont des variables de qualité pertinentes puisqu'elles peuvent avoir un impact sur l'état de santé des résidents. Nous disposons pour un échantillon restreint de résidents en EHPAD<sup>13</sup> (7043 individus) de 3 variables d'état de santé : une variable catégorielle ordonnée du degré d'incontinence du résident (codée de 0 à 3, 0 correspondant à un résident non incontinent et 3 à une incontinence élevée), une variable catégorielle ordonnée reflétant l'état dépressif du résident (codée également de 0 à 3), et une variable binaire indiquant si le résident a fait une ou plusieurs chutes au cours de l'année. Nous effectuons 3 estimations à partir de ces données, 2 régressions logistiques ordonnées et une régression logistique binomiale.<sup>14</sup> Les coefficients et les indicateurs de significativité associés aux variables de qualité sont présentés ci-dessous (Table 2). Une hausse du ratio d'écart à la norme du taux d'encadrement a, toutes choses égales par ailleurs, un impact positif sur l'état de santé des résidents (corrélation négative avec la dégradation de l'état de santé). La présence de personnel non qualifié augmente également les risques d'incontinence.

TABLE 2 – Régressions logistiques de variables d'état de santé des résidents sur  $N/N^*$  et *NonQualif*

Variable dépendante :	Degré d'incontinence	Etat dépressif	Chutes
$N/N^*$	-0.212	-0.35*	-0.247*
<i>NonQualif</i>	0.300*	-0.029	-0.054
$N/N^* \cdot NonQualif$	0.336*	-0.011	-0.046

Seuil de significativité : \* : 5%

Intégrer de telles variables de qualité dans notre modèle semble cependant poser un problème d'endogénéité.

Nous avons vu précédemment que le système de financement actuel d'un EHPAD  $e$  au prix de journée favorise une reconduction périodique de son budget comme suit :

$$B_{t=1}^e = B_{t=0}^e(\tau^e + 1) \quad (3)$$

Or on peut connaître le budget optimal de l'EHPAD  $e$  qui lui permettrait de prendre en charge ses résidents de manière satisfaisante sans lui laisser de rentes. Celui-ci est principalement fonction de ses besoins d'encadrement spécifiques  $N^*$  et des salaires prévalant sur son marché du travail environnant :

$$B^{*e} = B(N^{*e}, \sum_i w_i^e) \quad (4)$$

11. Ratios d'encadrement préconisés par le Plan Solidarité Grand Age 2007-2012.

12. Nous avons ajouté aux ETP de personnels salariés dans l'établissement les ETP correspondant aux heures de sous-traitance, d'interim, ou de recours à des personnels libéraux.

13. Ces données proviennent de la fiche PATHOS de l'enquête EHPA 2007.

14. Les variables de contrôle intégrées dans ces régressions sont les suivantes : âge et catégorie GIR du résident, variable binaire indiquant si le résident est mobile, s'il se plaint de la douleur, variable de statut de l'établissement, effectif total des résidents hébergés, proportion de bénéficiaires à l'aide sociale, présence dans l'établissement de chambres de fin de vie.

On peut alors avoir  $B^e$  supérieur, égal ou inférieur à  $B^{*e}$ .

Si  $B^e \geq B^{*e}$ , on peut faire l'hypothèse que le recrutement en personnel soignant  $N^e$  n'est pas fonction du budget  $B^e$  alloué à l'établissement  $e$ , le système de paiement serait alors un prix de journée stricto-sensu par lequel tous les coûts de l'établissement quels qu'ils soient lui sont remboursés.

Si en revanche  $B^e < B^{*e}$ , le recrutement de l'établissement en personnel soignant sera contraint. Soit  $\tilde{N}^e = f(N^{*e}, \sum_i w_i^e)$  sa demande notionnelle de travail. On a dans cette situation :

$$\tilde{N}^e \geq N^e = \frac{B^{*e} - \text{autres frais}}{w^e} \quad (5)$$

Il pourrait alors exister un biais de simultanéité entre  $N/N^*$  et  $B^e$  lorsque  $N/N^*$  est inférieur à 1.

Pour traiter ce problème d'endogénéité des variables de qualité, nous utilisons une méthode de variables instrumentales<sup>15</sup>. Nous comparons ensuite, pour chacune des méthodes d'estimation (analyse stochastique de frontière et régression quantile), les résultats de 3 régressions : estimation sans variable de qualité, avec variables de qualité  $N/N^*$ , *NonQualif*, et avec variables de qualité instrumentées  $\widehat{N/N^*}$ ,  $\widehat{NonQualif}$ . Nous utilisons comme instruments l'ancienneté moyenne en nombre d'années du personnel (hors personnel administratif et de direction) dans l'établissement (*ANC* et *ANC<sup>2</sup>*), en supposant qu'une moindre qualité de la prise en charge et un épuisement du personnel, liés à un faible ratio d'encadrement ou/et à une faible qualification, engendreraient un turn-over important (Spilisbury et al., 2011[31]).

**Variables explicatives des scores d'inefficacité.** Une des difficultés dans l'estimation de frontières est de savoir quelles variables inclure comme variables explicatives du coût, et lesquelles réserver pour expliquer l'inefficacité (Vitaliano et Toren, 1994[33]). Puisque le terme d'inefficacité suit une loi normale tronquée, même lorsque l'on opte pour une procédure en une étape (Battese et Coelli, 1995[2]), le choix d'intégrer une variable comme élément explicatif du coût ou du terme d'inefficacité va modifier la fonction de vraisemblance et les résultats d'estimation. La question se pose notamment pour la variable de catégorie institutionnelle *STAT*. Nous cherchons dans cette étude à mettre en évidence des différences éventuelles d'efficacité-coût selon le statut des EHPAD (public rattaché à une structure hospitalière, public non rattaché ou privé associatif). Pour cette raison nous choisissons d'intégrer le statut comme variable explicative de l'inefficacité et non du coût. Nous devons néanmoins être prudents dans l'analyse des scores d'efficacité, puisque le statut peut aussi avoir un impact direct sur les coûts, qu'il s'agisse des coûts fixes liés au bâtiment ou des coûts salariaux<sup>16</sup>. En contrôlant par les salaires bruts moyens propres à l'établissement et par le type de propriété du bâti, nous avons néanmoins partiellement réduit cet impact direct du statut sur les coûts. Nous ajoutons, comme variables explicatives de l'inefficacité, la proportion de bénéficiaires à l'aide sociale départementale *BENEF* et le logarithme du PIB/hab. du département de résidence de l'établissement *PIB*. Ces deux variables n'ont a priori aucun impact sur les coûts, mais peuvent néanmoins avoir un effet sur la latitude budgétaire dont disposent les établissements, et donc sur leur efficacité.

15. Le traitement de l'endogénéité des variables de qualité est présenté en annexe.

16. Les différences de charges patronales ne sont par exemple pas incluses dans notre analyse. Or celles-ci sont moins élevées pour les établissements publics : le taux de charges patronales s'élève en moyenne à 50.9%, contre 57.9% pour les établissements privés associatifs (Rapport du cabinet AUMERAS, 2010). Nous avons fait le choix de ne pas les inclure puisqu'il existe en contrepartie des charges financières supplémentaires pour les établissements publics, à l'instar de la rémunération du personnel en congé maladie, qui réduisent ces différences.

TABLE 3 – Récapitulatif des variables utilisées

Y	modèle 1 : Y modèle 2 : $Y_{HP}$ , $Y_{HT}$ , $Y_{AJ}$ et $Y_{GIR12}$ , $Y_{GIR34}$ , $Y_{GIR56}$
$w_i$	$w_{AS}$ , $w_{ASH}$ , $w_{INF}$
$z_{k_1}$	$PUI$ , $BATI$ , $URB$ , $PROP$
$z_{k_2}$	$ALZ$ , $OS$
G	$propGIR1$ , $propGIR2$ , $propGIR3$ , $propGIR4$ , $propGIR5$ , $propGIR6$
q	$N/N^*$ , $NonQualif$
Variable explicatives du score d'efficacité (m)	$STAT$ , $BENEF$ , $PIB$

## 5 Données

Pour les méthodes de régression en coupe transversale, nous utilisons un échantillon national de 1171 maisons de retraite. Les données de salaires proviennent de l'enquête DADS 2008 (Déclaration annuelle de données sociales) de l'INSEE. Les établissements fournissent annuellement des informations relatives à chacun de leurs salariés (salaire brut, ETP, âge, type de contrat, type de convention collective). A partir de ces données nous pouvons calculer un salaire brut moyen par catégorie de personnel et par établissement, ce qui constitue un véritable atout pour ce type d'étude. Les autres variables nous sont renseignées par l'enquête exhaustive quadriennale EHPA 2007 (DREES). Nous ne disposons pas de données brutes sur les coûts des établissements. Néanmoins puisque les tarifs dépendance, soins et hébergement des établissements habilités sont des prix de journée, et qu'ils tiennent compte des résultats antérieurs qui ne peuvent être affectés librement par les EHPAD, on fait l'hypothèse que  $\sum \text{tarifs} * \text{output} = CT$ . Nous multiplions alors les journées-résidents en hébergement permanent avec les tarifs concernant ce type d'hébergement, et celles en accueil de jour ou en hébergement temporaire avec les tarifs respectivement fixés pour l'accueil de jour ou pour l'hébergement temporaire. L'information sur les tarifs de ces types d'hébergement alternatifs à l'accueil permanent nous est donnée par l'enquête annuelle MAUVE 2010 (DREES). Cette enquête est effectuée auprès de tous les établissements recevant des personnes atteintes d'Alzheimer ou de maladies apparentées, ce qui est le cas de la plupart des établissements proposant de l'accueil de jour ou de l'hébergement temporaire. Nous avons supprimé de notre échantillon les établissements non répertoriés dans l'enquête MAUVE et pour lesquels ces types d'hébergement représentent plus de 5% de leur nombre total de journées-résidents. Pour les établissements non répertoriés et pour lesquels l'hébergement permanent représente plus de 95% de leur activité, nous ignorons leur activité d'hébergement alternatif.

Nous présentons ci-dessous les statistiques descriptives pour chacune des variables utilisées en fonction du statut de l'établissement (Table 4).

On remarque que le coût journalier par résident est plus élevé pour les établissements publics rattachés à un hôpital que pour les établissements privés associatifs, eux-mêmes plus coûteux que les établissements publics non rattachés. Nous cherchons dans cette étude à mettre en évidence les facteurs de coûts responsables de ces différences. Nous observons une activité plus forte pour les établissements publics, principalement pour les établissements publics rattachés. La structure de dépendance des résidents est quasiment identique selon le statut, bien que les établissements privés associatifs soient plus nombreux à n'accueillir aucun patient atteint de la maladie d'Alzheimer ou maladie apparentée. Concernant les variables de qualité, on observe des ratios  $N/N^*$  plus élevés pour les établissements publics, mais les établissements publics non rattachés à un hôpital recrutent cependant plus de personnels non qualifiés. Les établissements publics sont plus fréquemment propriétaires du bâtiment ou, si tel n'est pas le cas, sont alors souvent en location publique. Quant aux coûts salariaux, si les rémunérations nettes du personnel travaillant en institution publique sont en moyenne plus élevées qu'en établissement privé<sup>17</sup>, les coûts salariaux bruts annuels semblent relativement plus élevés pour les établissements privés associatifs (concernant les rémunérations du personnel aide-soignant-AMP et

17. A titre comparatif, pour ce même échantillon, le salaire net annuel moyen d'un aide-soignant en établissement privé associatif est de 18120 euros, 20695 euros en établissement public non rattaché à un hôpital et 21141 euros en établissement public hospitalier. Le salaire moyen net infirmier est respectivement de 25002 euros, 26368 euros et 27193 euros.

des agents de services hospitaliers). Ceci est dû aux exonérations d'une partie des charges salariales dont bénéficient les établissements publics.

