Quality of service and cost-efficiency of French nursing homes.

Brigitte DORMONT, Cécile MARTIN

Abstract

This paper examines cost-efficiency of French nursing homes, in the context of a possible reform of the current retrospective payment system. We have at our disposal cross-sectional data relative to 1,171 private non-profit and public nursing homes in 2007, with information about costs, staff composition and the dependency level of residents. Observations for years 2003 and 2007 are available for a substantial sub-sample of 740 nursing homes, which makes panel data estimations possible. Quality of service is measured by the gap between a norm for the nurse/resident ratio, and the actual level. In our sample, this indicator is correlated with incontinence, falls and depression. We use a stochastic frontier analysis, a quantile regression approach and an error component model to estimate a translog cost function. Depending on estimation methods, inefficiency raises costs by 2% to 11% for public and private non-profit nursing homes, but the extra cost is higher, 7% to 19% for public nursing homes that are associated to hospitals. Taking the quality of service into account leads to a slight improvement in estimated efficiency. Private non profit nursing homes appear to be significantly more efficient than public nursing homes.

1 Introduction

Public spending due to dependence issue amounts to 1.7% of the GDP in France in 2011. And this percentage would probably increase because of the the rising proportion of aged people in the population. Paying for these needs calls for reforms of the long-term care sector. Various measures are taken and planned by french public authorities to contain costs of long-term care facilities (nursing homes, but also long stay units and special residential homes). One of them is a reform of price regulation. Inspired by recent changes in hospital sector, a coming transition from a cost-reimbursement to a prospective-payment system based on the degree of dependency of residents is schedulded. The underlying idea of this reform is to reduce the non-explained cost heterogeneity between institutions. Indeed we observe a huge per resident per daily cost variability according to their institutional form (public, private non-profit and profit sectors) and their spatial location. Some associations are strongly opposing this change because it involves some risks of patients selection if their characteristics are not well taken into account in fixed prices. This reform could also generate a decline of quality, which would be already too low because of insufficient resources to several institutions. Public authorities try to achieve two objectives: decreasing costs of nursing homes and improving quality of care for the dependent elderly. Unfortunately, the negative correlation that seems to exist between them would imply that a choice would have to be made between quality or efficiency. We test empirically whether this idea is confirmed.

A comprehensive overview of the sector is needed before the implementation of this reform in order to predict its possible effects. We have a sample of 1171 homes for 2007, and a sample of 740 nursing homes in 2003 and 2007. We use these samples to estimate with frontier parametric methods (frontier stochastic analysis, quantile regression and random effects estimations) the part of cost heterogeneity due to characteristics of the patient base (age, degree of dependence, diseases), quality of care and the part due to technical and allocative inefficiencies of nursing homes. We define nursing homes quality by two variables: a ratio of observed staff on recommended staff (*a nurse resident ratio gap*), and a ratio of low skilled on skilled staff (*a low skilled staff ratio*). Endogeneity is corrected by an instrumental variables method. Quality seems to have an effect on costs, but a substantial part of the cost variability remains unexplained by the model. After taking quality into account, public nursing homes would be less efficient than associative nursing homes.

2 Context

2.1 Nursing homes institutional form

Nursing homes with different institutional forms share the long-term care market for dependent elderly. In 2007[1], 30% of french nursing homes are private non-profit *(Priv.NP)* organizations. 22% belongs to the private for-profit

[1]Source: EHPA 2007 survey

sector; we do not study here these nursing homes because we have no data on them. Public nursing homes represent 48% of the market, and are either hospital based *(Pub.HB)* (18%) or non-hospital based institutions *(Pub.NHB)* (30%). In the early 20^{th} century, some beds were reserved for elderly patients. This quota of beds has grown into units for frail elderly, and then into legally dependent institutions of the hospital but may be geographically distant. Today these nursing homes operate identically to the non-hospital based ones, and they cater for elderly with the same dependency levels. However, they are relatively bigger institutions and aged people come more often after an hospital stay. According to a recent report of the General Inspection of Social Affairs [3], some cost and quality differences seem to exist between hospital based and non-hospital based institutions. For all of these reasons, we have splitted up public nursing homes into these two categories. Porportions of each of these institutional forms are highly variable depending on the geographical location: for instance, the ratio of public nursing homes ranges from 20% in some local subdivisions (called *departement*) to almost 90% in 2007. This suggests some endogeneity of the status.

Whatever institutional form nursing homes are, they face the same legal constraints. Especially, creation or extension of a nursing home must first be accepted by public authorities. Thus the number of beds is an exogenous variable for any institution. The occupancy rate is always close to 100%. Public nursing homes employ mostly state employees (civil servants)[2], whose wages are fixed by an official wage scale. When they are not state employees, staff have less likely temporary contrats in the public sector than in the private sector[3]. Social allowances for institutinalized aged people are the same regardless the nursing home status, to the extent that prices are regulated (which is the case of most public and *PNP* nursing homes, far more rarely in the private for-profit sector).

2.2 The pricing reform project

For now the cost-based payment mechanism allows regulated nursing homes to set 3 day rates equal to their current average costs to their 3 main activities: housing, nursing care and medical care. Thus, this is a retrospective pricing. Operating costs and investment charges are splitted up into these three categories by an officialy specified way. For each of these rates, residents are more ou less subsidized. For the medical care rate, everybody if fully covered by the National Health Insurance. Concerning the price of nursing care, all of the institutionalized aged people get the Personal Independance Allowance (PIA); its level is contingent upon their degree of dependency and the price of the nursing home, it is a cost-sharing system. If their resources are insufficient to pay the housing fees, residents may benefit from a second social allowance: the Health Area Social Allowance. The eligibility threshold for the latter depends on the *departement* of residence of beneficiaries before they were institutionalized.

Public authorities in each *departement* fix 3 nursing care rates and 3 medical care rates, according to the dependancy degree of residents.[4]

With such a payment system, long-term care organizations have no incentives for cost-containment. Another drawback is the historical inequity of nursing homes budgets. This could probably explain a part of the per resident average cost variability among institutions (Figure 1), in particular regarding on the institutional form. Average cost is higher and more variable for *Pub.HB* nursing homes. This daily cost is almost the same for the two first quartiles of *PNP* and *Pub.NHB* nursing homes, however, it is higher for the two last quartiles of *PNP* organizations.

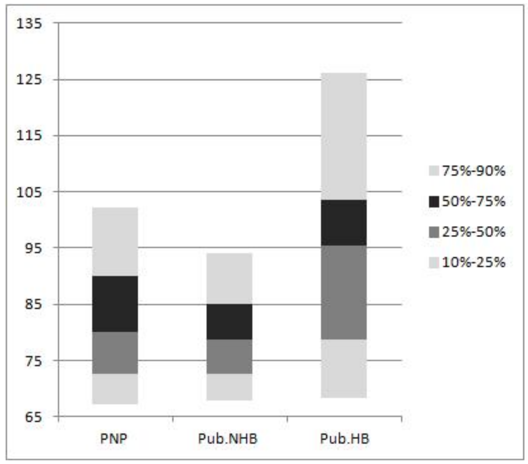
To reduce this cost variability and to provide more incentives for cost-efficiency, a pricing reform project with a transition to a per case prospective payment system is suggested. Nursing homes would be funded according to degrees of dependency and diseases of their residents. But this kind of payment mechanism involves some usual risks - of a decline of quality and patients selection - which slow down the implementation of the pricing reform. There is an additionnal issue with this reform because the housing price is suggested to remain a cost-based one. Thus what we expect is a transfer of expenses on this pricing category if nursing care and medical care fixed rates are not perfectly suitable, and this could generate some affordability problems. Yet, what residents and their families have to pay is already quite high (2200 euros per month on average according to the General Inspection of Social Affairs [3]). This could probably explain why aged people are entering increasingly later in nursing homes. They prefer to opt for solutions of formal or informal care until their disabilities do not allow them anymore. The aim of our study is to analyze the possible consequences of such a pricing reform, by examining cost differences and their causes between nursing homes with several empirical frontier methods. We sum up in Table 1 the current pricing mechanism and the reform project for each of the pricing categories.

[2]Acoording to the EHPA 2007 survey, the part of civil servant nurses is on average 75,6% in *Pub.NHB* nursing homes and 87,5% in *Pub.HB* nursing homes.

[3]According to the EHPA 2007 survey, the part of nurses with temporary contracts is on average 9,4% in public nursing homes compared to 88,5% in *PNP* nursing homes.

[4]The notion of *dependence* and its measure are defined by a national scale, that allows to classify aged people in 6 resource utilization group *(Groupes Iso-Ressources, GIR)*, contingents on activities they managed to do themselves. More dependent residents are classified in *GIR 1*, whereas the more independant of them are in *GIR 6*.

Figure 1: Average daily cost per resident in nursing homes (in current euros)



Sample of 3339 nursing homes, source: EHPA 2007

Table 1: Summarization: the current payment system and the reform project for regulated nursing homes

	Current payment system	**Suggested reform**	**Who pays?**
Housing price	Cost-based payment	Cost-based payment	Social Allowance *-if insufficient resources*
Nursing care price **GIR 1 and 2, 3 and 4,** **5 and 6**	Cost-based payment	Case-based prospective payment	Personal Independance Allowance *-depending on the dependence degree* *-depending on nursing home prices*
Medical care price **GIR 1 and 2, 3 and 4,** **5 and 6**	Cost-based payment	Case-based prospective payment	National Health Insurance

3 Literature review

Cost functions and inefficiencies in nursing homes have been studied for years. Thus, from the beginning of the 80's, some cost factors are highlighted with ordinary least squares (OLS) methods (MacKay 1988 [18], Dor 1989 [6]). However, the use of more specific efficiency analysis methods to study this sector is relatively recent. Filippini (1999) [10] et Farsi et al.(2005) [8] have studied efficiency of swiss nursing homes in this way, with several parametric methods: panel data model fitted to efficiency analysis , stochastic and determinist frontier methods. To test the robustness of their results, some authors have also resorted to quantile regressions, in the manner of Knox et al.(2007) [13] .