TABLE 4 – Statistiques descriptives des variables utilisées (Estimations sur données en coupe)

Variables	Privé associatif	Public non hospitalier	Public hospitalier	Total
	Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )
CT	2272833 (1478669)	2339552 (1349127)	3728178 (2086564)	2334382 (1432294)
Y	27082 (13127)	29046 (14376)	40185 (22286)	28386 (14125)
CT/Y	82.8 (15.4)	80 (11.4)	91.1 (19.1)	81.4 (13.5)
$w_{AS}$	23925 (1815)	24350 (1396)	24577 (1154)	24170 (1602)
$w_{inf}$	33224 (3435)	31263 (3301)	32055 (3049)	32126 (3489)
$w_{ASH}$	21002 (1560)	20259 (865)	20488 (812)	20584 (1268)
propGIR1	0.175 (0.093)	0.173 (0.085)	0.142 (0.102)	0.174 (0.089)
propGIR2	0.304 (0.097)	0.307 (0.080)	0.282 (0.110)	0.306 (0.089)
propGIR3	0.139 (0.064)	0.146 (0.061)	0.165 (0.059)	0.144 (0.063)
propGIR4	0.203 (0.083)	0.213 (0.080)	0.221 (0.108)	0.209 (0.082)
propGIR5	0.090 (0.063)	0.093 (0.060)	0.083 (0.053)	0.091 (0.061)
propGIR6	0.089 (0.089)	0.067 (0.067)	0.107 (0.119)	0.077 (0.079)
BATI	15.6 (14.6)	18.2 (18.5)	15.9 (8.92)	17 (16.8)
N/N*	0.813 (0.170)	0.937 (0.175)	0.959 (0.174)	0.884 (0.184)
NonQualif	1.27 (0.637)	1.5 (0.534)	1.07 (0.496)	1.39 (0.592)
ANC	6.71 (2.92)	8.32 (2.54)	9.65 (3.44)	7.65 (2.85)
PIB	25016 (8365)	23493 (4716)	24468 (3298)	24169 (6575)
BENEF	0.326 (0.306)	0.328 (0.277)	0.325 (0.207)	0.327 (0.289)
	N (%)	N (%)	N (%)	N (%)
PROP=propriété	231 (45.6%)	400 (62.1%)	17 (85%)	648 (55.3%)
PROP=loc.public	48 (9.5%)	218 (33.8%)	2 (10%)	268 (22.9%)
PROP=loc.HLM	98 (19.3%)	25 (3.9%)	1 (5%)	124 (10.6%)
PROP=loc.asso	70 (13.8%)	1 (0.2%)	0 (0%)	71 (6.1%)
PROP=loc.privé	60 (11.8%)	0 (0%)	0 (0%)	60 (5.1%)
URB=x<20000	213 (42%)	316 (49.1%)	15 (75%)	544 (46.5%)
URB=20000<x<200000	110 (21.7%)	144 (22.4%)	0 (0%)	254 (21.7%)
URB=200000<x	147 (30%)	149 (23.1%)	3 (15%)	299 (25.5%)
URB=Paris	37 (7.3%)	35 (5.4%)	2 (10%)	74 (6.3%)
PUI=0	475 (93.7%)	589 (91.5%)	4 (20%)	1068 (91.2%)
PUI=1	32 (6.3%)	55 (8.5%)	16 (80%)	103 (8.8%)
OS=0	396 (78.1%)	450 (69.9%)	5 (25%)	851 (72.7%)
OS=1	111 (21.9%)	194 (30.1%)	15 (75%)	320 (27.3%)
ALZ=0	231 (45.6%)	369 (57.3%)	11 (55%)	611 (52.2%)
ALZ=1	276 (54.4%)	275 (42.7%)	9 (45%)	560 (47.8%)
<b>Nb. obs.</b>	<b>507</b>	<b>644</b>	<b>20</b>	<b>1171</b>

Pour notre modélisation en données de panel, nous avons recours à un autre échantillon de 710 maisons de retraite, pour lequel nous disposons de données pour les années 2003 et 2007. Nous utilisons les mêmes variables que pour les analyses sur données en coupe, ormis *BATI*, *PROP*, *OS* pour lesquelles nous n'avons pas d'information en 2003. Nous n'avons pas non plus d'indication sur l'ancienneté du personnel en 2003, nous ne pouvons donc pas utiliser de méthode d'estimation par variables instrumentales. Pour les observations de l'année 2003 nous utilisons l'enquête EHPA 2003. Nous n'avons pas d'information quant aux tarifs de l'hébergement temporaire et de l'accueil de jour, par conséquent nous utilisons comme variable dépendante le nombre de journées-résidents uniquement en hébergement permanent. Concernant les coûts salariaux nous utilisons l'enquête DADS 2005 (INSEE). Nous ne disposons que d'un échantillon au 1/12ème des données de cette enquête. Afin de conserver un nombre satisfaisant de maisons de retraite, nous inférons les salaires bruts moyens par établissement, en estimant un salaire pour chaque individu en fonction de son activité, de son type de contrat, de sa convention collective, de son âge et de la localisation géographique de l'établissement. Nous procédons de la même façon pour l'année 2007 afin de perdre le moins possible d'observations. Nous présentons les statistiques descriptives des variables en 2003 et 2007 à la Table 5.

Les coûts journaliers par résident ont augmenté d'environ 20% entre 2003 et 2007, cette hausse étant légèrement moins importante pour les établissements publics non rattachés à un hôpital. Cette augmentation est probablement due en partie à la hausse observée des rémunérations brutes, ainsi qu'à une augmentation du niveau de dépendance des résidents et à un accueil plus fréquent de patients atteints de la maladie d'Alzheimer. Ceci confirme l'idée que les personnes âgées entrent de plus en plus tardivement en institution. La hausse du ratio d'écart à la norme du taux d'encadrement (taux qui se rapprochent de 1) et l'augmentation relative du personnel soignant peut sans doute s'expliquer au regard de cet évolution de la structure de dépendance des résidents.

TABLE 5 – Statistiques descriptives des variables utilisées (Estimations en données de panel)

Variables	Année	Privé associatif	Public non hospitalier	Public hospitalier	Total
		Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )	Moy. ( $\sigma^2$ )
CT (pour heb.permanent)	2003	1918060 (1078983)	1921209 (1299842)	2700977 (1713001)	2141771 (1428630)
	2007	2538410 (1509929)	2377857 (1461653)	3648355 (2283580)	2771637 (1824957)
Y (en heb. permanent)	2003	28380 (14067)	28583 (14636)	36571 (19712)	30803 (16505)
	2007	29917 (15809)	29142 (14422)	39676 (22876)	32280 (18078)
$CT/Y$	2003	67.2 (12.3)	65.9 (11.3)	72.4 (13.3)	68.1 (12.4)
	2007	84.3 (14.5)	80.1 (11.6)	92.9 (25.5)	84.7 (18)
$w_{AS}$	2003	22915 (956)	19087 (775)	19344 (566)	20039 (1753)
	2007	24420 (1173)	23187 (1312)	23615 (1292)	23589 (1364)
$w_{inf}$	2003	32122 (1015)	28178 (1066)	28205 (766)	29092 (1923)
	2007	33270 (1167)	31858 (1828)	32248 (1472)	32290 (1693)
$w_{ASH}$	2003	20910 (1058)	17233 (660)	17275 (591)	18090 (1715)
	2007	21358 (1380)	19850 (1299)	20019 (1343)	20242 (1463)
propGIR1	2003	0.166 (0.087)	0.152 (0.079)	0.108 (0.094)	0.143 (0.088)
	2007	0.190 (0.093)	0.173 (0.085)	0.142 (0.109)	0.168 (0.095)
propGIR2	2003	0.292 (0.106)	0.287 (0.085)	0.247 (0.103)	0.277 (0.097)
	2007	0.319 (0.102)	0.302 (0.077)	0.278 (0.105)	0.299 (0.093)
propGIR3	2003	0.132 (0.059)	0.147 (0.065)	0.158 (0.068)	0.147 (0.065)
	2007	0.136 (0.053)	0.149 (0.063)	0.169 (0.074)	0.151 (0.065)
propGIR4	2003	0.208 (0.094)	0.210 (0.087)	0.235 (0.099)	0.217 (0.092)
	2007	0.206 (0.082)	0.214 (0.081)	0.226 (0.094)	0.215 (0.085)
propGIR5	2003	0.100 (0.067)	0.112 (0.069)	0.135 (0.094)	0.116 (0.078)
	2007	0.081 (0.053)	0.096 (0.059)	0.106 (0.083)	0.095 (0.066)
propGIR6	2003	0.102 (0.108)	0.092 (0.083)	0.117 (0.108)	0.101 (0.097)
	2007	0.067 (0.077)	0.066 (0.065)	0.079 (0.082)	0.070 (0.073)
N/N*	2003	0.758 (0.159)	0.857 (0.187)	0.907 (0.263)	0.849 (0.212)
	2007	0.813 (0.150)	0.940 (0.183)	0.918 (0.206)	0.805 (0.189)
NonQualif	2003	1.39 (1.06)	1.43 (0.838)	0.944 (0.552)	1.28 (0.858)
	2007	0.964 (0.464)	1.06 (0.447)	0.786 (0.353)	0.96 (0.442)
BENEF	2007	0.259 (0.276)	0.249 (0.196)	0.332 (0.240)	0.275 (0.231)
	2003	0.328 (0.305)	0.341 (0.289)	0.419 (0.292)	0.360 (0.295)
PIB	2003, 2007	23506 (3358)	23620 (4670)	23196 (4559)	23474 (4372)
		N (%)	N (%)	Moy. (%)	N (%)
URB=x<20000	2003, 2007	68 (40.2%)	176 (48.5%)	97 (46.6%)	341 (46.1%)
URB=20000<x<200000	2003, 2007	39 (23.1%)	83 (22.9%)	41 (19.7%)	163 (22%)
URB=200000<x	2003, 2007	49 (29%)	79 (21.7%)	59 (28.4%)	187 (25.3%)
URB=Paris	2003, 2007	13 (7.7%)	25 (6.9%)	11 (5.3%)	49 (6.6%)
PUI=0	2003	151 (88.8%)	312 (86.7.3%)	16 (7.6%)	479 (64.7%)
	2007	151 (89.3%)	325 (89.5%)	23 (11.1%)	499 (67.4%)
PUI=1	2003	19 (11.2%)	48 (13.3%)	194 (92.4%)	261 (35.3%)
	2007	18 (10.6%)	38 (10.5%)	185 (88.9%)	241 (32.6%)
ALZ=0	2003	164 (96.5%)	331 (91.9%)	208 (99%)	703 (95%)
	2007	72 (42.6%)	197 (54.3%)	97 (46.6%)	366 (49.5%)
ALZ=1	2003	6 (3.5%)	29 (8.1%)	2 (1%)	37 (5%)
	2007	97 (57.4%)	166 (45.7%)	111 (53.4%)	374 (50.5%)
Nb. obs.	2003	<b>170</b>	<b>360</b>	<b>210</b>	<b>740</b>
	2007	<b>169</b>	<b>363</b>	<b>208</b>	<b>740</b>

## 6 Résultats

### 6.1 Etude de l'efficacité par statut et des rendements d'échelle : modèle 1

**Tests de spécification.** A partir de l'estimation stochastique de frontière (estimation (1)), nous testons l'hypothèse  $H_0 : \sigma_u^2 = 0$  contre l'hypothèse alternative  $H_1 : \sigma_u^2 > 0$  via un test de vraisemblance. Nous obtenons un ratio de vraisemblance  $LR = 5.28$  avec une p-value de 0.011. L'hypothèse nulle est donc rejetée, utiliser un modèle de frontière stochastique semble être pertinent<sup>18</sup>.

Dans la même idée, un test réalisé à partir de la skewness des résidus obtenus lors d'une estimation avec distribution normale tronquée des erreurs permet d'étudier la significativité du terme d'inefficacité  $u$ . Coelli (1995)[6] a en effet montré que la présence d'un terme d'efficacité déforme négativement les résidus obtenus par régression MCO. On obtient ici par un test de normalité des résidus (z-test) une p-value de 0.020 ( $z=1.917$ ) ; on rejette l'hypothèse de normalité, ce qui confirme la significativité des termes d'inefficacité dans le modèle.

**Résultats d'estimation.** Le modèle translog (modèle 1) est estimé par diverses méthodes : analyse stochastique de frontière (SFA), régression quantile (RQ) et estimation en panel avec effets aléatoires (EA) et effets individuels corrélés (MEIC). Les résultats d'estimation sont reportés ci-dessous (Tables 6 et 7). Nous précisons pour chacune de ces estimations si les variables de qualités  $q$  ont été intégrées, et le cas échéant si l'endogénéité a été corrigée par méthode de variables instrumentales ( $\hat{q}$ )<sup>19</sup>. Compte tenu du biais d'endogénéité qui peut être engendré par l'intégration de variables de qualité non instrumentées, nous nous intéressons plus particulièrement aux résultats des estimations (1), (3), (4) et (6), et aux estimations en données de panel (7), (9) et (10).

Les résultats sont dans l'ensemble relativement proches, ormis pour les coefficients associés à l'output  $Y$  et aux prix des facteurs  $w_{AS/INF}$  et  $w_{ASH/INF}$ . Puisque nous avons utilisé un modèle de coût translog, les coefficients de premier ordre associés à ces variables sont interprétables comme des élasticités-coûts. Selon le modèle retenu, l'output a un effet positif significatif au premier ou au second ordre. Conformément à nos attentes, une hausse des salaires induit une augmentation des coûts, mais cet effet est décroissant pour le salaire moyen aide-soignant. Concernant le salaire moyen des agents de service, les coefficients de 1<sup>er</sup> et de 2<sup>eme</sup> ordre sont tous les deux positifs, l'impact d'une augmentation de  $w_{ASH/INF}$  sur les coûts est donc croissant. Les termes d'interaction entre les rémunérations d'aide-soignants et d'ASH peuvent s'interpréter comme des coefficients de substitution entre ces 2 catégorie de personnel : lorsque ces termes sont par exemple négatifs, cela signifie qu'une hausse du salaire aide-soignant engendre un moindre effet du salaire ASH sur les coûts, ce qui peut être dû à une diminution du recrutement d'ASH engendrée par la réduction des effectifs d'aide-soignants (effet de complémentarité) en réponse à la hausse de  $w_{AS/INF}$ . Il est néanmoins difficile de tirer une telle conclusion ici puisque ce coefficient est positif pour les estimations sur données en coupe mais négatif pour les modèles avec données de panel. Les termes d'interaction entre l'output  $Y$  et les salaires, lorsqu'ils sont significativement négatifs, peuvent quant à eux s'interpréter comme des effets d'échelle : plus l'activité de l'établissement est importante, moins la hausse du salaire d'une catégorie de personnel n'a d'effet sur ses coûts.

Concernant le degré de dépendance des résidents, plus la proportion de GIR 3 à 6 est importante, moins les coûts sont élevés. La proportion de personnes en GIR 1 relativement à la proportion de personnes en GIR 2 ne semble pas en revanche avoir d'effet<sup>20</sup>. L'accueil de patients Alzheimer aurait toutes choses égales par ailleurs un effet négatif sur les coûts<sup>21</sup>, mais de moindre importance lorsque  $Y$  augmente. Au sujet du coût du bâtiment, être en location HLM ou en location privée à but lucratif est plus coûteux, ainsi qu'être localisé en région parisienne. Une construction ou une rénovation récente est générateur de coûts, puisque l'ancienneté du bâtiment a un effet négatif significatif. Disposer d'une pharmacie à usage intérieure augmente également

18. Le test de ratio de vraisemblance ne correspond pas à un test de vraisemblance classique puisqu'il s'agit d'un test asymétrique ; on utilise donc un test du ratio de vraisemblance asymétrique, pour lequel le ratio de vraisemblance doit suivre une distribution de probabilité moitié Chi-2(1ddl)/moitié 0.

19. Les écarts-types associés aux régressions effectuées après traitement de l'endogénéité ( $\hat{q}$ ) correspondent ici aux écarts-types de seconde étape. Ces écarts-types, ainsi que les indicateurs de significativité, devront par la suite être corrigés.

20. Le classement d'un résident dans l'une ou l'autre des catégories GIR (1 ou 2) n'a pas d'impact sur le montant des allocations qui lui sont versées. Par conséquent, lors du girage des résidents, cette répartition entre les catégories GIR 1 et GIR 2 est bien souvent arbitraire. Ceci peut expliquer pourquoi la proportion relative de résidents en GIR 1 n'a pas d'impact sur les coûts.

21. Ceci est probablement dû au fait qu'un établissement accueillant des patients Alzheimer au sein d'unités spécialisées (unités Cantou) reçoivent souvent des subventions départementales ou nationales, qui peuvent permettre de réduire le niveau des tarifs. Or les coûts que nous observons correspondent uniquement aux tarifs fixés par les établissements.



naturellement les coûts. En revanche le choix de l'option soins *globale* ne semble pas avoir d'impact sur les dépenses, alors que le panier de soins pris en charge par l'établissement est pourtant plus large ; les établissements ayant opté pour l'option soins *globale* dépenseraient ainsi à première vue moins que les établissements à option soins *partielle*.

Les variables de qualité ont l'effet attendu sur les dépenses des établissements : une moindre qualification du personnel a un effet négatif significatif, tandis qu'un ratio élevé d'écart à la norme du taux d'encadrement a un effet positif significatif. La non prise en compte de ces variables dans un nouveau mécanisme de régulation tarifaire pourrait avoir des effets néfastes en termes de qualité de prise en charge en établissement.

Afin de comparer les coefficients associés aux variables expliquant les termes d'inefficacité dans l'analyse stochastique de frontière, nous avons régressé par moindres carrés ordinaires les résidus obtenus par estimation quantile et les effets spécifiques aléatoires obtenus par estimations en panel sur les variables *STAT*, *BENEF* et *PIB*. La proportion de bénéficiaires a un impact significatif positif sur les coûts, et donc négatif sur l'inefficacité. Cela peut peut-être s'expliquer par un pouvoir de négociation accru des établissements accueillant des bénéficiaires à l'aide sociale vis à vis des autorités de tarification. Sur les estimations en coupe, le statut aurait peu d'effets sur l'inefficacité, alors que le statut public a un effet positif fortement significatif sur les coûts à partir des estimations en données de panel. Nous vérifions ce résultat par des tests paramétriques et non paramétriques présentés ci-dessous (Table 9).

Les résultats d'estimation de l'équation auxiliaire de Mundlak nous indique qu'il existe de l'hétérogénéité non observée corrélée aux variables explicatives, qui serait liée à la structure de qualification du personnel mais également aux variables de rémunération du personnel. Les résultats d'estimation en effets aléatoires (7) et (8) peuvent donc être biaisés.

TABLE 6 – Résultats d'estimation : Analyses stochastiques de frontière (SFA) et régressions quantiles (RQ) - Modèle 1

	(SFA - sans q)		(SFA - $\bar{q}$ )		(RQ - sans q)		(RQ - q)		(RQ - $\bar{q}$ )	
	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)
$\alpha_0$	2.897*	(1.165)	-4.657**	(0.986)	1.635	(1.474)	-0.899	(1.355)	-5.137†	(3.042)
Y	-0.498*	(0.226)	0.352	(0.191)	0.315	(0.284)	0.174	(0.269)	1.022†	(0.566)
$w_{AS}/INF$	0.629	(1.266)	1.373	(1.119)	3.583*	(1.613)	-0.193	(1.838)	0.027	(1.603)
$w_{ASH}/INF$	2.816*	(1.347)	2.596*	(1.201)	1.543	(1.909)	3.565*	(1.653)	3.250	(2.826)
$w_{S}^2/INF$	-1.679*	(0.770)	-2.013**	(0.650)	-2.087**	(0.750)	-1.922†	(0.991)	-1.476*	(1.445)
$w_{ASH}^2/INF$	0.970**	(0.336)	0.882**	(0.319)	0.996**	(0.351)	-0.329	(0.281)	-0.498†	(0.265)
$Y^2$	0.132**	(0.023)	0.089**	(0.019)	-0.012	(0.031)	0.117**	(0.028)	0.068*	(0.027)
$Y.w_{AS}/INF$	-0.058	(0.126)	-0.128	(0.112)	-0.345*	(0.168)	0.056	(0.181)	0.031	(0.159)
$Y.w_{ASH}/INF$	-0.189	(0.133)	-0.155	(0.119)	-0.112	(0.156)	-0.249	(0.185)	-0.297†	(0.163)
$w_{AS}/INF.w_{ASH}/INF$	0.148	(0.511)	0.575	(0.436)	1.121*	(0.506)	0.971	(0.653)	0.918†	(0.514)
PU1	0.055**	(0.015)	0.034**	(0.013)	-0.028	(0.019)	0.020	(0.019)	0.040*	(0.018)
OS	0.102	(0.200)	0.198	(0.175)	0.221	(0.196)	-0.067	(0.223)	0.099	(0.225)
Y.OS	-0.009	(0.020)	-0.019	(0.017)	-0.020	(0.019)	0.008	(0.022)	-0.008	(0.022)
PROP = propr.	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
PROP = loc.public	-0.010	(0.009)	-0.006	(0.008)	0.001	(0.010)	0.006	(0.011)	0.005	(0.011)
PROP = loc.HLM	0.021†	(0.012)	0.027*	(0.011)	0.031	(0.021)	0.037*	(0.015)	0.051**	(0.015)
PROP = loc.asso	0.034*	(0.016)	0.024	(0.015)	0.027	(0.018)	0.031	(0.019)	0.040*	(0.018)
PROP = loc.priv.BL	0.016	(0.017)	0.025	(0.016)	0.024**	(0.023)	0.026	(0.021)	0.062**	(0.020)
URB : $x < 20000$	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
URB : $20000 < x < 200000$	-0.001	(0.009)	0.005	(0.008)	0.030**	(0.010)	-0.021†	(0.011)	-0.008	(0.011)
URB : $x > 200000$	0.013	(0.009)	0.019*	(0.008)	0.045**	(0.010)	0.017	(0.011)	0.021*	(0.011)
URB : PARIS	0.115**	(0.016)	0.089**	(0.014)	0.057**	(0.018)	0.112**	(0.019)	0.084**	(0.018)
ALZ	-0.387*	(0.185)	-0.357*	(0.162)	-0.431*	(0.182)	-0.268	(0.217)	-0.080	(0.217)
Y.ALZ	0.042*	(0.018)	0.038*	(0.016)	0.043*	(0.018)	0.030	(0.021)	0.011	(0.021)
BATI	-0.001*	(0.000)	-0.001*	(0.000)	-0.001**	(0.000)	-0.0005*	(0.000)	-0.001*	(0.000)
propGIR2	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
propGIR1	0.019	(0.060)	0.021	(0.053)	0.006	(0.058)	0.058	(0.075)	0.027	(0.073)
propGIR3	-0.144*	(0.066)	-0.183**	(0.060)	-0.196*	(0.079)	-0.049	(0.085)	-0.114	(0.079)
propGIR4	-0.405**	(0.055)	-0.493**	(0.050)	-0.632**	(0.095)	-0.359**	(0.069)	-0.450**	(0.066)
propGIR5	-0.653**	(0.067)	-0.762**	(0.060)	-0.889**	(0.108)	-0.678**	(0.086)	-0.788**	(0.083)
propGIR6	-0.512**	(0.055)	-0.739**	(0.054)	-0.911**	(0.194)	-0.504**	(0.064)	-0.746**	(0.062)
NonQualif	-	-	-0.039**	(0.006)	-0.254**	(0.049)	-	-	-0.042**	(0.007)
N/N*	-	-	0.362**	(0.020)	1.002**	(0.203)	-	-	0.345**	(0.026)
$\delta_0$	-4.467	(2.880)	-3.260*	(1.507)	-4.107	(3.232)	-0.708**	(0.206)	-0.756**	(0.183)
STAT = Priv.BNL	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
STAT = Pub.non hosp.	0.024	(0.046)	-0.213†	(0.123)	0.092	(0.080)	0.002	(0.007)	-0.019**	(0.007)
STAT = Pub.hosp.	0.337	(0.228)	0.184*	(0.093)	0.416	(0.314)	0.079**	(0.028)	0.041	(0.025)
BENEF	0.193	(0.138)	0.127	(0.088)	0.217	(0.182)	0.033**	(0.013)	0.030**	(0.011)
PIB	0.404†	(0.245)	0.301*	(0.128)	0.363	(0.268)	0.079**	(0.020)	0.083**	(0.018)
$\sigma^2$	0.036†	(0.021)	0.033*	(0.015)	0.036	(0.023)	0.079**	(0.020)	0.083**	(0.018)
$\gamma$	0.683**	(0.195)	0.761**	(0.112)	0.693**	(0.204)	-	-	0.761*	(0.383)