The aim of these studies is to identify the main cost factors of nursing homes so as to adjust the case-based payment mechanism in an appropriate way. However, few studies have examined in details the impact of quality on costs. Some authors like Farsi et al. (2005) [8] integrate quality variables in their cost model, but without analyzing the potential endogeneity or even the effects of such a variable on estimated efficiency terms. Looking for quality variables and dealing with their endogeneous issue is an original contribution of our paper. In general terms, missing variables can bias the results because nursing homes considered as the more inefficient ones would provide in fact a higher level of quality. This is all the more true for the long-term care sector in which quality of care varies according to staff availability for residents. And wages bill represent about 80% of costs in nursing homes. So, looking for efficiency incentives in this sector is a completely different issue than in the hospital one. We can suppose that each gain in "efficiency" would be due to staff cuts or skill mix changes and, so, to a quality reduction. Staff ratios seem to be already low in France - 5.7 per 10 residents (National Report on Dependance Issue, 2011 [23]) - compared to other European countries like Germany where the average staff ratio in nursing homes is 12 (Létard et al., 2004 [17]). Recent books which recount experience of some french nursing homes employees describe adverse effects of staff shortage on quality of care: diaper wearing by non incontinent aged people, use of sleeping pills, malnutrition; *"Staff ratios in nursing homes are well below those of hospitals. Of course people need fewer treatments, but we should be capable of ensuring a sufficient relational quality (...) To work today in a nursing home without depression risk, we need to belittle residents as furniture"* (Wiliam Réjault, 2009) [24]. If efficiency gains can only be obtained with a cost-quality trade-off, the relevance of the pricing reform in France would be limited. Thus, the question is that of a possible reduction in charges unrelated to the exogenous costs characteristics of quality of care.

Some authors compared the impacts of institutional form on nursing homes costs and efficiency. With a Texan nursing homes sample from 1999 to 2002, Knox et al.(2007) [13] observed that private non-profit institutions were less efficient relative to private for-profit nursing homes. Nevertheless, few empirical studies have compared efficiency of public and private non-profit nursing homes. This is the issue that we are interested in. There is no consensus among theoritical analysis on the higher efficiency of one institutional sector with respect to the other. According to Newhouse (1970) [20], when the competition level is low enough, the resource allocation in private non-profit health organization is often not optimal because of too high quality of care: *"there is a bias against producing lower quality products"*. Conversely, theories of public economics allocate different objective functions depending on the institution status: most of them assign to the public sector budget balance or size maximization objectives, but not profit maximization peculiar to the private sector. According to these theories, public nursing homes would be less cost-efficient than *PNP* nursing homes. In this mind, Filippini (1999) [10] found a higher efficiency of private non-profit nursing homes compared to public institutions. We test in this paper if the same observation can be made in France. To do that, we compare estimate results obtained with regressions of cost model integrating and excluding quality variables. The underlying idea is that quality would be expensive. We examine whether the observation of lower costs for a nursing homes status category is or not characterized by a higher quality of provided care.

4 Methods

4.1 Model

It is assumed that nursing homes are transforming some inputs (several staff categories L_i, *i=nurses, nursing auxiliaries and support staff*) into resident-days of permanent, temporary stays and day care (output Y). Quality of care q is included in the model as an output characteristic. We also made the assumption that several categorical variables ($z = \sum_k z_k$) [5] which characterize some cost specificities of the nursing home (variables related to the building, patients with Alzheimer disease), as well as variables related on degrees of dependency of residents G may affect the production of resident-days. Thus, production technology of a nursing home may be written as follows: $F(Y, q, L, G, z, u) = 0$, with u: random technical inefficiency of production. Thanks to the dual properties of production functions, alternatively we can explain the production technology by the total cost function: $TC = C(Y, q, w, G, z, v)$, with w_i the average wage per staff category i for the nursing home, and v a random cost. We prefer to use a cost function for this approach because its estimation requires as explanatory variables, not the quantities of inputs, but their prices, probably less endogenous.

To be consistent with microeconomic theory of the producer, the cost function must check various properties that we look after estimation: growth, concave and homogenous of degree 1 with respect to the amount of outpout and prices of inputs.

We choose a translog function (Christensen et al. 1973 [4]) which consists in approximating the cost function by a Taylor limited development at order 2. This functional form is more flexible than *Cobb-Douglas* or *Constant Elasticity of Substitution (CES)* forms. Our cost model is the following:

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

Standardisation of costs and wages by the price of one of the factors of production guarantees the homogeneity of degree 1 of the total cost with respect to input prices.

[5] z_k can be divided into two categories: z_{k^1} is a vector of variables assumed to directly affect costs but not the relationship between output and costs, unlike z_{k^2} that can affect this relationship output-costs.

4.2 Estimation methods

Stochastic frontier analysis. We conduct a first step stochastic frontier analysis with cross-sectional data. The chosen specification breaks down the error in two components: one corresponding to the inefficiency of the nursing home e (u_e), and the other to a statistical noise (v_e). Different parametric assumptions of u_e and v_e distributions allow to identify these two components.

$$lnTC_e = lnTC(X_e) + u_e + v_e$$

with $u_e \geq 0$, TC_e the dependent variable associated with the nursing home e, and X_e a vector of explanatory variables for this nursing home e.

Then:

$$TC_e = TC(X_e) * exp(v_e) * exp(u_e)$$
$$et\ CE_e = \frac{TC_e}{TC(X_e).exp(v_e)} = exp(u_e) \geq 1$$

with CE_e the cost-efficiency term for the nursing home e.

If the nursing home e is fully efficient, then u_e=0 and $CE_e = 1$. (u_e) and efficiency terms CE_e are estimated using the formulation suggested by Jondrow et al.(1982) [12]

This method of maximum likelihood estimation, suggested in 1977 by Aigner et al. [1] and Meeusen and Van den Broeck [19], requires strong assumptions about functional form of the cost model and distribution of the efficiency term. Thus the shape of the likelihood function, and therefore estimate results, will depend on the distribution of the efficiency term u_e (Newhouse, 1994 [21]). We opt for a distribution of the disturbance normal- truncated normal, with $v_e \sim iidN(0, \sigma_v^2)$, $u_e \sim iidN^+(\mu, \sigma_u^2)$ and $\mu = \delta m_e + \eta_e$ (m: explanatory environmental variables of inefficiency, and η: statistical noise). And we use *a one-step approach* (Battese et Coelli, 1995 [2]) for which the inefficiency term is replaced in the cost-function by its explanatory variables. We could have used another approach found in the literature (Pitt and Lee 1981 [22], Vitaliano and Toren 1994 [27]), which consists of estimating first-stage stochastic frontier cost and inefficiency scores assuming they are independently and identically distributed, and then regress these scores on various explanatory variables. The second step is just to contradict the assumption of independant inefficiency terms. A two-step approach can lead to inefficient estimates (Battese et Coelli, 1995 [2]), or biaised estimates of inefficiency scores (Wang et Schmidt 2002 [28]).

Given the constraints on the shape of the cost function and inefficiency term distribution that need to be made for this stochastic frontier estimation, it is important to examine the robustness of estimate results and inefficiency terms. So, we confront these results with those of estimation methods that do not use these strong assumptions about distribution of error term: quantile regression and panel data estimates.

Quantile regression. We use a quantile regression method proposed by Koenker and Basset in 1978 [14]. Assuming that the k^{th} percentile represents efficient nursing homes, the quantile regression QR(k) allows an estimate of the cost frontier. This estimate is even closer to the *true* boundary since it is less sensitive to the presence of outliers relative to a method of maximum likelihood.

If this method allows to avoid any assumptions about the distribution of the inefficiency term, it needs a choice of the percentile at which nursing homes will be considered as efficient. Like Liu et al. (2008) [16] that hold the 80^{th} percentile to estimate a production frontier, we choose the 20^{th} percentile to estimate the cost frontier. Residuals obtained from this estimate can then be understood as inefficiency scores.

Error component model We then perform an analysis with panel data in order to get rid of the assumption on distribution of error term, but also now on non-correlation of inefficiency term with regressors (Schmidt and Sickles, 1984 [25]). This assumption of non-correlation is probably too strong for a cost function analysis, as various environmental factors can affect both costs and characteristics of patient base, nursing homes size,... This is especially true for quality variables introduced q, strongly linked to the amount of authorized expenses and, thus, to the extent of bargaining power of the institution with respect to public authorities. Under the assumption that these quality variables q are correlated with individual specific effect u_e and not with disturbance v_{et}, i.e. they are correlated with time-constant unobserved heterogeneity, using a random effects model with correlated individual effects allows to deal with this endogeneity. Taking into account the unobserved heterogeneity in the calculation of efficiency scores, which includes some aspects of quality difficult to measure, is another big advantage to the use of panel data. Since we only have two years of observations, it is not relevant to opt for a fixed effects estimation because individual specific effects would be not consistent. So we choose to run a random effects regression, which was adapted to the efficiency analysis by Schmidt and Sickles (1984) [25]:

$$lnTC_{et} = \beta_0 + \beta_X lnX_{et} + v_{et} + u_e$$

The error component model allow to distinguish two components from the error term: a statistical noise v_e and a random constant term u_e. We make the assumption that the inefficiency term u_e^* is also temporally constant and, thus, may be identified from u_e. The nursing home e for which u_e is minimum is considered to be fully efficient: $e, u_e^* = min_e[u_e] = 0$.

This estimate of β by generalized least squares (GLS) leads to biased results (heterogeneity bias) if inefficiency terms are correlated with regressors. To overcome this difficulty, we use a model with correlated individual effects (model based on Mundlak's specification, 1978) adapted to the model of Schmidt and Sickles by Farsi et al. (2005) [8]. Then $u_e = \delta_{X_e} \overline{lnX_e} + u_e^{'}$ and $\overline{lnX_e} = \frac{1}{T}\sum_t lnX_{et}$. Thus, inefficiency terms u_e^* are obtained from $u_e^{'}$.

We add for regression methods with cross-sectional data (stochastic frontier analysis and quantile regression) an instrumental variables estimation [6] to address the endogeneity issue of quality variables.

The use of different estimation methods allows us to examine the robustness of estimate results and inefficiency scores. We do not expect to obtain identical terms of inefficiency, the absolute value of this term is difficult to interpret: *"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 [11]). However, we study how the variability of these inefficiency terms is important, and if institutional form and quality of care affect efficiency. These impacts are studied by parametric and nonparametric tests (Kruskal-Wallis). By comparing these methods, we are trying to show the robustness of these effects.

4.3 Variables

When we want to study efficiency of agents of a particular economic sector, it is essential to fully understand all its specificities in order to have a production function (or a cost function with dual properties) that best characterizes the production process. If the complexity of the production technology is not taken well into account in the formalization of the cost function, then it will be on error terms and efficiency scores will be biased (Kumbhakar et Lovell, 2000 [15]).

We assumed above that a nursing home produces an output Y, corresponding to the annual number of resident-days.

Production of this output generates a cost TC, which is treated as the total annual cost of the institution. This cost is a long-term cost, including both variable costs (operating costs, payroll) and also fixed costs of the nursing home (investment expenses).

Wages w_i are average annual gross wages of nursing home staff performing a function i, for the 3 main staff categories of nursing homes: nurses w_N, nursing auxiliaries w_{NA} and support staff who are low skilled staff and do not provide any care service w_{SS}. To guarantee the homogeneity condition in wages, we divide total cost, nursing auxiliary and support staff wages by nurse wage.

Since we estimate a long-term cost function, we have no capital variable in our model. Wee need, however, a capital cost variable, but it is difficult to find a relevant one. Thus, we prefer to integrate several variables related to the building cost, such as:
-$BUILD$: the number of years since the construction of the building or, where appropriate, its last renovation;
-OWN: a categorical variable providing information on ownership structure: the manager is owner? tenant of a public, a private non-profit or a private for-profit buliding? or does he benefit from moderate rent habitations?[7]
-URB: the urbanization level of the town (city) location of the nursing home: less than 20000 inhabitants, between 20000 and 200000 inhabitants, more than 200000 inhabitants, or located in or near Paris.

[6] The treatment of endogeneity is presented in Appendix.
[7] In French: *Habitations à loyers modérés (HLM).*

A variables set of residents characteristics is also inserted into the model. It includes proportions of residents in each *GIR* group: $propGIR1$ to $propGIR6$, and a binary variable indicating whether the nursing home receives patients with Alzheimer or related diseases ALZ. Before the institutionalization of a resident, the physician (or a nurse) of the nursing home gives an opinion about this entry after a medical examination. Since it seems overall to be a shortage of beds in long-term care market, it is common that entry be denied to patients with Alzheimer's because of the extra help they often need. This variable may have different effects on costs depending on the size of the nursing home. Thus, we add an interaction term $Y.ALZ$.

In addition, we take into account various characteristics of environment of nursing homes that may have some impacts on costs: a binary variable $MedOption$ and its interaction term with the output variable $Y.MedOption$ that inform about the choice of option related to the medical care price[8] by the nursing home, and another binary variable $DRUG$ that indicates whether or not the nursing home has a pharmacy for its own use. Regarding the medical care option, we expect a positive effect of the *global* option on costs, since additional costs are funded by the nursing home. The presence of a drugstore has also an expected positive impact on costs, since it generates an additional fixed costs (cost structure, wages of pharmacists). Its implementation may certainly lead to better drug management and reduce associated costs, but drugs are direclty paid by residents and therefore have no effects on the nursing home costs.

Quality variables. Assuming that quality has a positive impact on costs, a simple way to reduce costs is to lower the quality level. We consider that the regulator includes quality of services in the desirable outputs. Therefore, to avoid providing incentives towards a deterioration of quality, one has to evaluate efficiency for a given level of quality. However, it is not easy to build relevant indicators of service quality. Quality of care in nursing home is difficult to observe. A proxy can be given by the staff/resident ratio. A study conducted in 2009 by the French Hospital Federation and the National Union of Gerontology pointed out differences in quality of care that can be generated by different levels of the staff/resident ratio [9]. Moreover, a recent survey carried out among 1802 residents in nursing homes in 2007 has shown that a major source of dissatisfaction is the lack of staff [7]. Our indicator takes into account differences in needs: it is defined as the ratio of actual staff N to the theoretical optimal staff N^*, the latter being defined by a norm established for the formal home care sector. According to this norm, the quantity of full-time equivalent (FTE) staff needed daily for an elderly person is 1 for a *GIR 1* person, 0.84 for a *GIR 2* person, 0.66 for a *GIR 3*, 0.42 for a *GIR 4*, 0.25 for a *GIR 5*, and 0.07 for a *GIR 6* person. Our first indicator, named the *staff/resident ratio gap* is defined by N/N^*, where the actual staff N includes all non-administrative staff[9]. Our second quality indicator, $LowSkill$, is based on the idea that the skill level of employees might affect the quality of care. It is computed as the ratio of non-nursing staff (support staff) on nursing staff (nurses, nursing auxiliaries).

To check whether these two indicators are effectively connected to quality, we have examined their correlation with outcomes regarding residents' health. Information about incontinence, depression and falls has been recorded for 7043 residents of our EHPA 2007 sample of nursing homes (individuals)[10]. A categorical variable measures the degree of incontinence (coded 0 to 3, 0 being a non incontinent resident, 3 corresponding to a high incontinence level), an ordered categorical variable reflects the depressed state of the resident (also coded 0 to 3), and a binary variable indicates whether the resident experienced at least one fall during the year. We have performed two logistic regressions and one binomial logistic regression to examine the link between our quality indicators and adverse outcomes[11]. Results are given in table 2. An increase in the staff/resident ratio gap has a positive impact on health outcomes (negative correlation with incontinence, depression and falls). As concerns the proportion of unskilled staff, it has a positive influence on one adverse outcome only: the risk of incontinence. These results allow us to consider that our indicators are relevant proxies of care quality.

[8] Nursing homes have to choose between the *partial* and the *global* option for the medical care section. The reimbursed health benefit basket is larger for nursing homes that have chosen the *global* option, because it includes GPs fees, radiographies and biological examinations. When the nursing home opts for the *partial* option, theses expenses are paid by the residents who are then reimbursed by their own health insurance fund.

[9] We add FTE corresponding to hours of outsourcing, interim and liberal staff.

[10] We cannot use directly this information in our stochastic frontier specification because it is not available for a representative sample of residents in each observed nursing home.

[11] Control variables included in these regressions are: age and *GIR* group of the resident, two binary variables indicating if the resident is moving and if he complains of pain, status variable of the nursing home, total number of residents, proportion of beneficiaries of the social allowance from *departement*, and presence of end-of-life rooms.

Table 2: Logistic regressions of health status variables of residents on N/N^* and $LowSkill$ variables

	Incontinence level	**Depressed state**	**Falls**
N/N^*	-0.424^{**}	-0.361^{**}	-0.270^{*}
$LowSkill$	0.263^{**}	-0.010	-0.064
$LowSkill^2$	-0.033^{**}	-0.001	0.004
$N/N^*/LowSkill$	0.022^{**}	0.009	-0.005

Significance levels : $*$: 5% $**$: 1%

Sample of 7043 residents; source EHPA 2007

Table 3 displays the proportions of nursing homes by institutional forms for several levels of quality. The lack of staff appears to be more frequent in private nonprofit nursing homes: more than half of them have a staff/resident ratio gap below 0.8! Above all, the quality level seems to be very low: all ownership together, almost 70% of nursing homes have a staff/resident ratio gap below 0.95. This result is in accordance with experiences reported in section 3 and with the common view that staff shortage contribute to jeopardize quality of services in French nursing homes. The skill level indicator gives somewhat different information: the staff appears to be more skilled in private nonprofit nursing homes, in comparison with public nursing homes, when they are not linked to a hospital.

Table 3: Proportions of nursing homes by status and quality groups

Variable categories		***PNP***	***Pub.NHB***	***Pub.HB***	**Total**
N/N^*	< 0.8	258 (50.9%)	118 (18.3%)	3 (15%)	379 (32.4%)
	$[0.8; 0.95[$	172 (33.9%)	252 (39.3%)	7 (35%)	431 (36.8%)
	$[0.95; 1.05[$	41 (8.1%)	138 (21.5%)	4 (20%)	183 (15.6%)
	> 1.05	36 (7.1%)	136 (21.1%)	6 (30%)	178 (15.2%)
$LowSkill$	< 1	170 (33.5%)	94 (14.6%)	10 (50%)	274 (23.4%)
	$[1; 1.5[$	189 (37.3%)	267 (41.4%)	6 (30%)	462 (39.5%)
	$[1.5; 2[$	101 (19.9%)	191 (29.7%)	3 (15%)	295 (25.2%)
	> 2	47 (9.3%)	92 (14.3%)	1 (5%)	140 (11.9%)
Total of nursing homes		507	644	20	1171

Our quality indicators are likely not to be exogenous. Indeed, the general low quality level is likely to derive from financial constraints associated to the budget level. At the time of the negotiation (*time t=1*) between a nursing home e and public authorities, day rates (i.e. its budget $B^e_{t=1}$) are actually setting in regard to previous prices (its budget $B^e_{t=0}$) and a rate τ^e that can be characterized as the bargaining power of the institution. Thus the budget of the nursing home e would be partly reconducted: $B^e_{t=1} = B^e_{t=0} * (\tau^e + 1)$. This budget B^e may be higher or lower than the optimal one B^{*e}. One can know the optimal budget of a nursing home e that would allow it to take care residents in a satisfactory manner without allowing private income. This is mainly based on its specific staff needs N^* and wages preveiling in its labor market:

$$B^{*e} = B(N^{*e}, \sum_i w^e_i)$$

If $B^e \geq B^{*e}$, we can assume that staff recruitment N^e is not bound by its budget B^e, the payment system would be strictly speaking a restrospective day rate whereby all costs would be reimbursed.
However, if $B^e < B^{*e}$, staff recruitment would be constrained. Let $\tilde{N^e} = f(N^{*e}, \sum_i w^e_i)$ be the notional demand for labor:

$$\tilde{N^e} \geq N^e = \frac{B^{*e} - other\ expenses}{w^e}$$

It might then be a simultaneity bias between N/N^* and B^e when N/N^* is less than 1.
To address this endogeneity problem of quality variables, we use a method of instrumental variables[12]. We then

[12]The treatment of endogeneity is presented in Appendix.

compare, for each methods, results of 3 regressions: estimation without quality variables, with N/N^* and $LowSkill$, and with instrumented quality variables $\widehat{N/N^*}$, $\widehat{LowSkill}$. We use as instrument the average staff seniority in years (excluding administrative staff) of the nursing home ($SENIOR$ and $SENIOR^2$), assuming that lower quality of care and burnout of staff, due to low staff ratios and/or low qualifications, would generate a significant turnover (Spilsbury et al., 2011 [26]).