† : 10% \* : 5% \*\* : 1%

TABLE 7 – Résultats d'estimation : Effets aléatoires (EA) et Effets individuels corrélés (MEIC) - Modèle 1

	(EA - sans q)		(EA - $\hat{q}$ )		(MEIC - sans q)		(MEIC - $\hat{q}$ )	
	(7)		(8)		(9)		(10)	
	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)
$\alpha_0$	1.037	(1.574)	-0.858	(1.401)	0.116	(1.815)	-2.786 <sup>†</sup>	(1.577)
$Y$	-0.139	(0.285)	0.164	(0.252)	-0.113	(0.298)	0.262	(0.262)
$w_{AS/INF}$	2.967	(1.813)	2.670	(1.686)	4.593*	(2.129)	3.838 <sup>†</sup>	(2.022)
$w_{ASH/INF}$	2.165	(1.723)	1.793	(1.597)	1.380	(2.130)	1.586	(2.028)
$w_{AS/INF}^2$	5.381**	(1.560)	5.196**	(1.467)	5.149**	(1.563)	5.598**	(1.456)
$w_{ASH/INF}^2$	-0.910	(0.782)	-0.381	(0.732)	-0.796	(0.786)	-0.347	(0.726)
$Y^2$	0.092**	(0.027)	0.067**	(0.024)	0.086**	(0.027)	0.052*	(0.024)
$Y.w_{AS/INF}$	-0.130	(0.174)	-0.119	(0.162)	-0.315	(0.206)	-0.252	(0.196)
$Y.w_{ASH/INF}$	-0.310 <sup>†</sup>	(0.163)	-0.227	(0.151)	-0.238	(0.206)	-0.236	(0.196)
$w_{AS/INF}.w_{ASH/INF}$	-1.499	(0.984)	-1.275	(0.922)	-2.069*	(1.019)	-2.265*	(0.953)
$PU1$	0.116**	(0.011)	0.103**	(0.010)	0.114**	(0.011)	0.099**	(0.010)
$URB : x < 20000$	ref	ref	ref	ref	ref	ref	ref	ref
$URB : 20000 < x < 200000$	-0.006	(0.013)	0.000	(0.011)	-0.004	(0.013)	0.003	(0.011)
$URB : x > 200000$	0.015	(0.013)	0.018	(0.011)	0.014	(0.013)	0.018 <sup>†</sup>	(0.011)
$URB : PARIS$	0.085**	(0.021)	0.056**	(0.018)	0.089**	(0.021)	0.045*	(0.018)
$ALZ$	-0.001	(0.185)	0.061	(0.173)	-0.042	(0.188)	0.010	(0.175)
$Y.ALZ$	0.005	(0.018)	-0.003	(0.017)	0.009	(0.018)	0.003	(0.017)
$propGIR2$	ref	ref	ref	ref	ref	ref	ref	ref
$propGIR1$	0.076	(0.071)	0.071	(0.066)	0.031	(0.095)	0.040	(0.090)
$propGIR3$	-0.081	(0.072)	-0.123 <sup>†</sup>	(0.067)	-0.163 <sup>†</sup>	(0.096)	-0.164 <sup>†</sup>	(0.092)
$propGIR4$	-0.261**	(0.060)	-0.327**	(0.055)	-0.309**	(0.083)	-0.356**	(0.080)
$propGIR5$	-0.508**	(0.069)	-0.637**	(0.064)	-0.395**	(0.099)	-0.425**	(0.097)
$propGIR6$	-0.554**	(0.061)	-0.737**	(0.058)	-0.594**	(0.090)	-0.681**	(0.090)
$NonQualif$	-	-	-0.035**	(0.006)	-	-	-0.038**	(0.008)
$N/N^*$	-	-	0.329**	(0.021)	-	-	0.179**	(0.030)
$\overline{Y}$					0.102	(0.146)	0.168	(0.131)
$\overline{w_{AS/INF}}$					-4.118	(3.606)	-3.211	(3.197)
$\overline{w_{ASH/INF}}$					1.883	(3.355)	-0.237	(2.997)
$\overline{propGIR1}$					0.079	(0.145)	0.053	(0.130)
$\overline{propGIR3}$					0.156	(0.148)	0.043	(0.133)
$\overline{propGIR4}$					0.108	(0.123)	0.089	(0.112)
$\overline{propGIR5}$					-0.228	(0.139)	-0.361**	(0.129)
$\overline{propGIR6}$					0.119	(0.124)	-0.054	(0.117)
$\overline{Y.w_{AS/INF}}$					-0.145	(0.329)	0.077	(0.294)
$\overline{Y.w_{ASH/INF}}$					0.465	(0.357)	0.376	(0.316)
$\overline{w_{AS/INF}.w_{ASH/INF}}$					1.835*	(0.807)	1.673*	(0.716)
$\overline{NonQualif}$					-	-	-0.001	(0.011)
$\overline{N/N^*}$					-	-	0.270**	(0.041)
$\delta_0$	0.054	(0.196)	0.504**	(0.135)	0.126	(0.198)	0.617**	(0.134)
$STAT = Priv.BNL$	ref	ref	ref	ref	ref	ref	ref	ref
$STAT = Pub.non hosp..$	0.056**	(0.007)	0.024**	(0.005)	0.048**	(0.007)	0.015**	(0.005)
$STAT = Pub.hosp..$	0.080**	(0.008)	0.045**	(0.005)	0.075**	(0.008)	0.039**	(0.005)
$BENEF$	0.066**	(0.010)	0.042**	(0.007)	0.066**	(0.011)	0.038**	(0.007)
$P1B$	0.123**	(0.019)	0.075**	(0.013)	0.118**	(0.020)	0.065**	(0.013)

Seuils de significativité : † : 10% \* : 5% \*\* : 1%

**Scores d'efficacité.** Nous présentons à la Table 8 les statistiques descriptives des termes d'inefficacité obtenus par les différentes méthodes d'estimation selon la catégorie institutionnelle de l'établissement. Ces termes sont respectivement associés au terme résiduel asymétrique  $u_e$  en analyse stochastique de frontière, à la perturbation en régression quantile, et à l'effet spécifique aléatoire pour les méthodes en données de panel. A la vue des scores obtenus sur les estimations en coupe, entre 5 à 10% des coûts seraient dûs à de l'inefficacité pour les établissements privés associatifs et publics non rattachés à une structure hospitalière, i.e. ne sont expliqués par aucune variable introduite dans le modèle. Ce taux s'élève à plus de 15% pour les établissements publics hospitaliers.

La variabilité des scores d'inefficacité obtenus par analyse stochastique de frontière est relativement faible pour les établissements privés associatifs et publics non rattachés à une structure hospitalière, mais la variance des termes résiduels et des effets spécifiques, obtenus respectivement par régression quantile et par données de panel, suggèrent cependant qu'une part non négligeable de la variabilité des coûts reste non expliquée par les variables intégrées dans notre modélisation. Quant aux établissements publics rattachés à une structure hospitalière, quelque soit la méthode d'estimation utilisée, nous observons une forte dispersion de

leurs termes d'inefficacité. Des gains non négligeables en efficacité pourraient probablement être réalisés par certains EHPAD au regard de ces résultats.

L'introduction de variables de qualité (estimation (3),(6),et (10)) modifie légèrement l'inefficacité moyenne des établissements de chacune des catégories institutionnelles. Si les niveaux moyens et médians d'inefficacité des établissements privés associatifs et des établissements publics non hospitaliers semblent quasiment identiques lorsque la qualité n'est pas prise en compte, cette inefficacité devient relativement plus importante pour les EHPAD publics lorsque les coûts sont épurés de la qualité. Ce résultat est confirmé par des tests paramétrique (comparaison de moyenne) et non paramétrique (test de Kruskal-Wallis<sup>22</sup>) effectués sur les scores d'inefficacité regroupés par statut (Table 9) : lorsque la qualité est prise en compte, les différences d'inefficacité entre les établissements de divers statuts sont statistiquement significatives : les établissements publics seraient alors moins efficaces que les établissements privés à but non lucratif. Ce constat est surprenant dans la mesure où les établissements privés ont en moyenne des coûts journaliers par résident plus élevés que les établissements publics non rattachés à une structure hospitalière. Une fois ces coûts épurés de l'hétérogénéité liée à la qualité et à l'environnement d'activité des établissements, ils deviennent relativement plus faibles pour les établissements privés associatifs. Ceci confirme l'idée de Newhouse (1970)[24] selon laquelle les établissements privés à but non lucratif pourraient être moins efficaces en termes de coûts, et ce en raison de niveaux de qualité relativement élevés.

Le fait d'intégrer des variables de qualité dans le modèle estimé ne modifie cependant que très peu le classement des établissements selon leur plus ou moins grande inefficacité. Nous avons calculé des coefficients de corrélation de Spearman<sup>23</sup> entre les termes d'efficacité estimés via un modèle sans variable de qualité et via un modèle avec variable de qualité instrumentée ; ces coefficients nous indiquent en effet une forte dépendance statistique significative des termes d'efficacité, qu'il s'agisse de termes obtenus par analyse stochastique de frontière (coefficient=0.9613) ou par régression quantile (coefficient=0.9645). Cela peut s'interpréter comme une influence faible des variables de qualité sur l'inefficacité : les établissements les plus (moins) efficaces ne seraient pas ceux proposant le moins (plus) de qualité. Il convient néanmoins d'être prudent dans cette interprétation, puisque les variables de qualité introduites sont des variables instrumentées et n'intègrent donc qu'une partie de l'hétérogénéité liée à la qualité. La corrélation entre les termes d'efficacité obtenus par estimation avec effets individuels aléatoires corrélés de modèles de données de panel excluant et intégrant des variables de qualité (non instrumentée) montre en effet une corrélation moins parfaite, avec un coefficient de corrélation de Spearman de 0.83 qui reste toutefois élevé. Afin d'approfondir notre analyse, nous représentons graphiquement (Figure 3) les corrélations entre les scores d'efficacité obtenus par analyse stochastique de frontière d'un modèle sans variable de qualité (estimation 1) et les variables de qualité suivantes :  $N/N^*$ ,  $NonQualif$ , et une variable de qualité agrégée  $N/N^*.(1/NonQualif)$ . Il ne semble pas y avoir de corrélation entre la qualité et les scores d'efficacité. Pour le vérifier, nous estimons plusieurs coefficients de Spearman entre la variable de qualité agrégée  $N/N^*.(1/NonQualif)$  et les scores d'efficacité, selon le statut de l'établissement et selon son ratio d'écart à la norme du taux d'encadrement. Nous présentons les résultats à la Table 10. La corrélation entre la qualité et les scores d'efficacité pour l'ensemble des établissements est significative mais faible (coefficient de Spearman de 0.323). Celle-ci est d'autant plus faible pour les établissements publics ; cela signifie que les inefficacités-coûts observées sont moins souvent expliquées par l'encadrement en personnel dans le secteur public que dans le secteur privé. Nous remarquons également que le lien entre les coûts non justifiés et la qualité est moins fort lorsque les établissements ont un faible taux d'encadrement en personnel relativement à leurs besoins. Nous pouvons apporter deux interprétations à ce résultat. Les variances de la variable de qualité et des termes d'inefficacité sont d'une part plus faibles pour cette catégorie d'établissements ; ceci peut donc en partie expliquer l'absence d'une nette corrélation statistique entre les deux variables. Il est, d'autre part, tout à fait probable que les établissements ayant un taux d'encadrement insuffisant aient également une plus grande inefficacité, compte tenu des surcoûts qu'un manque de personnel peut engendrer : turn-over du personnel, absentéisme, recours à l'intérim, etc...