Explanatory variables of inefficiency scores One difficulty in estimating frontiers is knowing which variables include as regressors of cost, and which ones to keep as explanatory variables of inefficiency (Vitaliano and Toren, 1994 [27]). Since the inefficiency term follows a truncated normal distribution, even when one opts for a one stage procedure (Battese and Coelli, 1995 [2]), the choice to integrate a variable as an explanatory element of cost or inefficiency will modify the likelihood function form and thus the estimate results. The status variable $STAT$ is strongly involved in this issue. In this study, we try to highlight any differences in cost-efficiency of nursing homes depending on its institutional form (*PNP, Pub.NHB and Pub.HB*). For this reason, we choose to include institutional form as regressor of inefficiency and not cost. We must be careful in analyzing efficiency scores, since status may also have a direct effect on costs, either on fixed costs related to building or on labor costs. However, by controlling for average gross wages of each nursing home and for the type of ownership of the building, we partially reduce this direct impact of status on costs. We add as regressors of inefficiency the proportion in the nursing home of beneficiaries of social allowance from *department* $BENEF$, the logarithm of GDP per capita GDP and a ratio of the part of expenditure for elderly on the part of people over 75 years *ElderlyExp* at the *departement* level. These three variables have *a priori* no effect on costs, but may have an impact on the greater or lesser financial generosity of public authorities of each *departement*.

Table 4: Summarization of explanatory variables

Y	Y
w_i	w_N, w_{NA}, w_{SS}
z_{k_1}	$DRUG$, $BUILD$, URB, OWN
z_{k_2}	ALZ, $MedOption$
G	$propGIR1$, $propGIR2$, $propGIR3$, $propGIR4$, $propGIR5$, $propGIR6$
q	N/N^*, $LowSkill$
Explanatory variables of inefficiency terms (m)	$STAT$, $BENEF$, GDP

5 Data

We use a french national sample of 1171 nursing homes in 2007 for regressions with cross-sectionnal data. This sample represents 32,5% of the total number of public and private non-profit nursing homes in France. We have administrative data on wages from the French National Institute for Statistics and Economic Studies *(INSEE, DADS 2008 data)*. Nursing homes provide annual information on each of their employees (gross wage, FTE, age, type of contract, type of collective agreement). From this, we can calculate an average gross wage per staff category and per nursing home, which is a great adventage for such an analysis. We have information on other variables in the comprehensive four-year survey *EHPA 2007* (*DREES*, Directorate for Reasearch, Studies, Assessment and Statistics of Ministry of Health). We do not have data on costs of nursing homes. However, we know that pricing is retrospective for two reasons: prices are set based on the *ex post* observation of costs, and nursing homes cannot make a profit since any outcome would be taken into account for the subsequent annual budgets. We can therefore assume that total annual cost of the institution is the product of the daily rate by the annual number of resident-days. Thus, we multiply resident-days from permanent stays with the day rate for this category of stay, and the ones of temporary stays and day care with their respective prices. Information on day rates for these alternative stays is provided by the *MAUVE 2010* survey *(DREES)*.

We report below the descriptive statistics for each variable included in the model and for each status category (Table 5). We notice that the daily cost per resident is higher for public hospital based nursing homes than private ones, and these latter are also more costly than public non-hospital based ones. We try to highlight some factors explaining these differences in costs, but these three categories seem to be very close. Nevertheless, we observe a relatively higher activity for public nursing homes, mainly for the public hospital based ones. Dependency levels of residents are almost identical, but private non-profit nursing homes are more likely to refuse Alzheimer patients. Regarding quality variables, the nurse resident ratio gap is higher in public sector, but public nursing homes, however, hire more unskilled employees. Public institutions are more likely to own building or, if such is not the case, are then often in public rental. As for labor costs, if net wages of staff working in public nursing homes are on average higher than in the private sector[13], annual gross wages appear to be relatively higher for private non-profit nursing homes (for nursing auxiliairies and support staff). This is due to exemptions from some of the labor costs afforded to public institutions.

For our panel data model, we use another sample of 740 nursing homes, for which we have data for the years 2003 and 2007. We insert the same variables as for cross-sectional analysis, except $BUILD$, OWN and $MedOption$ for whom we have no information in 2003. We have no indication of seniority staff, so we cannot use any estimation method by instrumental variables. Observations in 2003 are given by the *EHPA 2003* comprehensive survey *(DREES)*. We have no information about day rates for temporary stays and day care, therefore as dependent variable we use the number of resident-days from permanent stays only. Concerning labor costs, we use wages administrative data of 2005 *(DADS 2005, INSEE)*. We have only $1/12^{th}$ of the sample data for this year 2005. In order to keep a satisfactory number of observations, we infer average gross wages per nursing home by estimating wage for each employee based on its activity, age, type of contract, collective agreement and location of the institution. We do the same for 2007 so as to lose as few observations as possible.

We present descriptive statistics for this new sample for 2003 and 2007 in Table 6.

Daily costs per resident increased by about 15% between 2003 and 2007, but this rise is less important for public non-hospital based nursing homes. This increase does not seem to result of wage developments, at least for the private sector in which wages decreased between 2003 and 2007 (in constant euros). The cost rise is probably partly due to the increased level of dependency of residents and the more frequent care of Alzheimer residents. This confirms the idea that aged people enter later in nursing homes. We also note an increase in nurse resident ratio gap (closer to 1) and a relative rise of skilled staff between 2003 and 2007.

6 Results

Specification tests. From the stochastic frontier estimation (estimate (1)), we test the hypothesis $H_0 : \sigma_u^2 = 0$ against $H_1 : \sigma_u^2 > 0$ with a likelihood ratio test. We obtain a ratio of $LR = 5.28$ with a p-value of 0.011. The null hypothesis is rejected, thus, the use of a stochastic frontier analysis seems relevant. In the same idea, from residuals obtained by an estimation with a truncated normal distribution of error terms, a test of skewness allows us to study the significance of inefficiency term u: Coelli (1995) [5] has shown that the presence of an efficiency term negatively distorts residuals obtained by OLS regression. A p-value of 0.020 is obtained by a test of residuals normality (z-test) (z=1.917); thus, we reject the normality assumption, which confirms the statistical significance of inefficiency terms in our model.

Estimate results The Translog model (model 1) is estimated by various methods: stochastic frontier analysis (SFA), quantile regression (QR) and panel data estimation with random effects (RE) and correlated random effects (CRE). Estimation results are shown below (Tables 7 and 8). We specify for each of these estimates if quality variables q have been included, and if so if endogeneity was corrected by instrumental variable method $(\widehat{q})$[14]. Given endogeneity bias that can be generated by incorporating not instrumented quality variables, we are interested more particularly in the estimation results (1), (3), (4) and (6), and panel data estimates (7), (9) and (10).

Results are on the whole fairly closed except for coefficients associated with output Y and factor prices $w_{NA/N}$ and $w_{SS/N}$. Since we used a translog cost model, first-order coefficients associated with these variables can be interpreted as cost elasticities. Depending on the model, the output has a significant positive effect in first or second order. As expected, an increase in wages leads to increased costs, but this effect is decreasing for nursing auxiliaries wage. About average wage of support staff, coefficients of first and second order are both positive, so

[13] For comparison, for the same sample, the average annual net wage of a nursing auxiliary in *PNP* nursing homes is 18120 euros, 20695 euros in *Pub.NHB* nursing homes and 21441 euros in *Pub.HB* nursing homes. The average annual net wage of a nurse is respectively 25002 euros, 26368 euros and 27193 euros.

[14] Standard errors associated with regressions performed after treatment of endogeneity $(\widehat{q})$ correspond here to standard deviations of second stage. These ones, as well as indicators of significance, will have thereafter to be corrected.

the impact of an increase of $w_{SS/N}$ on costs is increasing. Interaction term between support staff and nursing auxiliaries wages can be interpreted as coefficient of substitution between these two staff categories: when this term is negative, for example, this means that a nursing auxiliaries wage increase generates a lesser effect of support staff wage on costs, which may be due to a decrease in hiring support staff conducted by reduction in number of nursing auxiliaries (complementarity effect) in response to the increase of $w_{NA/N}$. However, it is difficult to draw such a conclusion here, since this coefficient is positive for estimates of cross-sectional data but negative for models with panel data. Interaction terms between output Y and wages, when significantly negative, can be interpreted as scale effects: the more important activity of nursing homes is, the less a wage increase affects costs.

Concerning the degree of dependency of residents, the greater proportions of GIR 3 to 6 are, the lower costs are. However, proportion of people in *GIR 1* respect to proportion of people in *GIR 2* does not seem to have an effect on costs[15]. Care of Alzheimer patients would have, other things being equal, a negative effect on costs[16], but smaller when Y increases. About capital costs, be renting a *HLM* or a private for-profit building is more costly, as well as being located in or near Paris. A recent construction or renovation is generating costs, since the age of the building has a significant negative effect. Of course having its own pharmacy also increases costs. Concerning the choice of the *global* option of medical care price, one could have expected it to have a positive impact on costs since the reimbursed health benefit basket is larger than with the *partial* option. The effect of this variable is not significant, however. Institutions which have opted for *global* option would spend *a priori* less than nursing homes with *partial* option.