22. Le test de Kruskal-Wallis compare la moyenne des rangs des observations de 2 échantillons ; il s'agit test non paramétrique puisque aucune hypothèse n'est faite sur la forme de la distribution.

23. Le coefficient de corrélation de Spearman est une mesure non paramétrique de la dépendance statistique entre 2 variables ; il s'agit d'un coefficient de corrélation de Pearson calculé sur les rangs des observations associés à chacune des variables.

TABLE 8 – Comparaison des scores d'inefficacité estimés

Méthode d'estimation	Privé Associatif				Public non hosp.				Public hosp.			
	Moy.	1 <sup>er</sup> q.	Méd.	3 <sup>eme</sup> q.	Moy.	1 <sup>er</sup> q.	Méd.	3 <sup>eme</sup> q.	Moy.	1 <sup>er</sup> q.	Méd.	3 <sup>eme</sup> q.
SFA (1)	1.07	1.04	1.05	1.07	1.06	1.04	1.06	1.07	1.15	1.08	1.13	1.22
SFA (2)	1.09	1.05	1.07	1.011	1.05	1.04	1.05	1.06	1.15	1.07	1.16	1.21
SFA (3)	1.06	1.04	1.05	1.07	1.07	1.05	1.06	1.08	1.16	1.09	1.14	1.25
RQ (4)	0.10	0.01	0.09	0.18	0.10	0.01	0.08	0.17	0.18	0.07	0.20	0.29
RQ (5)	0.09	0.01	0.09	0.16	0.07	0.01	0.07	0.13	0.13	0.03	0.14	0.24
RQ(6)	0.08	0.00	0.08	0.17	0.09	0.02	0.09	0.17	0.18	0.07	0.18	0.19
EA (7)	1.31	1.25	1.31	1.37	1.37	1.3	1.36	1.42	1.39	1.3	1.38	1.46
EA(8)	1.27	1.22	1.27	1.32	1.3	1.25	1.3	1.33	1.32	1.26	1.31	1.37
EIC (9)	1.33	1.26	1.33	1.39	1.38	1.31	1.38	1.44	1.41	1.32	1.39	1.48
EIC (10)	1.29	1.23	1.29	1.33	1.3	1.26	1.3	1.34	1.33	1.26	1.31	1.37

TABLE 9 – Tests paramétrique et non paramétrique sur les termes d'inefficacités

	Test de comparaison de moyenne		Test de Kruskal-Wallis	
	p-value	Catégorie la moins efficace ?	p-value	Catégorie la moins efficace ?
<b>SFA (1) : sans var q</b>				
Privé asso. vs Public non hosp.	0.4652	-	0.0886	Public non hosp.
Privé asso. vs Public hosp.	0.0002	Public hosp.	0.0001	Public hosp.
Public non hosp. vs hosp.	0.0002	Public hosp.	0.0001	Public hosp.
<b>SFA (3) : avec var q</b>				
Privé asso. vs Public non hosp.	0.0007	Public non hosp.	0.0001	Public non hosp.
Privé asso. vs Public hosp.	0.0001	Public hosp.	0.0001	Public hosp.
Public non hosp. vs hosp.	0.0002	Public hosp.	0.0001	Public hosp.
<b>R.Q. (1) : sans var q</b>				
Privé asso. vs Public non hosp.	0.8528	-	0.9616	-
Privé asso. vs Public hosp.	0.0365	Public hosp.	0.0116	Public hosp.
Public non hosp. vs hosp.	0.0330	Public hosp.	0.0078	Public hosp.
<b>R.Q. (3) : avec var q</b>				
Privé asso. vs Public non hosp.	0.0911	Public non hosp.	0.0598	Public non hosp.
Privé asso. vs Public hosp.	0.0223	Public hosp.	0.0051	Public hosp.
Public non hosp. vs hosp.	0.0434	Public hosp.	0.0119	Public hosp.
<b>EIC (1) : sans var q</b>				
Privé asso. vs Public non hosp.	0.0000	Public non hosp.	0.0017	Public non hosp.
Privé asso. vs Public hosp.	0.0000	Public hosp.	0.0001	Public hosp.
Public non hosp. vs hosp.	0.0000	Public hosp.	0.0001	Public hosp.
<b>EIC (3) : avec var q</b>				
Privé asso. vs Public non hosp.	0.0006	Public non hosp.	0.0001	Public non hosp.
Privé asso. vs Public hosp.	0.0000	Public hosp.	0.0001	Public hosp.
Public non hosp. vs hosp.	0.0000	Public hosp.	0.0053	Public hosp.

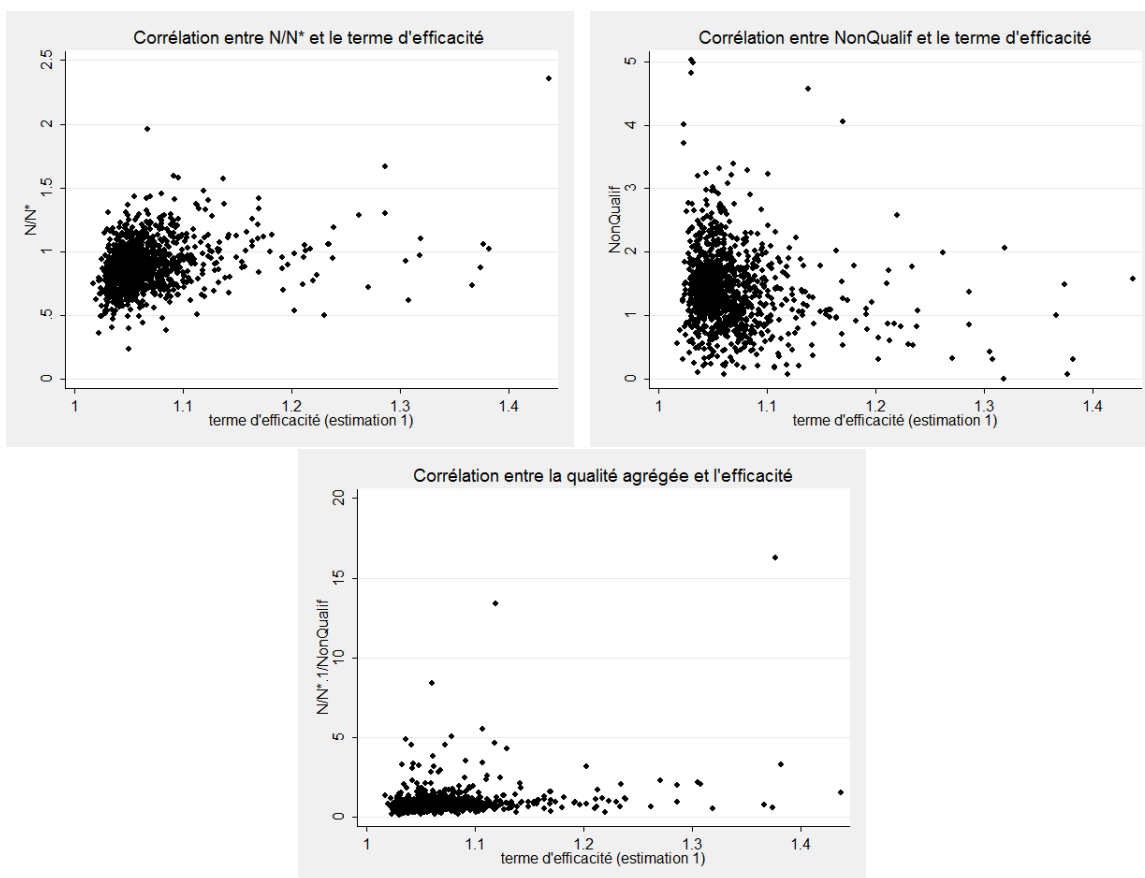


FIGURE 3 – Corrélations entre les scores d'efficacité obtenus par l'estimation (1) et les variables de qualité

TABLE 10 – Coefficients de corrélation de Spearman entre les scores d'efficacité  $u$  de l'estimation (1) et la variable de qualité  $q$  agrégée

$N/N$	<i>Privé Asso.</i>	<i>Public non hosp.</i>	<i>Public hosp.</i>	Total	% d'EHPAD	$\sigma_q$	$\sigma_u$
$N/N^* < 0.8$	0.378**	0.063	-	0.299**	32.4%	0.036	0.66
$0.8 \leq N/N^* < 1$	0.298**	0.174**	0.548	0.245**	46.5%	0.035	0.78
$N/N^* \geq 1$	0.456**	0.294**	0.083	0.323**	21.1%	0.059	1.06
Total	0.365**	0.267**	0.298	0.322**	100%	0.042	0.82

**Etude des rendements d'échelle.** Grâce au théorème de la dualité, nous pouvons analyser les rendements d'échelle directement à partir de la fonction de coût, en étudiant la relation entre l'output  $Y$  et le coût total  $CT$ .

L'élasticité d'échelle  $EE$  se définit comme l'inverse de la hausse proportionnelle des coûts résultant d'une hausse proportionnelle de l'output, soit :

$$EE = \frac{\Delta Y/Y}{\Delta CT/CT} = \frac{1}{\frac{\partial \ln CT}{\partial \ln Y}} \quad (6)$$

Il y a présence d'économies d'échelle lorsque l'élasticité d'échelle  $EE > 1$ , i.e. lorsque les coûts peuvent être réduits si l'établissement produit plus. A l'inverse, si  $EE < 1$ , il y a des déséconomies d'échelle, l'établissement gagnerait à produire en moindre quantité. Enfin, si  $EE = 1$ , l'échelle de production est optimale.

A partir des estimations que l'on a obtenues pour les estimations (1), (4) (9) et (10)<sup>24</sup>, nous calculons l'élasticité d'échelle de la manière suivante :

$$\frac{1}{EE} = \frac{\partial \ln CT / w_{INF}}{\partial \ln Y} = \frac{\partial \ln CT}{\partial \ln Y} = \alpha_Y + \alpha_{Y^2} Y + \alpha_{Y.w_{AS/INF}} w_{AS/INF} + \alpha_{Y.w_{ASH/INF}} w_{ASH/INF} + \alpha_{Y.ALZ} ALZ + \alpha_{Y.OS} OS$$

en posant  $w_{AS/INF}$ ,  $w_{ASH/INF}$  égales à leurs valeurs moyennes, et  $ALZ$ ,  $OS$  égales à leur valeur dominante (i.e. 0).

Nous présentons ci-dessous les élasticités d'échelle et les coûts moyens estimés<sup>25</sup> pour différents niveaux d'output (Table 11). Les établissements disposant d'un nombre de lits inférieur à 80 ont des opportunités d'économies d'échelle non réalisées. Ceux ayant un nombre de lits supérieurs à 80 sont à l'inverse en déséconomies d'échelle. Nous obtenons donc une taille critique des établissements supérieure à celle observée par Dervaux et al.(2006)[9]. Les coûts moyens par résident forment ainsi une *courbe en U*, avec un minimum correspondant au nombre moyen de journées-résidents des établissements (28425 journées).

TABLE 11 – Etude des rendements d'échelle

Méthode d'estimation		50% moy.	75% moy.	Moyenne	125% moy.	150% moy.
SFA sans var q	<i>CM</i>	75.6	73.4	72	71.4	71.4
	<i>EE</i>	1.14	1.09	1.05	1.02	0.99
RQ sans var q	<i>CM</i>	67.9	66.7	66.1	66.2	66.5
	<i>EE</i>	1.09	1.05	1.01	0.99	0.97
MEIC sans var q	<i>CM</i>	71.9	70.9	56.1	71.2	71.6
	<i>EE</i>	1.66	1.03	1	0.98	0.97
MEIC avec var q	<i>CM</i>	69.3	68.6	56.5	68.6	68.8
	<i>EE</i>	1.03	1.01	1	0.99	0.98
<i>Nb de lits correspondants</i>		40	60	80	100	120

## 6.2 Etude des économies de gamme par type d'hébergement et par degré de dépendance des résidents : modèle 2

L'activité des établissements n'est pas homogène et peut varier selon le type d'hébergement proposé aux résidents (hébergement permanent, temporaire ou accueil de jour), et selon le degré de dépendance des personnes accueillies. Nous présentons ci-dessous (Tables 12 et 13) les proportions relatives de chacune de ces catégories d'output.

TABLE 12 – Proportions des types d'hébergement en maison de retraite

	N	% Heb. permanent	% Heb. temporaire	% Accueil de jour
EHPAD avec heb. permanent uniquement	978	100%	-	-
EHPAD avec heb. permanent + heb. temporaire	86	98%	2%	-
EHPAD avec heb. permanent + accueil de jour	54	98.7%	-	1.3%
EHPAD proposant les 3 modes d'heb.	53	94.3%	3.9%	1.8%

24. Afin d'obtenir des rendements d'échelle robustes, nous préférons utiliser les résultats d'estimations des modèles n'utilisant pas d'instrumentation des variables de qualité.

25. L'estimation des coûts moyens s'effectue à partir des valeurs moyennes de toutes les variables continues autres que Y et des valeurs dominantes des variables discrètes.

TABLE 13 – Proportions moyennes de catégories GIR des résidents accueillis

	<b>Dépendance forte</b>	<b>Dépendance moyenne</b>	<b>Dépendance faible</b>
	<b>N</b>	<b>(GIR 1-2)</b>	<b>(GIR 3-4)</b>
			<b>(GIR 5-6)</b>
EHPAD sans résidents en GIR 5 et 6	25	70.2%	29.8%
EHPAD avec toutes les catégories de résidents	1146	47.5%	35.3%
			17.2%

La majorité des établissements de notre échantillon pratiquent uniquement l'accueil permanent des résidents, et ceux qui proposent des modes d'hébergement alternatifs semblent le faire à la marge de leur activité principale d'hébergement permanent<sup>26</sup>. La plupart des établissements accueillent en revanche des résidents de chaque catégorie de dépendance, mais le nombre de journées-résidents correspondant aux personnes de GIR 1 et 2 est toutefois beaucoup plus important.

Nous présentons (Table 14) les résultats d'estimation du modèle hybride par analyse de frontière stochastique et par régression quantile<sup>27</sup>. Selon la décomposition effectuée,  $Y_1$  caractérise le nombre de journées en hébergement permanent ou le nombre de journées de résidents en GIR 1 ou 2,  $Y_2$  l'activité en accueil de jour ou les journées-résidents des GIR 3 et 4,  $Y_3$  le nombre de journées en hébergement temporaire ou les journées-résidents des GIR 5 et 6. Les coefficients négatifs associés aux variables d'interaction des différentes catégories d'outputs suggèrent la présence d'économies de gamme. Il serait ainsi plus avantageux de "produire" plusieurs de ces outputs plutôt qu'un seul. Ce résultat est confirmé par le calcul des économies de gamme effectué à partir de ces résultats d'estimation.

26. Il existe quelques structures spécialisées dans l'accueil de jour et/ou dans l'hébergement temporaire. Par souci de comparabilité, nous avons choisi de ne pas les intégrer dans notre échantillon et de ne conserver que les maisons de retraite.

27. Nous préférons ne pas intégrer de variables de qualité dans les modèles estimés afin d'éviter le biais d'hétérogénéité probablement engendré par l'introduction d'une variable de qualité endogène, ou le bruit ajouté par l'introduction d'une variable de qualité instrumentée.



TABLE 14 – Résultats d'estimation : Analyse stochastique de frontière et Régression quantile - Modèle 2

	(Décomposition par type d'hébergement)				(Décomposition par GIR)			
	(11)		(12)		(13)		(14)	
	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)
$\alpha_0$	3.489**	(0.043)	3.364**	(0.050)	3.186**	(0.057)	2.980**	(0.045)
$Y_1$	0.469**	(0.008)	0.515**	(0.010)	0.613**	(0.042)	0.699**	(0.025)
$Y_2$	1.692*	(0.824)	0.452	(0.838)	0.539**	(0.035)	0.673**	(0.037)
$Y_3$	1.113**	(0.421)	1.059†	(0.542)	0.509**	(0.070)	0.552**	(0.054)
$Y_1^2$	-0.019**	(0.001)	-0.028**	(0.001)	-0.016*	(0.008)	-0.026**	(0.006)
$Y_2^2$	0.879	(1.872)	-1.985	(2.276)	-0.036	(0.039)	-0.054**	(0.020)
$Y_3^2$	0.294	(0.662)	0.114	(0.649)	0.030†	(0.016)	-0.022	(0.027)
$Y_1.Y_2$	-0.310*	(0.136)	0.113	(0.118)	-0.080*	(0.040)	-0.132**	(0.021)
$Y_1.Y_3$	-0.171†	(0.104)	-0.115	(0.167)	-0.192**	(0.036)	-0.196**	(0.027)
$Y_2.Y_3$	-22.057**	(1.366)	-26.776**	(7.070)	-0.130*	(0.060)	-0.143**	(0.042)
$Y_1.Y_2.Y_3$	5.680**	(0.718)	6.519**	(1.736)	0.048**	(0.010)	0.074**	(0.007)
$w_{AS/INF}$	0.506**	(0.057)	0.523**	(0.072)	0.476**	(0.099)	0.441**	(0.100)
$w_{ASH/INF}$	0.375**	(0.058)	0.431**	(0.065)	0.353**	(0.059)	0.449**	(0.094)
$PUI$	0.032†	(0.017)	-0.008	(0.019)	0.041	(0.026)	0.002	(0.028)
$OS$	-0.008	(0.021)	-0.046†	(0.025)	0.008	(0.022)	0.034	(0.037)
$Y_1.OS$	0.005	(0.007)	0.020*	(0.008)	-0.002	(0.025)	0.007	(0.024)
$Y_2.OS$	-0.245	(0.378)	-0.181	(0.299)	0.040	(0.027)	0.012	(0.037)
$Y_3.OS$	0.569	(0.457)	0.246	(0.319)	-0.098*	(0.045)	-0.087*	(0.042)
$PROP = propr.$	ref	ref	ref	ref	ref	ref	ref	ref
$PROP = loc.public$	-0.011	(0.011)	-0.007	(0.011)	-0.014	(0.010)	-0.004	(0.017)
$PROP = loc.HLM$	0.026†	(0.014)	0.034*	(0.015)	0.029	(0.019)	0.036	(0.023)
$PROP = loc.asso$	0.037*	(0.018)	0.024	(0.020)	0.028	(0.018)	0.009	(0.030)
$PROP = loc.priv.BL$	0.013	(0.019)	0.046*	(0.021)	0.007	(0.054)	0.025	(0.031)
$URB : x < 20000$	ref	ref	ref	ref	ref	ref	ref	ref
$URB : 20000 < x < 200000$	-0.008	(0.011)	-0.015	(0.012)	-0.007	(0.010)	-0.013	(0.017)
$URB : x > 200000$	0.021*	(0.010)	0.021†	(0.011)	0.013	(0.014)	0.013	(0.016)
$URB : x < 20000$	0.119**	(0.018)	0.119**	(0.019)	0.113**	(0.017)	0.104**	(0.029)
$ALZ$	0.042*	(0.019)	0.027	(0.022)	0.010	(0.043)	0.038	(0.036)
$Y_1.ALZ$	-0.003	(0.006)	0.003	(0.007)	0.016	(0.024)	0.009	(0.024)
$Y_2.ALZ$	-0.163	(0.265)	-0.199	(0.297)	-0.002	(0.023)	-0.018	(0.032)
$Y_3.ALZ$	-0.366	(0.688)	-0.129	(0.477)	0.020	(0.054)	0.019	(0.040)
$BATI$	-0.001**	(0.000)	-0.001*	(0.000)	-0.001*	(0.000)	-0.001	(0.000)
$propGIR2$	ref	ref	ref	ref				
$propGIR1$	0.027	(0.066)	0.102	(0.077)				
$propGIR3$	-0.180**	(0.074)	-0.067	(0.089)				
$propGIR4$	-0.401**	(0.062)	-0.355**	(0.071)				
$propGIR5$	-0.642**	(0.075)	-0.612**	(0.090)				
$propGIR6$	-0.542**	(0.061)	-0.502**	(0.069)				

Significance levels : † : 10% \* : 5% \*\* : 1%

**Économies de gamme.** Il y a économie de gamme lorsqu'il est plus avantageux de produire plusieurs types d'outputs simultanément que de les produire seuls. A l'instar de Baumol et al.(1982)[3], nous calculons les économies de gamme  $EG$  de la manière suivante :

$$EG = \frac{\sum_i C(Y_i) - C(Y)}{C(Y)}$$

Nous calculons la baisse (ou la hausse) des coûts due à la production jointe de  $Y_{HP}$ ,  $Y_{HT}$  et  $Y_{AJ}$ , et celle due à la production jointe de  $Y_{GIR12}$ ,  $Y_{GIR34}$  et  $Y_{GIR56}$ .  $EG$  est alors positive en cas d'économies de gamme, ou négative en cas de déséconomies de gamme.

Nous présentons ci-dessous (Table 15) les résultats trouvés pour les estimations d'économies de gamme totales et spécifiques à chaque catégorie d'output<sup>28</sup> :

28. Ces estimations ont été réalisées à partir des valeurs moyennes (ou des valeurs dominantes pour les variables discrètes) de toutes les variables autre que les outputs.

TABLE 15 – Economies de gamme totales et spécifiques

Décomposition par type d'hébergement	Décomposition par GIR			
	SFA	RQ	SFA	RQ
$(Y_{HP})+(Y_{AJ})+(Y_{HT})$	0.61	0.57	$(Y_{GIR12})+(Y_{GIR34})+(Y_{GIR56})$	0.42 0.31
$(Y_{HP})+(Y_{AJ} + Y_{HT})$	0.30	0.28	$(Y_{GIR12})+(Y_{GIR34} + Y_{GIR56})$	0.16 0.10
$(Y_{AJ})+(Y_{HP} + Y_{HT})$	0.31	0.29	$(Y_{GIR34})+(Y_{GIR12} + Y_{GIR56})$	0.15 0.09
$(Y_{HT})+(Y_{HP} + Y_{AJ})$	0.30	0.29	$(Y_{GIR56})+(Y_{GIR12} + Y_{GIR34})$	0.26 0.20

Nous trouvons des économies de gammes totales et spécifiques pour chacun de ces outputs. Il ne serait donc pas pertinent en termes de coûts de se spécialiser dans un seul type d'hébergement ou dans une seule catégorie de dépendance. On remarque néanmoins que les économies de gamme spécifiques aux résidents de GIR 3 et 4 sont relativement faibles ; cela signifie que le coût d'un résident en GIR 3 ou 4 ne varie que très peu entre les établissements accueillant uniquement ce type de résidents et ceux hébergeant une clientèle plus diversifiée.

**Economies d'échelles spécifiques.** Toujours à partir des résultats d'estimation présentés à la Table 14, nous calculons les élasticités d'échelle spécifiques pour chacun des types d'hébergement et chacune des catégories de résidents accueillis.