Quality variables have the expected effect on nursing homes expenditure: a lower skilled staff has a significant negative effect on costs, while a higher nurse resident ratio gap has a significant positive effect. An increase of 0.1 in nurse resident ratio gap would result in higher costs of 11.07% as estimated by stochastic frontier analysis, and 7.61% as estimated by quantile regression. Thus, not taking quality into account in a new pricing mechanism may have adverse effects on quality of care in nursing homes.

To examine robustness of coefficients associated with explanatory variables of inefficiency in stochastic frontier analysis, we run OLS regressions of inefficiency scores obtained by quantile regression and random effects estimates by the following variables: $STAT$, $BENEF$, GDP and *ElderlyExp*. Proportion of beneficiaries have a positive impact on cost-inefficiencies. This can perhaps be explained by a higher generosity of public authorities to nursing homes with residents with limited resources. As expected, we also observe significant positive effects of GDP and *ElderlyExp* on nursing homes costs. This suggests that generosity of public authorities varies by *departement*. About the effect of institutional forms, public hospital based nursing homes would be less efficient than the non-hospital based and private ones. According to cross-sectional estimates, there are no significant differences between private and public non-hospital based nursing homes when quality is not taken into account. When quality variables are included in cost model, private non-profit nursing homes appear to be significantly more efficient than public ones. We check this result by parametric and non-parametric tests presented below (Table 10).

Estimation results of Mundlak auxiliary equation indicate there are unobserved heterogeneity correlated with regressors, which is linked to qualification structure of staff and wages variables. Thus, estimation results with random effects methods (7) and (8) may be biased.

[15]For each institution, rates are identical for residents of *GIR 1* and *GIR 2*, so as for *GIR 3* and *GIR 4*, and between *GIR 5* and *GIR 6*. Thus, when the physician (or nurse) of the nursing home determines *GIR* groups of each resident, he pays little attention to the classification of residents between categories 1 and 2. This may explain why the level of dependency changes little between residents of *GIR 1* and *GIR 2*, which consequently generates few differences in costs.

[16]This is probably due to the fact that a nursing home that takes care of Alzheimer patients in specialized units often receive subsidies from *departement* or national agencies, which allow them to reduce their prices. And costs that we observe correspond only to prices fixed for each nursing homes.

Table 5: Descriptive statistics (for cross-sectionnal estimations)

Variables	***PNP***	***Pub.NHB***	***Pub.HB***	**Total**
	Average (σ^2)	Average (σ^2)	Average (σ^2)	Average (σ^2)
TC *(in current euros)*	2272833 (1478669)	2339552 (1349127)	3728178 (2086564)	2334382 (1432294)
Y	27082 (13127)	29046 (14376)	40185 (22286)	28386 (14125)
TC/Y (in current euros)	*82.8 (15.4)*	*80 (11.4)*	*91.1 (19.1)*	*81.4 (13.5)*
w_{NA} *(in current euros)*	23925 (1815)	24350 (1396)	24577 (1154)	24170 (1602)
w_N *(in current euros)*	33224 (3435)	31263 (3301)	32055 (3049)	32126 (3489)
w_{SS} *(in current euros)*	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)
BUILD	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)
LowSkill	1.27 (0.637)	1.5 (0.534)	1.07 (0.496)	1.39 (0.592)
SENIOR	6.71 (2.92)	8.32 (2.54)	9.65 (3.44)	7.65 (2.85)
BENEF	0.326 (0.306)	0.328 (0.277)	0.325 (0.207)	0.327 (0.289)
GDP	25016 (8365)	23493 (4716)	24468 (3298)	24169 (6575)
ElderlyEXP	1.97 (0.284)	1.9 (0.264)	1.89 (0.202)	1.93 (0.274)
(NON INCLUDED VARIABLES)				
Number of residents	74.9 (36)	80.2 (39.7)	112 (62.5)	78.5 (38.9)
Number of beds	77.6 (37.6)	83 (41.1)	118 (67.5)	81.2 (40.6)
Occupancy rate	0.979 (0.045)	0.98 (0.054)	0.984 (0.032)	0.979 (0.050)
Length of stay	3.7 (1.17)	3.94 (0.955)	3.8 (1.08)	3.83 (1.06)
Total staff ratio	0.531 (0.105)	0.614 (0.106)	0.593 (0.169)	0.578 (0.115)
Ratio nurse/staff	0.090 (0.033)	0.086 (0.025)	0.123 (0.054)	0.088 (0.030)
Ratio nursing aux./staff	0.295 (0.079)	0.299 (0.072)	0.336 (0.117)	0.298 (0.076)
Ratio support staff/staff	0.295 (0.184)	0.388 (0.114)	0.324 (0.136)	0.347 (0.155)
	N (%)	N (%)	N (%)	N (%)
OWN=owner	231 (45.6%)	400 (62.1%)	17 (85%)	648 (55.3%)
OWN=Pub.rental	48 (9.5%)	218 (33.8%)	2 (10%)	268 (22.9%)
OWN=HLM.rental	98 (19.3%)	25 (3.9%)	1 (5%)	124 (10.6%)
OWN=PNP.rental	70 (13.8%)	1 (0.2%)	0 (0%)	71 (6.1%)
OWN=PFP.rental	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%)
DRUG=0	475 (93.7%)	589 (91.5%)	4 (20%)	1068 (91.2%)
DRUG=1	32 (6.3%)	55 (8.5%)	16 (80%)	103 (8.8%)
MedOption=0	396 (78.1%)	450 (69.9%)	5 (25%)	851 (72.7%)
MedOption=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%)
N.	***507***	***644***	***20***	***1171***

Sample of 1171 nursing homes; source EHPA 2007

Table 6: Descriptive statistics (for panel data estimations)

Variables	Year	***PNP***	***Pub.NHB***	***Pub.HB***	**Total**
		Average (σ^2)	Average (σ^2)	Average (σ^2)	Average (σ^2)
TC	2003	2068244 (1163467)	2071640 (1401619)	2912464 (1847129)	2309472 (1540492)
(in constant euros (/2007))	2007	2538410 (1509929)	2377857 (1461653)	3648355 (2283580)	2771637 (1824957)
Y (permanent stays only)	2003	28380 (14067)	28583 (14636)	36571 (19712)	30803 (16505)
	2007	29917 (15809)	29142 (14422)	39676 (22876)	32280 (18078)
TC/Y	*2003*	*72.4 (13.3)*	*71.1 (12.2)*	*78.1 (14.4)*	*73.4 (13.4)*
(in constant euros (/2007))	*2007*	*84.3 (14.5)*	*80.1 (11.6)*	*92.9 (25.5)*	*84.7 (18)*
w_{NA}	2003	24709 (1031)	20581 (835)	20859 (610)	21608 (1890)
(in constant euros (/2007))	2007	24420 (1173)	23187 (1312)	23615 (1292)	23589 (1364)
w_N	2003	34637 (1094)	30384 (1150)	30414 (826)	31369 (2073)
(in constant euros (/2007))	2007	33270 (1167)	31858 (1828)	32248 (1472)	32290 (1693)
w_{SS}	2003	22547 (1140)	18583 (712)	18628 (638)	19506 (1850)
(in constant euros (/2007))	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)
LowSkill	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)
GDP	2003, 2007	23506 (3358)	23620 (4670)	23196 (4559)	23474 (4372)
ElderlyExp	2003, 2007	1.94 (0.289)	1.89 (0.270)	0.92 (0.261)	1.91 (0.273)
		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%)
DRUG=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%)
DRUG=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%)
N	2003	***170***	***360***	***210***	***740***
	2007	***169***	***363***	***208***	***740***

Sample of 740 nursing homes; source: EHPA 2003 and EHPA 2007

Table 7: Estimate results: stochastic frontier analysis (SFA) and quantile regression (QR)

	(SFA - without q) (1)		(SFA - q) (2)		(SFA - $\widehat{q}$) (3)		(QR - without q) (4)		(QR - q) (5)		(QR - $\widehat{q}$) (6)	
	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)
α_0	2.822*	(1.142)	0.441	(0.996)	-5.495**	(1.012)	1.635	(1.474)	-0.899	(1.355)	-5.137†	(3.042)
Y	-0.510*	(0.221)	-0.073	(0.193)	1.079**	(0.195)	-0.306	(0.284)	0.174	(0.269)	1.022†	(0.566)
$w_{NA/N}$	0.586	(1.086)	1.270	(0.952)	3.898**	(0.978)	-0.193	(1.838)	0.027	(1.603)	1.573	(2.994)
$w_{SS/N}$	1.961†	(1.172)	2.320*	(1.014)	1.736†	(0.991)	3.073	(1.909)	3.565*	(1.653)	3.250	(2.826)
$w^2_{NA/N}$	-1.708*	(0.827)	-1.950**	(0.636)	-2.241**	(0.836)	-1.922†	(0.991)	-1.476*	(0.735)	-1.918	(1.445)
$w^2_{SS/N}$	0.829*	(0.351)	0.742*	(0.337)	0.902*	(0.383)	-0.329	(0.281)	-0.498†	(0.265)	1.258**	(0.431)
Y^22	0.136**	(0.022)	0.092**	(0.019)	-0.023	(0.020)	0.117**	(0.028)	0.068*	(0.027)	-0.015	(0.056)
$Y.w_{NA/N}$	-0.054	(0.108)	-0.118	(0.096)	-0.383**	(0.105)	0.056	(0.181)	0.031	(0.159)	-0.136	(0.310)
$Y.w_{SS/N}$	-0.110	(0.116)	-0.134	(0.101)	-0.035	(0.096)	-0.249	(0.185)	-0.297†	(0.163)	-0.188	(0.281)
$w_{NA/N}.w_{SS/N}$	0.192	(0.544)	0.567	(0.436)	1.197*	(0.565)	0.971	(0.653)	0.918†	(0.514)	0.840	(0.987)
$DRUG$	0.060**	(0.015)	0.035**	(0.013)	-0.028	(0.019)	0.020	(0.019)	0.040*	(0.018)	-0.055	(0.036)
$MedOption$	0.110	(0.193)	0.184	(0.173)	0.218	(0.195)	-0.067	(0.223)	0.099	(0.225)	0.002	(0.327)
$Y.MedOption$	-0.010	(0.019)	-0.017	(0.017)	-0.020	(0.019)	0.008	(0.022)	-0.008	(0.022)	0.002	(0.032)
$OWN = owner$	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
$OWN = Pub.rental$	-0.016†	(0.009)	-0.007	(0.008)	-0.006	(0.010)	0.006	(0.011)	0.005	(0.011)	0.015	(0.018)
$OWN = HLM.rental$	0.019	(0.012)	0.027*	(0.011)	0.033†	(0.020)	0.037*	(0.015)	0.051**	(0.015)	0.026	(0.038)
$OWN = PNP.rental$	0.029†	(0.016)	0.023	(0.014)	0.023	(0.018)	0.031	(0.019)	0.040*	(0.018)	0.004	(0.035)
$OWN = PP.rental$	0.013	(0.017)	0.027†	(0.016)	0.087**	(0.022)	0.026	(0.021)	0.062**	(0.020)	0.077†	(0.041)
$URB : x < 20000$	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
$URB : 20000 < x < 200000$	-0.004	(0.009)	0.003	(0.008)	0.029**	(0.010)	-0.021†	(0.011)	-0.008	(0.011)	0.007	(0.019)
$URB : x > 200000$	0.006	(0.009)	0.014†	(0.008)	0.040**	(0.010)	0.017	(0.011)	0.021*	(0.011)	0.045*	(0.018)
$URB : PARIS$	0.090**	(0.016)	0.072**	(0.014)	0.030†	(0.017)	0.112**	(0.019)	0.084**	(0.018)	0.044	(0.035)
ALZ	-0.435*	(0.186)	-0.399*	(0.160)	-0.448*	(0.182)	-0.268	(0.217)	-0.080	(0.217)	-0.225	(0.314)
$Y.ALZ$	0.047*	(0.018)	0.042**	(0.016)	0.044*	(0.018)	0.030	(0.021)	0.011	(0.021)	0.023	(0.031)
$BUILD$	-0.001*	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)	0.000*	(0.000)	-0.001**	(0.000)	-0.001*	(0.000)
$propGIR2$	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
$propGIR1$	0.010	(0.060)	0.005	(0.053)	0.003	(0.057)	0.058	(0.075)	0.027	(0.073)	0.058	(0.109)
$propGIR3$	-0.133*	(0.067)	-0.178**	(0.059)	-0.196*	(0.079)	-0.049	(0.085)	-0.114	(0.079)	-0.038	(0.141)
$propGIR4$	-0.405**	(0.055)	-0.492**	(0.049)	-0.665**	(0.088)	-0.359**	(0.069)	-0.450**	(0.066)	-0.423*	(0.169)
$propGIR5$	-0.633**	(0.066)	-0.760**	(0.059)	-0.899**	(0.101)	-0.678**	(0.086)	-0.788**	(0.083)	-0.791**	(0.196)
$propGIR6$	-0.493**	(0.056)	-0.731**	(0.053)	-0.960**	(0.178)	-0.504**	(0.064)	-0.746**	(0.062)	-0.629†	(0.363)
$LowSkill$	-	-	-0.041**	(0.006)	-0.259**	(0.048)			-0.042**	(0.007)	-0.287**	(0.092)
N/N^*	-	-	0.356**	(0.020)	1.107**	(0.172)			0.345**	(0.026)	0.761*	(0.383)
δ_0	-3.357**	(0.920)	-3.657**	(1.326)	-1.980**	(0.317)	-0.589**	(0.174)	-0.639**	(0.153)	-0.583**	(0.199)
$STAT = PNP$	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
$STAT = PubNHB$	0.024	(0.025)	-0.148*	(0.072)	0.041*	(0.018)	0.000	(0.006)	-0.017**	(0.006)	0.018*	(0.007)
$STAT = PubHB$	0.176*	(0.081)	0.200*	(0.090)	0.152**	(0.043)	0.078**	(0.024)	0.042*	(0.021)	0.096**	(0.027)
$BENEF$	0.121**	(0.040)	0.136*	(0.067)	0.096**	(0.022)	0.026*	(0.011)	0.023*	(0.009)	0.039**	(0.012)
GDP	0.250**	(0.068)	0.271**	(0.089)	0.143**	(0.031)	0.061**	(0.017)	0.065**	(0.015)	0.056**	(0.020)
$ElderlyExp$	0.341**	(0.098)	0.364*	(0.154)	0.229**	(0.030)	0.038**	(0.011)	0.042**	(0.010)	0.045**	(0.013)
σ^2	0.017**	(0.002)	0.027**	(0.009)	0.014**	(0.001)						
γ	0.190	(0.144)	0.693**	(0.114)	0.053	(0.064)						

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Sample of 1171 nursing homes in 2007; source: EHPA 2007

Table 8: Estimate results: random effects (RE) and correlate random effects (CRE) estimations

	(RE - without q) (7)		(RE - q) (8)		(CRE - without q) (9)		(CRE - q) (10)	
	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)	Coeff.	(St.Err)
α_0	1.397	(1.562)	-0.595	(1.406)	0.323	(1.808)	-2.734†	(1.574)
Y	-0.190	(0.284)	0.125	(0.253)	-0.197	(0.295)	0.217	(0.261)
$w_{NA/N}$	3.119†	(1.764)	2.773†	(1.678)	4.396*	(2.045)	3.895*	(1.987)
$w_{SS/N}$	2.787†	(1.682)	2.086	(1.592)	2.213	(2.045)	2.072	(1.994)
$w^2_{NA/N}$	4.420**	(1.521)	4.748**	(1.466)	4.473**	(1.516)	5.091**	(1.441)
$w^2_{SS/N}$	-0.801	(0.761)	-0.367	(0.728)	-0.502	(0.763)	-0.161	(0.718)
Y^22	0.094**	(0.027)	0.069**	(0.024)	0.090**	(0.027)	0.053*	(0.024)
$Y.w_{NA/N}$	-0.178	(0.169)	-0.141	(0.161)	-0.337†	(0.198)	-0.286	(0.193)
$Y.w_{SS/N}$	-0.335*	(0.159)	-0.242	(0.150)	-0.273	(0.198)	-0.255	(0.193)
$w_{NA/N}.w_{SS/N}$	-0.837	(0.961)	-0.981	(0.921)	-1.475	(0.988)	-1.893*	(0.942)
$DRUG$	0.117**	(0.011)	0.104**	(0.010)	0.114**	(0.011)	0.100**	(0.010)
$URB : x < 20000$	ref	ref	ref	ref	ref	ref	ref	ref
$URB : 20000 < x < 200000$	-0.003	(0.013)	0.001	(0.011)	-0.003	(0.013)	0.004	(0.011)
$URB : x > 200000$	0.014	(0.013)	0.017	(0.011)	0.014	(0.013)	0.018	(0.011)
$URB : PARIS$	0.086**	(0.021)	0.058**	(0.019)	0.091**	(0.021)	0.046*	(0.018)
ALZ	0.003	(0.179)	0.061	(0.172)	-0.047	(0.182)	-0.007	(0.172)
$Y.ALZ$	0.001	(0.017)	-0.004	(0.017)	0.006	(0.018)	0.002	(0.017)
$propGIR2$	ref	ref	ref	ref	ref	ref	ref	ref
$propGIR1$	0.064	(0.070)	0.067	(0.065)	0.015	(0.091)	0.028	(0.089)
$propGIR3$	-0.097	(0.071)	-0.126†	(0.066)	-0.151	(0.092)	-0.153†	(0.090)
$propGIR4$	-0.266**	(0.059)	-0.328**	(0.055)	-0.258**	(0.080)	-0.302**	(0.079)
$propGIR5$	-0.479**	(0.068)	-0.624**	(0.064)	-0.272**	(0.096)	-0.322**	(0.097)
$propGIR6$	-0.490**	(0.061)	-0.708**	(0.058)	-0.389**	(0.090)	-0.502**	(0.094)
$Year$	0.070**	(0.009)	0.029**	(0.009)	0.083**	(0.010)	0.059**	(0.011)
$LowSkill$	-	-	-0.031**	(0.006)	-	-	-0.027**	(0.008)
N/N^*	-	-	0.313**	(0.021)	-	-	0.122**	(0.032)
$\overline{Y}$					0.144	(0.144)	0.202	(0.130)
$\overline{w_{NA/N}}$					-3.978	(3.561)	-3.368	(3.176)
$\overline{w_{SS/N}}$					1.205	(3.304)	-0.584	(2.974)
$\overline{propGIR1}$					0.086	(0.143)	0.060	(0.129)
$\overline{propGIR3}$					0.137	(0.145)	0.028	(0.132)
$\overline{propGIR4}$					0.037	(0.121)	0.023	(0.111)
$\overline{propGIR5}$					-0.380**	(0.138)	-0.480**	(0.129)
$\overline{propGIR6}$					-0.099	(0.124)	-0.241*	(0.121)
$\overline{Y.w_{NA/N}}$					-0.104	(0.324)	0.097	(0.291)
$\overline{Y.w_{SS/N}}$					0.480	(0.352)	0.410	(0.314)
$\overline{w_{NA/N}.w_{SS/N}}$					1.454†	(0.798)	1.481*	(0.712)
$\overline{LowSkill}$							-0.012	(0.011)
$\overline{N/N*}$							0.328**	(0.042)
δ_0	0.063	(0.211)	0.520**	(0.141)	0.096	(0.212)	0.631**	(0.139)
$STAT = PNP$	ref	ref	ref	ref	ref	ref	ref	ref
$STAT = PubNHB$	0.058**	(0.008)	0.026**	(0.005)	0.053**	(0.008)	0.017**	(0.005)
$STAT = PubHB$	0.082**	(0.008)	0.047**	(0.006)	0.081**	(0.008)	0.041**	(0.006)
$BENEF$	0.072**	(0.011)	0.043**	(0.008)	0.068**	(0.011)	0.037**	(0.007)
GDP	0.117**	(0.021)	0.068**	(0.014)	0.114**	(0.021)	0.058**	(0.014)
$ElderlyExp$	0.042**	(0.011)	0.033**	(0.007)	0.041**	(0.011)	0.032**	(0.007)

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Sample of 740 nursing homes in 2003 ans 2007; source EHPA 2003 et EHPA 2007

Inefficiency scores. We present in Table 9 descriptive statistics of inefficiency terms obtained by different estimation methods depending on institutional groups of nursing homes. These terms are respectively associated with asymmetric residual term u_e in stochastic frontier analysis, disturbance in quantile regression, and individual specific random effects for panel data methods. In the light of scores obtained by regressions with cross-sectionnal data, private non-profit and public non-hospital based institutions would be relativeley little inefficient, since only 2 to 11 % of costs would be non explained by any variables in the model. This rate is over 9% for public hospital based facilities.