Il faut pour cela calculer le *coût incrémental moyen* pour chacun des outputs et pour différents niveaux de ces outputs. Le *coût incrémental* représente la différence entre le coût associé à la production de tous les outputs  $j$  ( $j \neq i$ ) à leur niveau moyen  $\bar{Y}_j$  plus  $Y_i$  avec une certaine quantité choisie, et le coût de production des outputs sans  $Y_i$ . Pour obtenir un coût incrémental *moyen*, nous divisons cette différence par  $Y_i$ .

$$CMI_i = \frac{C(\bar{Y}_j, Y_i) - C(Y_i)}{Y_i}$$

Les élasticités d'échelle spécifiques par produit *EEsp* (Baumol et al., 1982[3]) sont alors calculées à partir de ces coûts moyens incrémentaux et de l'estimation des coûts marginaux :

$$ESsp_i = \frac{CM_i}{Cm_i}$$

TABLE 16 – Economies d'échelle spécifiques par type d'hébergement

	Héb. permanent			Accueil de jour			Héb. temporaire		
	Cm	CM	EEsp	Cm	CM	EEsp	Cm	CM	EEsp
<b>méthode SFA</b>									
<b>75% moy.</b>	60.1	43.7	0.73	151.4	155	1.02	112.7	122.5	1.09
<b>moyenne</b>	72.8	49.4	0.68	156	153.6	0.98	123.9	122	0.98
<b>méthode RQ</b>									
<b>75% moy.</b>	58.2	43	0.74	87.3	132.8	1.52	114.3	134.3	1.17
<b>moyenne</b>	68.5	48	0.70	131.1	132.5	1.01	135	134.4	0.99

TABLE 17 – Economies d'échelle spécifiques par degré de dépendance des résidents accueillis

	GIR 1-2			GIR 3-4			GIR 5-6		
	Cm	CM	EEsp	Cm	CM	EEsp	Cm	CM	EEsp
<b>méthode SFA</b>									
<b>75% moy.</b>	74.6	65	0.87	60.1	62.5	1.04	43.7	38.1	0.87
<b>moyenne</b>	88.3	69.5	0.79	68.6	63.7	0.93	44	39.3	0.89
<b>méthode RQ</b>									
<b>75% moy.</b>	73.8	63.1	0.86	62.6	65	1.04	41.2	42.2	1.03
<b>moyenne</b>	88	67.4	0.77	70.5	66.2	0.94	42.7	42.3	0.99

On observe ainsi la présence d'économies d'échelle spécifiques pour l'hébergement temporaire et l'accueil de jour, pour un niveau d'activité correspondant à 75% de l'activité moyenne. Ceci peut aisément s'expliquer par le fait qu'une plus grande taille permet de mieux utiliser ses ressources grâce à un turn-over important

de résidents. De plus petites structures font face à des variations plus importantes de demande en proportion ; elles ont alors besoin d'un slack de personnel et de lits en cas de pics de demande et ont ainsi plus de ressources inutilisées. A l'inverse, nous mettons en évidence des déséconomies d'échelle pour l'hébergement permanent à tous les niveaux, qui peuvent s'expliquer par la présence d'économies de gammes et l'avantage à ne pas produire une quantité trop grande de journées-résidents avec un type particulier d'hébergement.

Concernant la décomposition par groupe de résidents, l'accueil de résidents de GIR 3 et 4 semble le plus propice aux économies d'échelle. L'hébergement des catégories de résidents à forte et à faible dépendance présente à l'inverse des déséconomies d'échelle spécifiques élevées ; ceci suggère à nouveau qu'il n'est pas optimal de se spécialiser dans ces catégories de résidents et qu'il est plus rentable d'accueillir des résidents présentant des degrés de dépendance divers.

**Economies d'échelles globales.** Nous étudions une nouvelle fois les économies d'échelle globales des établissements, à partir de ces résultats d'estimation. Les élasticités d'échelle sont calculées pour différentes valeurs de l'output total, en conservant une proportion fixe<sup>29</sup> des catégories d'output<sup>30</sup>. Les élasticités calculées à partir des estimations par analyse stochastique de frontière sont présentées ci-dessous (Table 18). Nous retrouvons une taille critique des établissements à environ 80 lits, ce qui correspond à la taille moyenne des EHPAD en France. Cela confirme la robustesse de notre résultat.

TABLE 18 – Etude des rendements d'échelle globales (estimées à partir du modèle 2)

Décomposition :		50% moy.	75% moy.	Moyenne	125% moy.	150% moy.
par type d'hébergement	<i>CM</i>	82.3	73	71.4	73	76.2
	<i>EE</i>	1.69	1.2	0.98	1.85	0.77
par GIR	<i>CM</i>	83.5	75.3	73.8	74.6	76.3
	<i>EE</i>	1.55	1.17	1	0.91	0.87
<i>Nb de lits correspondants</i>		<i>40</i>	<i>60</i>	<i>80</i>	<i>100</i>	<i>120</i>

## 7 Discussion et conclusion

Cette étude effectuée sur un large échantillon de maisons de retraite en France apporte plusieurs informations sur les facteurs de coûts de ces établissements. La robustesse des résultats a été testée par l'utilisation de diverses méthodes d'estimation.

Concernant l'effet de la catégorie institutionnelle, bien que les établissements publics non rattachés à une structure hospitalière aient des coûts par résident moins élevés que les établissements privés associatifs, ils paraissent en moyenne moins efficaces après contrôle par la qualité et les facteurs de coûts exogènes. Les différences sont d'autant plus marquées pour les établissements publics rattachés à des hôpitaux, pour lesquels les coûts non justifiés et liés à de l'inefficacité sont en moyenne de 15%.

La taille de l'établissement a également un impact non négligeable sur ses coûts. Nous observons une taille optimale de 80 lits, ce qui correspond à la capacité moyenne des maisons de retraite en France. L'optimalité n'a cependant été définie qu'en termes de minimisation de coûts. Si, dans notre optimisation, nous avons pris en compte la dégradation du bien-être des résidents lorsque ceux-ci sont hébergés dans des grandes structures, qui peuvent en outre être géographiquement éloignées de leur ancien habitat, la taille optimale aurait probablement été réduite. Cela explique pourquoi certains acteurs du secteur insistent sur l'importance de conserver des structures à *taille familiale*<sup>31</sup>.

29. Nous conservons les proportions moyennes de chacune des catégories d'output.

30. Ces élasticités d'échelles sont calculées comme suit :

$$ES = \frac{CM}{Cm} = \frac{CM}{\sum_i Cm_i * Y_i}$$

31. Certains schémas gérontologiques départementaux, à l'instar de celui de la Vendée, mentionnent ainsi des tailles critiques largement inférieures à 80 lits : "*établissement à taille humaine de 24 places*"[8].

Outre sa taille, la spécialisation d'un établissement dans un seul type d'hébergement ou l'accueil de résidents d'un seul niveau de dépendance semblent avoir un effet positif sur les coûts. Nous observons ainsi des économies de gamme pour les divers types d'hébergement (hébergement permanent, temporaire, et accueil de jour), et également pour l'accueil de résidents de différents degrés de dépendance (GIR 1 à 6). Cela signifie qu'il n'est pas rentable économiquement pour une institution de se spécialiser dans une seule catégorie d'activité.

Les différences en termes de qualité de prise en charge, que l'on évalue par la structure de qualification du personnel et par un ratio d'encadrement effectif sur encadrement nécessaire compte tenu des besoins spécifiques des résidents, n'expliqueraient qu'une faible partie de l'hétérogénéité des coûts entre établissements. Ces variables de qualité ont certes un impact positif significatif sur les coûts pour toutes les estimations effectuées. Néanmoins, les scores d'efficacité obtenus par analyse stochastique de frontière et par régression quantile du modèle de coût n'intégrant pas la qualité ne sont que très faiblement corrélés aux variables de qualité. Les établissements les plus efficaces ne sont donc pas ceux qui ont un fort encadrement en personnel soignant et, inversement, les établissements jugés comme étant les moins coûteux relativement à leurs besoins spécifiques n'ont pas un encadrement plus faible. Les scores d'inefficacité obtenus seraient alors le reflet d'une mauvaise gestion des établissements. Compte tenu de ce résultat, l'abandon du système de financement par prix de journée et l'adoption d'un nouveau mécanisme de tarification à l'activité pourraient être justifiés : cela permettrait d'inciter certains EHPAD à réduire leurs coûts et engendrerait de ce fait une plus grande équité entre les établissements. Il convient néanmoins d'être prudent dans une telle interprétation des résultats, et ce à plusieurs titres.

D'une part, les scores d'efficacité obtenus sont fonction des variables de contrôle introduites dans notre modélisation. Or nous n'avons pas pu prendre en compte la totalité des spécificités de coûts des établissements. Nous n'avons par exemple pas intégré le fait que quelques établissements puissent bénéficier d'une mise sous disponibilité gratuite du bâtiment par leur commune de résidence. Les coûts calculés comme le produit des tarifs journaliers par les journées-résidents ne représentent peut-être pas parfaitement la totalité des charges d'exploitation et d'investissement des EHPAD. Certains établissements peuvent recevoir ponctuellement des subventions départementales ou nationales, ce que nous ne prenons pas en compte dans notre étude. Pour toutes ces raisons, les scores d'inefficacité peuvent être *sur* ou *sous*-estimés.

D'autre part, nous avons restreint la qualité à l'encadrement en personnel permanent et à la part du personnel soignant sur cet encadrement. Mais il existe bien évidemment d'autres dimensions de la qualité qui peuvent expliquer une partie des coûts et qui, de ce fait, se retrouvent dans les scores d'inefficacité. Le recrutement ponctuel de personnel d'animation, l'organisation d'activités ou de sorties, la qualité des repas augmentent les coûts des établissements et ne sont pas pris en compte dans notre modélisation. Nous ne légitimons donc pas une réforme de la tarification qui consisterait à aligner tous les établissements sur les coûts de celui estimé comme étant le plus efficace, et qui ne prendrait pas en compte tous ces surcoûts nécessaires liés à la qualité.

Cette étude montre néanmoins le besoin d'une plus grande équité financière entre les EHPAD, compte tenu des hétérogénéités de coûts non justifiées par des différences de personnel. Elle montre également la nécessité d'une hausse des taux d'encadrement et d'une réduction de leur disparité : 78.9% des établissements étudiés ont un ratio d'encadrement effectif sur encadrement théorique<sup>32</sup> inférieur à 1. Et nous observons que la plupart de ces établissements présentent des termes d'inefficacité importants et non corrélés aux taux d'encadrement. Des défauts de gestion peuvent certes être en partie responsables de ces inefficacités, mais l'observation d'une plus faible corrélation de l'inefficacité avec la qualité lorsque les établissements ont un faible taux d'encadrement suggère qu'un manque de personnel puisse être générateur de surcoûts. Un encadrement insuffisant peut en effet engendrer des inefficacités liées à un fort absentéisme du personnel, un recours accru à l'intérim, des hospitalisations plus fréquentes, un turn-over important et donc des recrutements plus nombreux. Dans cette optique, une réforme de la tarification ne pourra permettre de gagner en efficacité que si des normes d'encadrement minimal sont adoptées par le secteur. Sans l'imposition de telles normes, il est à craindre que le passage d'un financement prospectif à une tarification à l'activité engendre non pas une réduction de l'inefficacité des établissements mais une réduction de leur qualité.

---

32. L'encadrement théorique est défini au regard des normes d'encadrement existantes pour l'aide à domicile et inscrites dans le Plan Solidarité Grand Age (2007-2012).