Variability of inefficiency scores obtained by stochastic frontier analysis is relatively low for private non-profit and public non-hospital based nursing homes. But variance of residuals and specific effects, respectively obtained by quantile regression and panel data estimations, however, suggests that a significant proportion of cost variability remains unexplained for some nursing homes of each of these institutional form. As to public hospital based institutions, whatever the estimation method, we observe a strong dispersion of their inefficiency terms. Significant gains in efficiency could probably be achieved by some nursing homes in light of those results.

Introduction of quality variables (estimates (3), (6) and (10)) slightly changes the average inefficiency of nursing homes of each institutional forms. If mean and median levels of inefficiency of private non-profit and public non-hospital based institutions seem almost identical when quality is not taken into account, inefficiencies become relatively more important for the public ones when quality is addressed. This result is confirmed by parametric (average comparison) and non-parametric tests (Kruskal-Wallis test[17]) presented in Table 10. When quality is taken into account, differences between institutional forms on nursing homes are statistically significant: public institutions would then be less efficient than private non-profit ones. This statement is surprising since private non-profit nursing homes have an average daily cost per resident higher than public non-hospital based facilities. This can confirm the idea of Newhouse (1970) [20] that private non-profit institutions may be less cost-efficient, and that due to relatively higher levels of quality.

To deepen the study of quality of care and its impact on costs, we plot (Figure 2) correlations between inefficiency scores obtained by estimate without quality variable and the following quality variables: N/N^*, $LowSkill$, and the extended variable $N/N^*/LowSkill$. We try to assess how it is possible to reduce nursing homes inefficiencies without reducing their quality. We choose to use inefficiency terms from panel data estimate (9) because of an higher variability of them. There seems to be some correlation between quality and inefficiency scores. To check this, we estimate several Spearman coefficients[18] between the extended quality variable $N/N^*/LowSkill$ and inefficiency scores of regression (9), depending on institutional form and nurse resident ratio gap of nursing homes. Results are presented in Table 11. Correlation between quality and inefficiency for all nursing homes is positive and significant (Spearman coefficient of 0.275). This is lower for institutions with insufficient staff ratios (when N/N^* is between 0.8 and 0.95). This means that cost-inefficiencies are then less well explained by quality factors. In other words, nursing homes with insufficient staff ratios would not be less costly, other things being equal. This is quite likely because of additional costs that may result in lack of staff: staff turnover, absenteeism, use of temporary work, etc.

Table 9: Comparison of estimated inefficiency terms

Estimation	***PNP***				***PubNHB***				***PubHB***			
method	Av.	1^{st} q.	Median	3^{rd} q.	Av.	1^{st} q.	Median	3^{rd} q.	Av.	1^{st} q.	Median	3^{rd} q.
SFA (1)	1.03	1.01	1.02	1.03	1.03	1.01	1.02	1.03	1.07	1.04	1.05	1.08
SFA (2)	1.08	1.04	1.06	1.09	1.05	1.03	1.04	1.05	1.13	1.07	1.12	1.20
SFA (3)	1.02	1.01	1.01	1.02	1.03	1.01	1.01	1.03	1.09	1.05	1.08	1.11
QR(4)	0.11	0.01	0.09	0.18	0.11	0.01	0.08	0.17	0.19	0.07	0.19	0.29
QR (5)	0.11	0.01	0.09	0.16	0.08	0.01	0.07	0.13	0.14	0.03	0.14	0.24
QR(6)	0.10	0.00	0.08	0.16	0.11	0.02	0.09	0.16	0.19	0.07	0.18	0.29
RE (7)	1.34	1.26	1.34	1.4	1.39	1.33	1.39	1.46	1.42	1.33	1.41	1.49
RE(8)	1.28	1.23	1.29	1.33	1.31	1.27	1.31	1.35	1.33	1.27	1.32	1.38
CRE (9)	1.35	1.28	1.35	1.41	1.4	1.33	1.4	1.46	1.43	1.34	1.41	1.5
CRE (10)	1.29	1.24	1.29	1.34	1.31	1.27	1.31	1.35	1.34	1.27	1.32	1.38

Two samples of 1771 and 740 nursing homes; source: EHPA 2003 and EHPA 2007

[17]The Kruskal-Wallis test compares average ranks of observations of two samples; it is non-parametric since no assumption is made on the shape of the distribution.

[18]Spearman correlation coefficient is a non-parametric measure of statistical dependence between two variables; it is a Pearson correlation coefficient calculated on the ranks of observations associated with each variable.

Table 10: Parametric and non-parametric tests on inefficiency terms

	Average comparison test		Kruskal-Wallis test	
	p-value	Less efficient?	p-value	Less efficient?
SFA (1): without var q				
PNP against *PubNHB*	0.9843	-	0.5702	-
PNP against *PubHB*	0.0032	*PubHB*	0.0001	*PubHB*
PubNHB against *PubHB*	0.0014	*PubHB*	0.0001	*PubHB*
SFA (3): with var q				
PNP against *PubNHB*	0.0136	*PubNHB*	0.0001	*PubNHB*
PNP against *PubHB*	0.0000	*PubHB*	0.0001	*PubHB*
PubNHB against *PubHB*	0.0000	*PubHB*	0.0001	*PubHB*
QR (4): without var q				
PNP against *PubNHB*	0.8060	-	0.9309	-
PNP against *PubHB*	0.0117	*PubHB*	0.0098	*PubHB*
PubNHB against *PubHB*	0.0078	*PubHB*	0.0062	*PubHB*
QR (6): with var q				
PNP against *PubNHB*	0.1372	-	0.0656	*PubNHB*
PNP against *PubHB*	0.0067	*PubHB*	0.0045	*PubHB*
PubNHB against *PubHB*	0.0103	*PubHB*	0.0166	*PubHB*
CRE (9): without var q				
PNP against *PubNHB*	0.0000	*PubNHB*	0.0025	*PubNHB*
PNP against *PubHB*	0.0000	*PubHB*	0.0001	*PubHB*
PubNHB against *PubHB*	0.0000	*PubHB*	0.0001	*PubHB*
CRE (10): with var q				
PNP against *PubNHB*	0.0004	*PubNHB*	0.0001	*PubNHB*
PNP against *PubHB*	0.0000	*PubHB*	0.0001	*PubHB*
PubNHB against *PubHB*	0.0000	*PubHB*	0.0037	*PubHB*

Two samples of 1771 and 740 nursing homes ; source EHPA 2003 and EHPA 2007

Table 11: Spearman's rank correlation coefficients between inefficiency term u from estimate (9) and the extended quality variable q

N/N^*	**PNP**	**PubNHB**	**PubHB**	**Total**	**% of N.H.**	σ_q	σ_u
$N/N^* < 0.8$	0.329^{**}	-0.060	0.376^{**}	0.275^{**}	29.9%	6.5	0.10
$0.8 \leq N/N^* < 0.95$	-0.060	0.136	0.110	$0.113^{\dagger}$	33.1%	1.12	0.10
$0.95 \leq N/N^* < 1.05$	-0.127	$0.180^{\dagger}$	0.361^{*}	0.193^{*}	18.9%	1.46	0.11
$N/N^* \geq 1.05$	0.615^{**}	0.274^{**}	0.135	0.270^{**}	18.1%	1.56	0.15
Total	0.211^{**}	0.234^{**}	0.303^{**}	0.275^{**}	100%	3.7	0.12

Sample of 740 nursing homes; source EHPA 2003 and 2007

7 Conclusion

This study on a large sample of nursing homes in France is providing several informations on cost factors. Results robustness is examined by the use of various estimation methods.

Concerning the effect of institutional category, although public non-hospital based nursing homes have costs per resident lower than the private non-profit ones, they seem less efficient after controlling for quality and exogenous