## 8 Annexes

**Estimation par variables instrumentales.** Nous avons choisi d'instrumenter la variable d'écart à la norme du taux d'encadrement  $N/N^*$  et la variable de basse qualification  $NonQualif$  par l'ancienneté du personnel. Nous utilisons les doubles moindres carrés, en projetant en première étape la variable  $N/N^*$  et la variable  $NonQualif$  sur les variables non endogènes du modèle et sur les instruments  $ANC$  et  $ANC^2$ . Nous présentons ci-dessous les équations d'instrumentation et les résultats d'estimation :

$$\begin{aligned}
N/N^* = & \alpha_0 + \alpha_Y \ln(Y) + \sum_{i,i \neq 1} \alpha_{w_i} \ln\left(\frac{w_i}{w_1}\right) + \frac{1}{2} \alpha_{Y^2} (\ln Y)^2 + \sum_{i,i \neq 1} \frac{1}{2} \alpha_{w_i^2} \left(\ln \frac{w_i}{w_1}\right)^2 \\
& + \sum_{i,i \neq 1} \alpha_{Y.w_i} (\ln Y) \left(\ln \frac{w_i}{w_1}\right) + \sum_{i,i \neq 1} \sum_{j,j \neq i} \alpha_{w_i.w_j} \left(\ln \frac{w_i}{w_1}\right) \left(\ln \frac{w_j}{w_1}\right) \\
& + \alpha_G G + \sum_{k^1} \alpha_{z_{k^1}} z_{k^1} + \sum_{k^2} \alpha_{z_{k^2}} z_{k^2} + \sum_{k^2} \alpha_{Y.z_{k^2}} Y.z_{k^2} \\
& + \alpha_{ANC} ANC + \alpha_{ANC^2} ANC^2 + \epsilon
\end{aligned} \tag{7}$$

$$\begin{aligned}
NonQualif = & \alpha_0 + \alpha_Y \ln(Y) + \sum_{i,i \neq 1} \alpha_{w_i} \ln\left(\frac{w_i}{w_1}\right) + \frac{1}{2} \alpha_{Y^2} (\ln Y)^2 + \sum_{i,i \neq 1} \frac{1}{2} \alpha_{w_i^2} \left(\ln \frac{w_i}{w_1}\right)^2 \\
& + \sum_{i,i \neq 1} \alpha_{Y.w_i} (\ln Y) \left(\ln \frac{w_i}{w_1}\right) + \sum_{i,i \neq 1} \sum_{j,j \neq i} \alpha_{w_i.w_j} \left(\ln \frac{w_i}{w_1}\right) \left(\ln \frac{w_j}{w_1}\right) \\
& + \alpha_G G + \sum_{k^1} \alpha_{z_{k^1}} z_{k^1} + \sum_{k^2} \alpha_{z_{k^2}} z_{k^2} + \sum_{k^2} \alpha_{Y.z_{k^2}} Y.z_{k^2} \\
& + \alpha_{ANC} ANC + \alpha_{ANC^2} ANC^2 + \epsilon
\end{aligned} \tag{8}$$

Nous effectuons un test de Sargan pour contrôler l'exogénéité des instruments  $ANC$ ,  $ANC^2$ . On ajoute à ces instruments 3 autres instruments exclus : pourcentages de résidents incontinents dans l'établissement, de résidents ayant effectué au moins une chute dans l'année, et ceux souffrant de dépression<sup>33</sup>. Nous acceptons l'hypothèse nulle de non endogénéité avec une p-value de 0.7118.

Les tests de Fisher de significativité globale des instruments effectués à partir des résultats d'estimation de chacune de ces équations d'instrumentation permettent d'écarter l'hypothèse nulle d'instruments faibles :  $H_0 : \alpha_{ANC} = \alpha_{ANC^2} = 0$ . On obtient en effet :

-pour la première équation d'instrumentation (projection de  $N/N^*$ ) :  $F(2, 1141) = 6.26$ , et une p-value de 0.002.

-pour la deuxième équation (projection de  $NonQualif$ ) :  $F(2, 1141) = 9.49$ , et une p-value de 0.0001.

Nous effectuons enfin un test d'Hausman pour s'assurer de l'endogénéité de  $N/N^*$ . Nous rejettons l'hypothèse nulle d'exogénéité de  $N/N^*$  et de  $NonQualif$  avec une p-value égale à 0. Afin de vérifier notre hypothèse de non endogénéité des variables de qualité au delà d'un certain seuil (que l'on fixe à  $N/N^*=1$ ), nous effectuons 2 tests d'Hausman : un test sur les établissements ayant un écart à la norme du taux d'encadrement inférieur à 1, et un autre test sur les observations telles que  $N/N^* \geq 1$ . Nous confirmons ainsi notre hypothèse :

-si  $N/N^* < 1$  : p-value=0;

-si  $N/N^* \geq 1$  : p-value=0.1570.

---

33. Ces 3 variables d'état de santé des résidents sont des proxy de qualité satisfaisants, malheureusement nous ne disposons de ces informations que pour un échantillon de 231 établissements.

TABLE 19 – Equations d’instrumentation

	$N/N^*$		$NonQualif$	
$Y$	-1.021**	(0.325)	1.597	(1.107)
$w_{AS/INF}$	-4.106*	(1.928)	-4.606	(6.577)
$w_{ASH/INF}$	1.502	(2.060)	4.636	(7.027)
$w_{AS/INF}^2$	0.731	(1.027)	1.858	(3.504)
$w_{ASH/INF}^2$	-0.551	(0.450)	-1.680	(1.534)
$Y^2$	0.101**	(0.032)	-0.168	(0.110)
$Y.w_{AS/INF}$	0.462*	(0.191)	0.670	(0.653)
$Y.w_{ASH/INF}$	-0.218	(0.203)	-0.513	(0.693)
$w_{AS/INF}.w_{ASH/INF}$	-0.278	(0.680)	2.367	(2.321)
$PUI$	0.026	(0.019)	-0.222**	(0.066)
$PROP = propr.$	ref	ref	ref	ref
$PROP = loc.public$	0.007	(0.012)	0.057	(0.041)
$PROP = loc.HLM$	-0.061**	(0.017)	-0.184**	(0.057)
$PROP = loc.asso$	-0.030	(0.021)	-0.150*	(0.071)
$PROP = loc.priv.BL$	-0.076**	(0.023)	-0.048	(0.077)
$URB : x < 20000$	ref	ref	ref	ref
$URB : 20000 < x < 200000$	-0.021†	(0.013)	0.023	(0.043)
$URB : x > 200000$	-0.022†	(0.012)	0.057	(0.041)
$URB : PARIS$	0.043*	(0.021)	-0.031	(0.071)
$ALZ$	0.002	(0.250)	-0.206	(0.854)
$Y.ALZ$	0.001	(0.025)	0.014	(0.084)
$OS$	-0.122	(0.268)	-0.017	(0.913)
$Y.OS$	0.013	(0.026)	0.006	(0.090)
$BATI$	0.000	(0.000)	0.000	(0.001)
$propGIR2$	ref	ref	ref	ref
$propGIR1$	0.024	(0.080)	-0.083	(0.275)
$propGIR3$	0.180*	(0.090)	0.460	(0.308)
$propGIR4$	0.354**	(0.075)	0.451†	(0.255)
$propGIR5$	0.419**	(0.091)	0.516†	(0.311)
$propGIR6$	0.819**	(0.074)	1.572**	(0.252)
$ANC$	0.008	(0.007)	0.095**	(0.024)
$ANC^2$	-0.001*	(0.000)	-0.005**	(0.001)
$\alpha_0$	5.757**	(1.682)	-6.702	(5.737)

Seuils de significativité : † : 10% \* : 5% \*\* : 1%

## Références

- [1] D Aigner, C. Lovell, and P. Schmidt. Formulation and estimation of stochastic frontier production models. *Journal of Econometrics*, 6 :21–27, 1977.
- [2] G. Battese and T. Coelli. A model for technical efficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20 :325–332, 1995.
- [3] W. Baumol, J. Panzar, and R. 1982 Willig. *Contestable markets and the theory of industry structure*. 1982.
- [4] C. Branchu, J. Voisin, J. Guedj, D. Lacaze, and S. Paul. Etat des lieux relatif à la composition des coûts mis à la charge des résidents des établissements d’hébergement pour personnes âgées dépendantes. 2009.
- [5] L. Chritensen, D. Jorgenson, and L. Lau. Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1) :28–45, 1973.
- [6] T. Coelli. Estimators and hypothesis tests for a stochastic frontier function : A monte-carlo analysis. *Journal of Productivity Analysis*, 6(3) :247–268, 1995.
- [7] Fédération Hospitalière de France et Syndicat National de la Gérontologie Clinique. Sécurité et qualité dans le secteur personnes âgées : produire des indicateurs de qualité associés à des indicateurs de résultats. 2009.
- [8] Conseil Général de Vendée. Plan "Bien vieillir en Vendée, 2010-2014". 2010.
- [9] B. Dervaux, H. Leleu, H. Nogues, and V. Valdmanis. Assessing French nursing home efficiency : An indirect approach via budget-constrained DEA models. *Socio-Economic Planning Sciences*, 40 :70–91, 2006.
- [10] A. Dor. The costs of Medicare patients in nursing homes in the United States : a multiple output analysis. *Journal of Health Economics*, 8 :253–270, 1989.

- [11] DREES. La vie en établissement d'hébergement pour personnes âgées du point de vue des résidents et de leurs proches. *Dossiers Solidarité et Santé*, 18, 2011.
- [12] M. Farsi, M. Filippini, and M. Kuenzleb. Unobserved heterogeneity in stochastic cost frontier models : an application to Swiss nursing homes. *Applied Economics*, 37 :2127–2141, 2005.
- [13] M. Filippini. Cost and scale efficiency in the nursing home sector : Evidence from Switzerland. *Quaderno*, N.99-02, 1999.
- [14] T. Grannemann, R. Brown, and M. Pauly. Estimating hospital costs : A multiple-output analysis. *Journal of Health Economics*, 5, 1986.
- [15] J. Jondrow, D. Knox Lovell, I. Masterov, and P. Schmidt. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19 :233–238, 1982.
- [16] K. Knox, E. Blankmeyer, and J. Stutzman. Technical efficiency in Texas nursing facilities : a stochastic production frontier approach. *Journal of economics and finance*, 31(1), 2007.
- [17] R. Koenker and G. Basset. Regression quantiles. *Econometrica*, 46(1), 1978.
- [18] S.C. Kumbhakar and C.A.K. Lovell. *Stochastic frontier analysis*. Cambridge University Press, 2000.
- [19] C. Liu, A. Laporte, and B. Ferguson. The quantile regression approach to efficiency measurement : Insights from monte-carlo simulations. *Health Economics*, 17, 2008.
- [20] V. Létard, H. Flandre, and S. Lepeltier. La France et les Français face à la canicule : les leçons d'une crise. 2004.
- [21] N. McKay. An econometric analysis of costs and scale economies in the nursing home industry. *The Journal of Human Resources*, 23(1) :57–75, 1988.
- [22] W. Meeusen and J. Van Den Broeck. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2), 1977.
- [23] R. Mutter, M. Rosko, and H. Wong. Measuring hospital inefficiency : The effects of controlling for quality and patient burden of illness. *Health Services Research*, 6 :1992–2013, 2008.
- [24] J.P. Newhouse. Toward a theory of nonprofit institutions : an economic model of a hospital. *American Economic Review*, 60(1) :64–74, 1970.
- [25] J.P. Newhouse. Frontier estimation : how useful a tool for health economics? *Journal of Health Economics*, 13(3) :317–322, 1994.
- [26] W. Niskanen. Bureaucracy and representative government. *Chicago, Aldine-Atherton.*, 1971.
- [27] M. Pitt and L. Lee. The measurement and sources of technical inefficiency in the industrial weaving industry. *Journal of Development Economics*, 9(1), 1981.
- [28] E. Ratte and D. Imbaud. Groupe de réflexion sur la dépendance : « accueil et accompagnement des personnes âgées en perte d'autonomie ». 2011.
- [29] W. Réjault. *Maman, est-ce que ta chambre te plaît? Survivre en maison de retraite*. Broché, 2009.
- [30] P. Schmidt and R. Sickles. Production frontier and panel data. *Journal of Business and Economic Statistics*, 2 :367–374, 1984.
- [31] K. Spilsbury, C. Hewitt, L. Stirk, and C; Bowman. The relationship between nurse staffing and quality of care in nursing homes : A systematic review. *International Journal of Nursing Studies*, 2011.
- [32] UNIOPSS. Secteur social et médico-social : les agences régionales de santé et l'appel à projet. Mars 2010.
- [33] D. Vitaliano and M. Toren. Cost and efficiency in nursing homes : a stochastic frontier approach. *Journal of Health Economics*, 13(3), 1994.
- [34] H. Wang and P. Schmidt. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, 18 :129144, 2002.