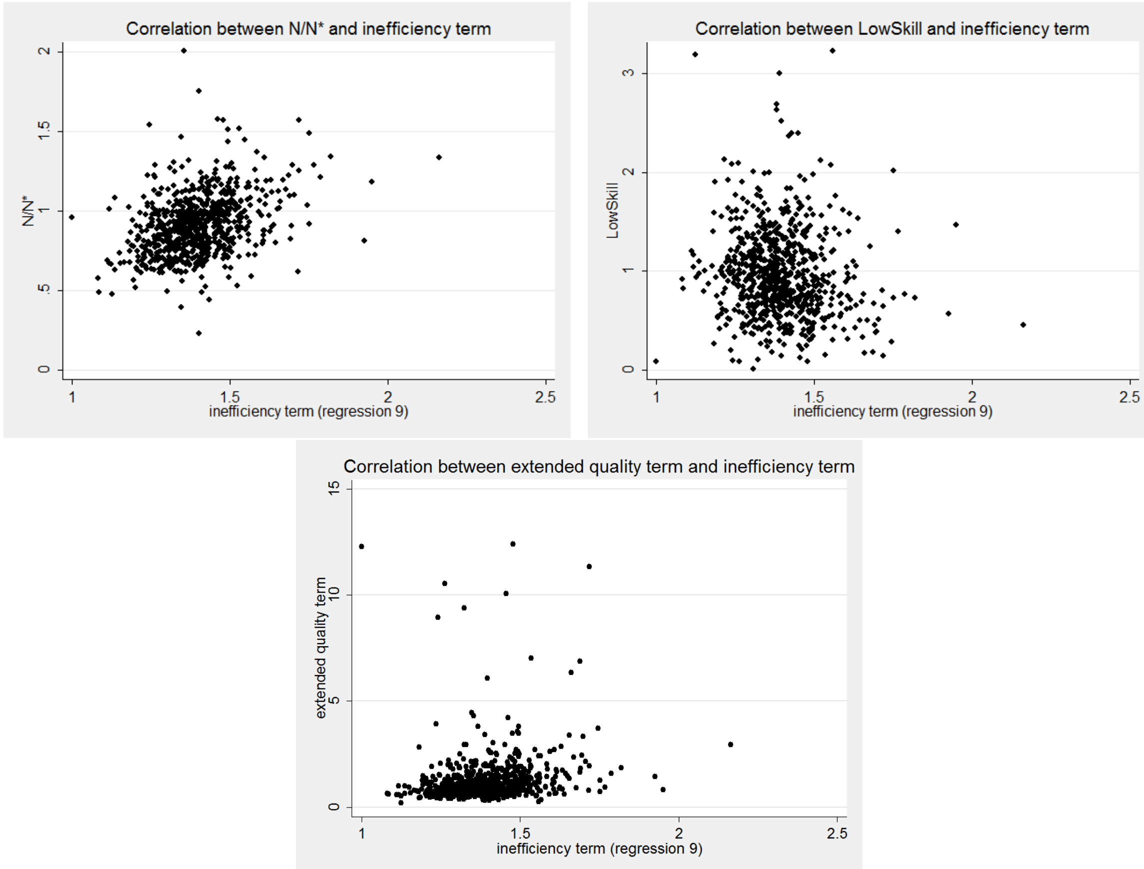


Figure 2: Correlations between inefficiency term u from estimate (9) and quality variables

cost factors. Differences are all the more important for public hospital based nursing homes, for which unjustified costs, and thus inefficiencies, are on average 15%.

Differences in quality of care we are trying to estimate by a low skilled staff ratio and a nurse resident ratio gap, explain only part of cost variability. So, the least efficient nursing homes are not always those with high staff ratios. Inefficiency terms would reflect poor management of nursing homes. Given this result, introduction of a new pricing mechanism might be justified: it would induce some nursing homes to reduce costs and thereby generate greater equity between them. Nevertheless we must be careful in such an interpretation of results, and this for several reasons.

Inefficiency scores we get are based on control variables introduced in our model. And we could not take into account all of cost specificities of nursing homes. For instance, we have not included the fact that some institutions may have building for free by their municipalities. Costs calulated as the product of daily prices with resident-days might not quite represent all operating costs and investment charges. Some nursing homes may receive subsidies from national agencies or from their *departement*, and we cannot consider this in our analysis. For these reasons, inefficiency scores can be over or underestimated.

In addition, we have define quality as staff ratios and staff skills in our model. But there are obviously other quality dimensions that could explain some costs and, therefore, are reflected in inefficiency scores. For example, occasional recruitments of entertainment staff, organization of outings, quality of meals increase costs and are not included in our cost model. So, we do not legititimize a pricing reform that would consist in paying all nursing homes on the basis of costs of the most efficient one, and that would not take into account additional costs required for quality.

Nevertheless, this study demonstrates the need for more financial equity between nursing homes, given unjustified cost heterogeneities. It also shows the need for an increase in staff ratios and a reduction of its variability: 78.9% of surveyed nursing homes have a nurse resident ratio gap less than 1. And we observe that most of these institutions have not low inefficiency terms, since these latter are not well correlated with quality varables. Management issues may certainly be partially responsible for these inefficiencies, but observation of a lower correlation with quality variables when staff ratios are insufficient suggests that a lack of staff may generate additional costs. Indeed, low staff ratios can be a source of inefficiencies associated with high staff absenteeism, increased use of temporary work, more frequent hospitalizations, important turn-over and hence more numerous recruitements. From this perspective, a pricing reform would make possible to improve efficiency if minimum standards of staff ratios were adopted. Without imposing these standards, it is feared that transition to a prospective payment system does not generate inefficiencies reduction but quality reduction.

8 Annexes

Instrumental variable method. We choose to instrument nurse resident ratio gap (N/N^*) and low skilled staff ratio ($LowSkill$) variables by the seniority of staff. We use two-stage least squares, by projecting in a first step N/N^* and $LowSkill$ on non-endogenous variables of the model and instruments ANC and ANC^2. We present equations of instrumentation and estimation results below (Table 12).

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

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

We perform a Sargan test to check the exogeneity of instruments ANC and ANC^2. We add three excluded instruments: percentage of incontinent residents, of residents who made at least one fall in the year, and of those suffering from depression in the nursing home[19]. We accept the null hypothesis of non endogeneity with a p-value of 0.7118.

Fisher tests of overall significance of instruments, carried out from results of each of these equations estimates, allow to reject the null hypothesis of weak instruments: $H_0 : \alpha_{ANC} = \alpha_{ANC^2} = 0$. We get indeed:
-for the first instrumentation equation (projection of N/N^*): $F(2, 1141) = 6.26$, and a p-value of 0.002.
-for the second equation (projection of $LowSkill$): $F(2, 1141) = 9.49$ with a p-value of 0.0001.

Finally, we perform a Hausman test to check endogeneity of N/N^* and $LowSkill$. We reject the null hypothesis of exogeneity of these variables with a p-value equal to 0. To test our hypothesis of non-endogeneity of quality variables beyond a certain threshold (which we can set to N/N^*=1), we conduct two Hausman tests: a test for nursing homes with a nurse resident ratio gap less than 1, and another for nursing homes with $N/N^* \geq 1$. Thus, we confirm our assumption:
-when $N/N^* < 1$: p-value=0;
-when $N/N^* \geq 1$: p-value=0.1570.

[19]These three variables of health status of residents seem to be good quality proxys, unfortunately we only have this information for a small sample of 231 institutions.

Table 12: Instrumentation equations

	N/N^*		$LowSkill$	
Y	-1.021**	(0.325)	1.597	(1.107)
$w_{NA/N}$	-4.106*	(1.928)	-4.606	(6.577)
$w_{SS/N}$	1.502	(2.060)	4.636	(7.027)
$w_{NA/N}^2$	0.731	(1.027)	1.858	(3.504)
$w_{SS/N}^2$	-0.551	(0.450)	-1.680	(1.534)
$Y^2 2$	0.101**	(0.032)	-0.168	(0.110)
$Y.w_{NA/N}$	0.462*	(0.191)	0.670	(0.653)
$Y.w_{SS/N}$	-0.218	(0.203)	-0.513	(0.693)
$w_{NA/N}.w_{SS/N}$	-0.278	(0.680)	2.367	(2.321)
$DRUG$	0.026	(0.019)	-0.222**	(0.066)
$OWN = owner$	ref	ref	ref	ref
$OWN = Pub.rental$	0.007	(0.012)	0.057	(0.041)
$OWN = HLM.rental$	-0.061**	(0.017)	-0.184**	(0.057)
$OWN = PNP.rental$	-0.030	(0.021)	-0.150*	(0.071)
$OWN = PP.rental$	-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)
$MedOption$	-0.122	(0.268)	-0.017	(0.913)
$Y.MedOption$	0.013	(0.026)	0.006	(0.090)
$BUILD$	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)
α_0	5.757**	(1.682)	-6.702	(5.737)

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

Sample of 1171 nursing homes; source EHPA 2007

References

[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] 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.

[4] L. Chritensen, D. Jorgenson, and L. Lau. Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1):28–45, 1973.

[5] T. Coelli. Estimators and hypothesis tests for a stochastic frontier function: A monte-carlo analysis. *Journal of Productivity Analysis*, 6(3):247–268, 1995.

[6] 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.

[7] 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.

[8] M. Farsi, M. Filippini, and M. Kuenzle. Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. *Applied Economics*, 37:2127–2141, 2005.

[9] French Hospital Federation and National Union of Gerontology. Sécurité et qualité dans le secteur personnes âgées: produire des indicateurs de qualité associés à des indicateurs de résultats. 2009.

[10] M. Filippini. Cost and scale efficiency in the nursing home sector : Evidence from Switzerland. *Quaderno*, N.99-02, 1999.

[11] R. Jacobs. Alternative methods to examine hospital efficiency: Data envelopment analysis and stochastic frontier analysis. *Health Care Management Science*, 10(8), 2001.

[12] 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.

[13] 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.

[14] R. Koenker and G. Basset. Regression quantiles. *Econometrica*, 46(1), 1978.

[15] S.C. Kumbhakar and C.A.K. Lovell. *Stochastic frontier analysis*. Cambridge University Press, 2000.

[16] C. Liu, A. Laporte, and B. Ferguson. The quantile regression approach to efficiency measurement: Insights from monte-carlo simulations. *Health Economics*, 17, 2008.

[17] V. Létard, H. Flandre, and S. Lepeltier. La France et les Français face à la canicule : les leçons d'une crise. 2004.

[18] 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.

[19] W. Meeusen and J. Van Den Broeck. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2), 1977.

[20] J.P. Newhouse. Toward a theory of nonprofit institutions : an economic model of a hospital. *American Economic Review*, 60(1):64–74, 1970.

[21] J.P. Newhouse. Frontier estimation: how useful a tool for health economics? *Journal of Health Economics*, 13(3):317–322, 1994.

[22] M. Pitt and L. Lee. The measurement and sources of technical inefficiency in the industrian weaving industry. *Journal of Development Economics*, 9(1), 1981.

[23] E. Ratte and D. Imbaud. Groupe de réflexion sur la dépendance: « accueil et accompagnement des personnes âgées en perte d'autonomie ». 2011.

[24] W. Réjault. *Maman, est-ce que ta chambre te plaît? Survivre en maison de retraite.* Broché, 2009.

[25] P. Schmidt and R. Sickles. Production frontier and panel data. *Journal of Business and Economic Statistics*, 2:367–374, 1984.

[26] 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.

[27] D. Vitaliano and M. Toren. Cost and efficiency in nursing homes: a stochastic frontier approach. *Journal of Health Economics*, 13(3), 1994.

[28] 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.